



Statistical Process Control Analysis Unraveled

By Peter Baxter, Distributive Management

1. Introduction

The idea of using statistical process control (SPC) immediately strikes terror into the hearts and minds of many managers. Fearful thoughts of long equations, foreign acronyms and terms, and having to re-read the instruction manual for your pocket calculator are evoked. Of course, SPC is implemented by the analysts and “metricians”, and not actually used by program managers and executives. In this section, I will describe why SPC techniques were initially developed, and how they were meant to be used in day-to-day management.

The anticipated “challenges” of SPC implementation deflect attention away from the basic problem that it addresses: Given data with some variation, how can you tell if the variation is due to normal process fluctuations, or due to a serious flaw in the process design or your organization? I often say that SPC helps a manager determine if the changes in their data are significant.

Consider a measure graph in your organization. Each month, you will collect some number of data points, and each month the graph depicting those points changes. The graph is never (or rather “almost never”) a straight line. Sometimes the new values are all higher than the previous month, and sometimes they are all lower, and sometimes the points are mixed. SPC helps managers answer the question: As a manager, how different from previous months does the data have to be before I act? You are interested in acting when the data is significantly different from previous months. But what is “significant”?

Dr. Walter Shewhart (pronounced “shoe heart”) is considered the creator of what is now called SPC. In 1924, Dr. Shewhart devised a framework for the first application of the statistical method to the problem of quality control. Dr. Shewhart wrote a note to R.L. Jones, responding to his request for some type of inspection report, which “might be modified from time to time, in order to give a glance at the greatest amount of accurate information.” Dr. Shewhart attached a sample chart “designed to indicate whether or not the observed variations in the percent of defective apparatus of a given type are significant; that is, to indicate whether or not the product is satisfactory.”

From this basic idea, mathematicians and statisticians have constructed the techniques that we use today for SPC. The following is an excerpt from the book “Dr. Deming: The American Who Taught The Japanese About Quality” (by Rafeal Aguayo):

When management looks at information on output or performance, it is confronted with a bunch of numbers. The usual assumption is that each number is due to one specific cause, such as the effort or lack of effort of an individual. But in a system like the bead factory, all the variation



is due to chance, which means it is caused by the system and not the workers. Management is responsible for the system. Why then blame the workers?

Summary: How do we determine if the variation in a system is due to chance or individual causes?

By using SPC, managers will better understand the behavior of the processes that they are managing. This will allow them to act on the variations of those processes that are not noise. Over time, the effective manager will reduce the variation in the process. In using SPC this way, the manager must address aspects of quality, technology, resources and process in an integrated way.

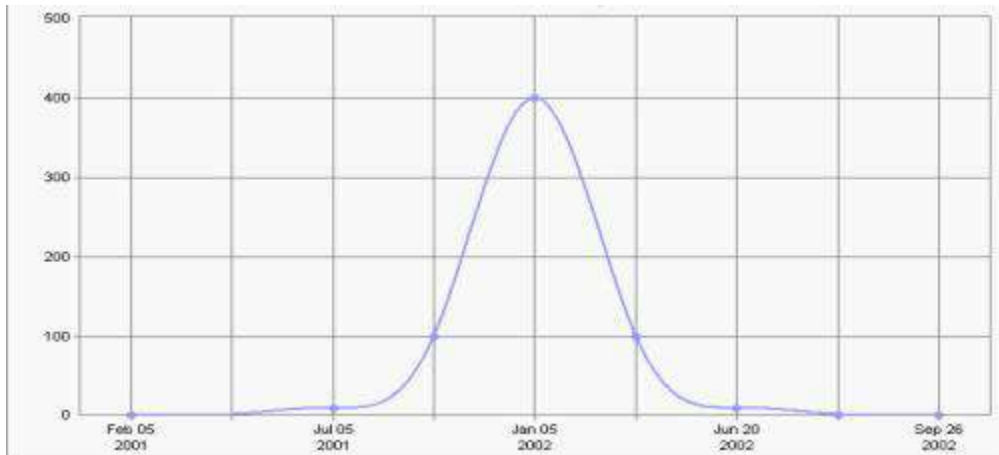
2. How Do We Determine Significant?

To start understanding how SPC calculates a number of significance, let's work through techniques for examining the variation in data over time. We are interested in establishing a **lower and upper limit** that, when a data point goes under or over it (respectively), indicates that we have a significant change.

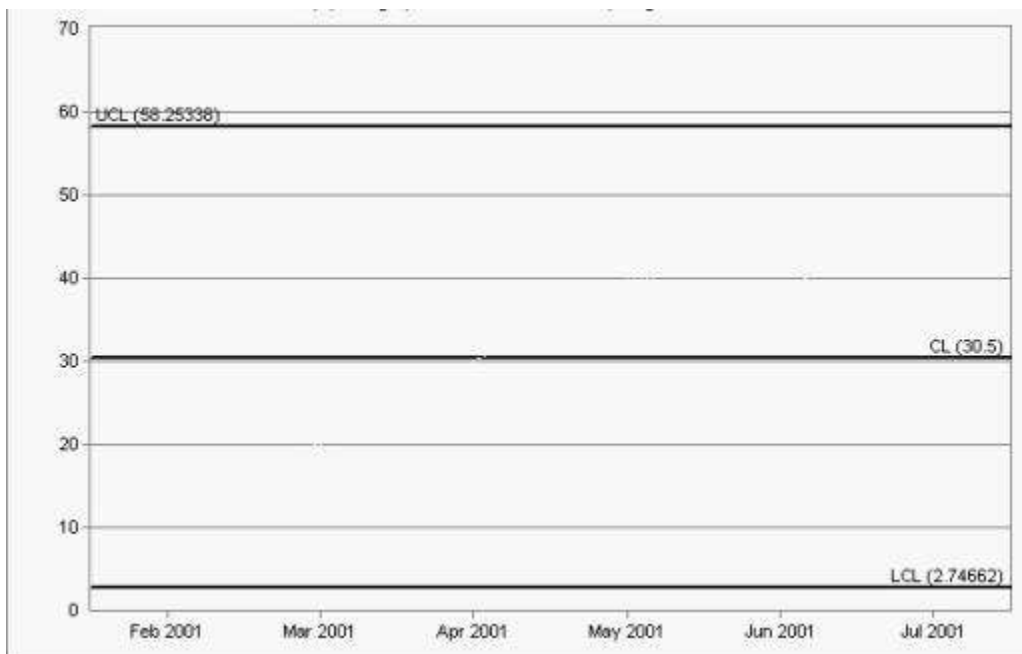
The easiest method would be to set constant "limits", such that when a data point exceeded the limit it was considered significant. For example, for a defect measure, the upper limit might be 1000 and a lower limit might be 0 (we don't expect to have fewer than 0 defects!). This mechanism is simple, but has a few challenges. First, a constant limit must be set for every measure in your measurement program, as each measure has a different range. Second, the limit has to be determined by examining historical data, but in such a way as to take into account the different size/scope of each project – a challenging task for any good metrician. Finally, and more importantly, this solution really sidesteps the question of significance by essentially saying, "Significant is whatever you say it is." Let's keep looking for a better technique.

Another way to assess "significance" would be to use the range of data values in your historic data. For example, if you had data with a range of values between 0 and 200 then a point outside 200 could be considered significantly different. While this solution is good, we can improve upon it by considering where most of the data values fall. We use the **average** value of the data set to indicate where most of the values fall. (Side note: technically the mode is the single point that occurs most frequently in the data set, but the average more commonly represents the center of a data set.) For example, if the sample data set has an average of 25, we should create limits (upper and lower) that are centered around 25.

This idea of "centering" the limits around an average is one of the assumptions of SPC. In particular, when the upper and lower limits are the same "distance" from the average, we are assuming we have what is termed a normal distribution. In a **normal distribution**, the chance that data falls above the average is the same as the chance that it will fall below the average. The most common representation of a normal distribution is the bell curve shown in the figure below.



Once we calculate the average of our data set, we still need to find a reasonable method for determining the upper and lower limits. To make this problem more approachable, let's agree that we will use the average as our center line, and what we are now looking for is a value that can be subtracted from the center line to get the lower limit, and added to the center line to get the upper limit. A chart showing the center line and lower and upper limits is called a control chart. A sample control chart is shown in the figure below.



We have narrowed our problem down to the equation for the value of one limit (which we will subtract from the center line for the lower limit, and add to the center line for the upper limit). To calculate this limit, we want to use an expression of **variation**. One method of determining the limit would be to look at the variation of each data point in the data set, by subtracting the data point's value from the center line. A sample of this is shown in the following table.



Distributive Management

White Paper – Statistical Process Control Analysis Unraveled

Period Num.	1	2	3	4	5	6	7	8	9	10
Data Value	5	25	30	20	10	25	20	25	5	35
Average (center line) 20										
Variation from average	-15	5	10	0	-10	5	0	5	-15	15

From the above table, the variation from average is useful in seeing that the values in periods 1, 9 and 10 have the largest variation (with a difference of 15 or -15). However, we still are not sure if this variation is significant or not. Referring to the normal distribution curve above, what we would like to do is draw vertical lines through the curve on the left and right of the average, or center line, that we could use as our control limits.

Walter Shewhart developed the notion of control limits and their relative position in 1924. He began with the variance of the data set, which is:

$$\text{Variance} = \text{Sum of the squared differences from the average} / \text{number of points} - 1$$

By using the sum of the squares, you are assured that your variance is always positive, and that differences below the average are not subtracted from differences above the average. Shewhart used a unit of **standard deviation** instead of the variation. The formula for standard deviation is:

$$\text{Standard Deviation} = \text{Square Root of the Variance}$$

One interesting effect of using the standard deviation is that it accounts for the differing range of values that data sets have. It would be nice if our control limits were wider for more naturally variable data, and that the “width” of control limits reflected this. For example, ten points with an average, or **mean**, of 25 where the values are between 24 and 27 is going to have a different acceptable variation than 10 points with an average, or mean, of 25 where the values are between 0 and 50. The standard deviation does this by setting a “**unit of deviation**” (i.e. one sigma) based on the average difference between the mean and all the data points.

If we revisit the normal distribution curve, you will find that a one sigma line drawn as control limits would include 68% of all the data in the normal distribution. If you use two sigma to define the control limits, then 95.4% of the values in the normal distribution will fall into that range. Finally, if you use three sigma as a control limit, then 99.73 percent of your data will fall between the average plus or minus three-sigma.

For example, if we have a data set that has a mean of 25 and a standard deviation of 3, then we would expect 68% of all values to fall between 22 and 28 (i.e. 25 +/- 3). Similarly, we would expect 99.73% of all values to fall between 16 and 34 (i.e. 25 +/- 9).

Shewhart decided to use three sigma (or three times the standard deviation) as the basis for the control chart. By using three sigma as the control limits, he is saying that 99.73% of our data should normally fall inside that range and that anything outside that range, most likely, represents a condition that is significant and that requires action or attention.



In our original challenge, we wanted to find a way to determine whether a change was significant or not. To start, we defined a “significant” region using upper and lower control limits (which are symmetrical around the mean, meaning the same distance up to the upper limit as down to the lower limit!) Then, we looked at how the difference, variance and then standard deviation can be used to build an equation for our control limits. Finally, we found that the standard deviation can be used (assuming a normal distribution) to represent the likelihood that data values will fall within a particular range. This provides a solution to our problem of “significant”.

3. More About Control Charts

The chart that is used to display the upper and lower control limits, the centerline, and the data points is called a control chart. Recall from the last section, that the centerline is the average of the data set, the upper limit is the centerline plus three standard deviations, and the lower limit is the centerline minus three standard deviations.

The control chart, as presented so far, is susceptible to two types of errors: 1) reporting instability when none actually exists (called a false alarm), and 2) failing to detect instability when one exists. There is a probability that a point that falls outside the upper or lower control limit does not represent an unstable process. A point such as this causes what is called a false alarm. A false alarm is a result of the fact that we are basing our control limits on probability, and in some cases this assumption is incorrect.

To counter false alarms, the control limits are set at three sigma, which encompasses almost 98% of the data that should occur. Any data outside this range is probably representative of instability. Consider for example if the limits were set at two sigma, then a larger number of data points could potentially indicate that the process is unstable. To keep you and other managers from investigating assignable or random causes of process variations, the control limits have been pre-determined at three sigma.

The other case we are interested is failing to detect instability when it exists. For example, suppose every data point was between the control limits, but increasing from the centerline to the lower control limit. In essence, the data is indicating that there is a clear trend downward and there is no apparent natural variation in the process (or if it exists, then it is negligible). To counteract the failure to detect instability, a number of statistical tests have been created that are now used (along with the 3 sigma test) to detect process instability.

The complete set of stability tests are:

- A single point outside three sigma
- At least 2 of 3 successive data points outside two sigma
- At least 4 of 5 successive data points outside one sigma
- At least 8 successive data points on the same side of the centerline



Using tests 2 through 4 above, you can accurately detect data that is not outside the control limits but still representative of process instability. These tests are included in most statistical process tools. In fact, these tests are recommended by leading maturity models when quantitative methods are a required process element.

4. The X, Moving R Control Chart

The control chart is meant to be a “current state” evaluation of process that is based on the most recent and relevant process performance. I emphasize current state because “more data is better” does not apply when dealing with control charts. In particular, the control chart should contain points that were generated by a process that is the same as the one being measured.

For example, if you have changed from a functional language to an object-oriented one, or started using use cases and OO Analysis instead of functional decomposition, you would want to make sure your control charts only contain data from processes that matches how your process currently perform.

The most common type of control chart used in statistical process control is the X, Moving Range, or X, Moving R chart or XmR chart. Previous sections described how the standard deviation (sigma) and the control chart can be used to determine when data is significantly different. With that background, we now move on to the type of graph (or types of graphs) used to visually display the control chart.

First, the control chart is probably the most common graphic element of SPC. However, it is not the only one. Other techniques used in SPC include these:

- Fishbone diagrams
- Scatter diagrams
- Histograms
- Pareto analysis
- Run charts
- Control and range charts

The primary purpose of these techniques is to aid in understanding the variation of one of more of your processes. For example, pareto analysis could aid in understanding the relative contribution of a number of factors to an unacceptable process variation. You can consult any number of books on the other techniques mentioned above.

You may have thought that there was only one kind of control chart – but that would be just too easy wouldn't it? In fact, there are a number of control chart types. Control chart types include the X, Moving Range chart, the X Bar R chart, the X Bar S chart, the NP chart, the C chart, the P Chart, the U chart, and the Z chart. I will go over a method for selecting which one to use in a later section. For now, know that there are more chart types than just the X, Moving Range chart.

Our goal, for this section, will be to understand the variation of code during the coding phase of a software development project. The following table contains a set of sample data, for the



Distributive Management

White Paper – Statistical Process Control Analysis Unraveled

KSLOCs (thousand source lines of code) generated each month for one software module. The last row in the table shows the difference between successive measures. This difference is called the “range”.

	Jan	Feb	March	April	May	June
Module	106	72	44	90	51	83
Range	n. a.	34	28	46	39	32

The average of the measurements, called “X bar” since the measurement values are plotted on the X-axis of a graph, is:

$$\begin{aligned} X \text{ Bar} &= (106 + 72 + 44 + 90 + 51 + 83) / 6 \\ &= 446 / 6 \\ &= 74.3 \end{aligned}$$

The average of the moving range, called “mR bar”, is:

$$\begin{aligned} mR \text{ Bar} &= (34 + 28 + 46 + 39 + 32) / (6 - 1) \\ &= 179 / 5 \\ &= 35.8 \end{aligned}$$

To put this data onto a control chart, we need to generate the upper and lower control limits (UCL and LCL, respectively). One technique for calculating the control limits is to use an equation that approximates standard deviation:

$$\begin{aligned} UCL &= X \text{ Bar} + (2.66 * mR \text{ Bar}) \\ &= 74.3 + (2.66 * 35.8) \\ &= 169.5 \end{aligned}$$

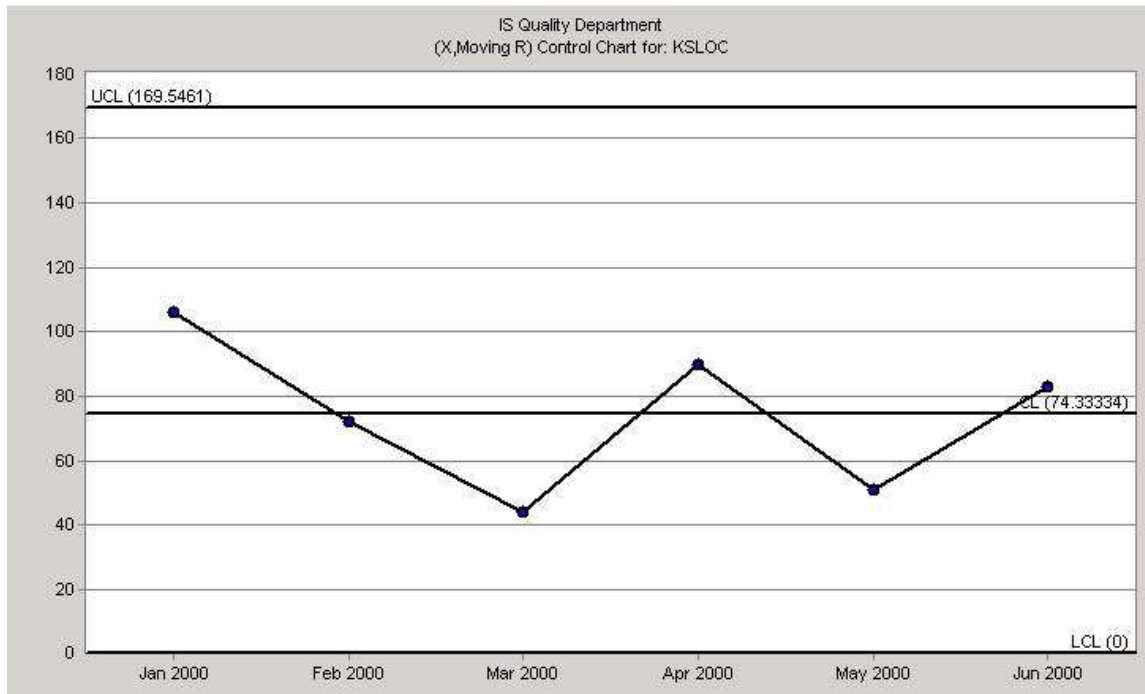
$$\begin{aligned} LCL &= X \text{ Bar} - (2.66 * mR \text{ Bar}) \\ &= 74.3 - (2.66 * 35.8) \\ &= 0 \quad (\text{lower limit cannot be less than } 0) \end{aligned}$$



Distributive Management

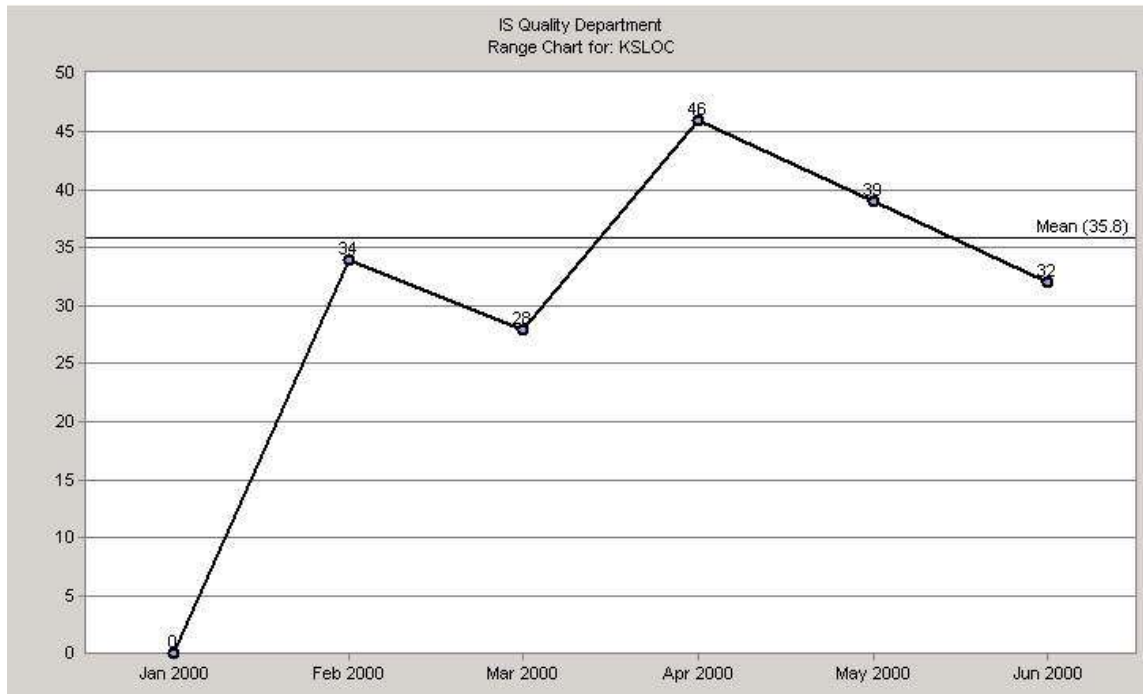
White Paper – Statistical Process Control Analysis Unraveled

With the initial measurement and the calculated values (from above), we can plot our X, Moving R control chart, as shown in the following diagram.





A range chart is commonly shown along with the control chart, and, for our sample data set, would look like the following.



So what does the X, Moving Range chart help us do? First, the control chart allows you to visually review the software development variation. Because the control limits are drawn on the chart, you can see whether any single point exceeds three sigma or not. Finally, you can see the overall trend of your process, and hopefully there is decreasing variation!

Over time, your organization may find and correct any number of activities that (previously) resulted in instability. This represents tuning a process, not necessarily a drastic change as described in the previous paragraph. For example, you change the way requirements are base lined to make improvements to software configuration management. This might have been done, for example, to address a “spiking” of requirement defects once they were formally checked into CM.

To allow the control charts to accommodate minor process change, you should “discard” the control chart points where previous “spikes” occurred. Discarding points will have the effect of “tightening” your control limits. A point outside three sigma is often called an outlier. So, once the outlier is discarded, the variation of the data set will be less, causing the single sigma and three sigma values to also be less. This in effect says, “Now that I have fixed the cause of my outlier, I expect my process to exhibit less variation.”

5. The X Bar R Chart



The name X Bar R is an abbreviation for the X-Bar and R chart, where the R stands for Range. The X Bar R chart is used to display variation occurring in a set of data points. (Remember from last time that the X Moving R displayed variation of essentially one measurement per period.) When you have more than one measurement per period, you use the X Bar R chart. The X Bar R chart provides insight into the process that generates the data set, not the items in the data set.

For example, if you apply the X Bar R chart to four software modules each week, you are studying the variation of the software development process that produces the software. You are not learning a great deal about the quality, complexity or completeness of the software modules themselves.

An Example of the X Bar R Chart

As an example, suppose you measure those four software module sizes each month for 6 months. Our example collected data looks like the following:

S/W Module	Month						Subtotals
	1	2	3	4	5	6	
Module A	104	89	124	139	168	168	792
Module B	12	21	41	57	62	68	261
Module C	156	26	103	76	51	112	524
Module D	68	77	83	92	108	119	547
Subtotals	340	213	351	364	389	467	
Average	85	53.25	87.75	91	97.25	116.75	
Range	144	68	83	82	117	100	

The size of the software for the four modules is shown for months 1 through 6. Under each month, is the subtotal of all software created, the average of the subtotal for that month, and the range (between the smallest and largest value).

To calculate the X Bar R value, we take the average of the monthly averages, or:

$$\begin{aligned} \text{X Bar ("Average X")} &= (85 + 53.25 + 87.75 + 91 + 97.25 + 116.75) / 6 \\ &= 88.5 \end{aligned}$$

$$\begin{aligned} \text{R Bar ("Average Range")} &= (144 + 68 + 83 + 82 + 117 + 100) / 6 \\ &= 99 \end{aligned}$$

The Center Line, or CL, is then:

$$\begin{aligned} \text{CL (Center Line)} &= \text{X Bar} \\ &= 88.5 \end{aligned}$$

To calculate the upper and lower control limits, we use a standard equation:

$$\begin{aligned} \text{UCL} &= \text{CL} + (3 * \text{sigma}) \\ \text{LCL} &= \text{CL} - (3 * \text{sigma}) \end{aligned}$$



For our sample data, the one sigma value is:

$$\text{Sigma} = 18.869$$

So, our values for UCL and LCL are:

$$\begin{aligned} \text{UCL} &= 88.5 + (3 * \text{sigma}) \\ &= 88.5 + 56.5072 \\ &= 145.1072 \end{aligned}$$

$$\begin{aligned} \text{LCL} &= 88.5 - (3 * \text{sigma}) \\ &= 88.5 - 56.5072 \\ &= 31.8928 \end{aligned}$$

Using Approximated UCL and LCL

In calculating the control limits, I used the equation for standard deviation and not an approximation. In many books on SPC, you will see an equation for an approximation of the UCL and LCL where the approximation depends on the number of elements in the subgroup. The equation for the approximations is given as:

$$\begin{aligned} \text{Approximated LCL} &= \bar{X} - (A2 * \bar{R}) \\ \text{Approximated UCL} &= \bar{X} + (A2 * \bar{R}) \end{aligned}$$

Where A2 is a constant based on the subgroup size

The approximated values for our data set are:

$$\begin{aligned} \text{UCL} &= 88.5 + (0.729 * 99) \\ &= 160.67 \end{aligned}$$

$$\begin{aligned} \text{LCL} &= 88.5 - (0.729 * 99) \\ &= 16.329 \end{aligned}$$

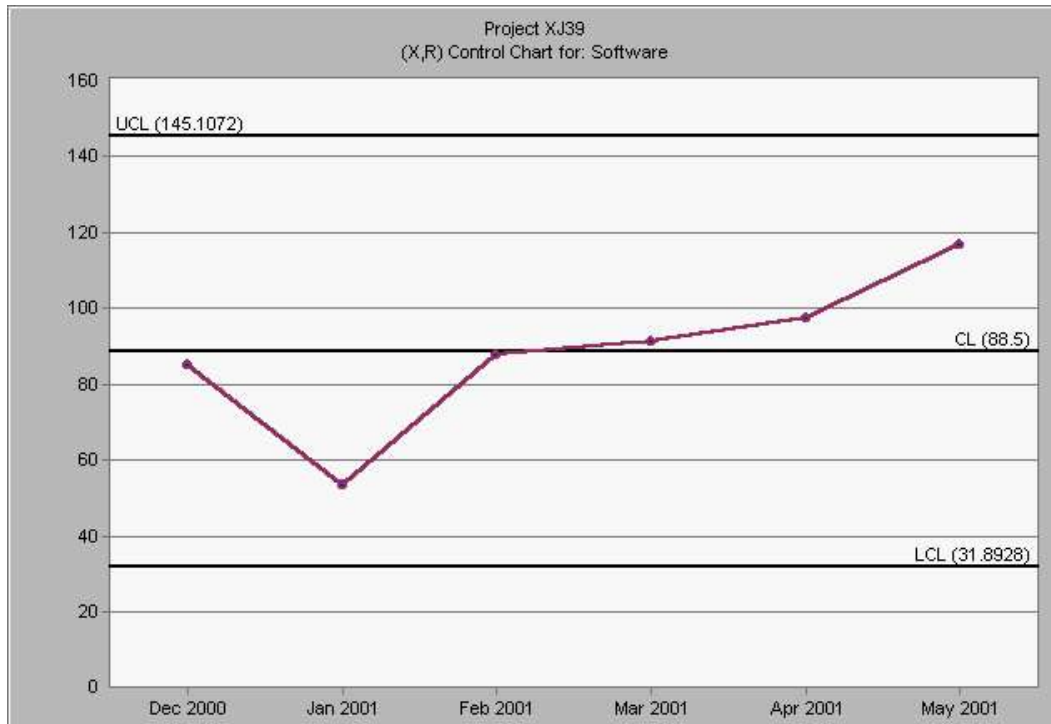
The effect of this approximation is that the control limits are wider than the data actually warrants. The approximated UCL is about 15 greater than the UCL, and the approximated LCL is about 15 lower than the LCL. The overall effect is that the “range” of the control limits went from 114 (145 – 31) to 144 (160 – 16). Or, the control limits are about 30% larger when using the approximations.

In my opinion you are safer using the calculated standard deviation and using the “true” equations for LCL and UCL, rather than using the approximations. When these equations had to be done by hand, the approximations could save you valuable time if you had to do 20 of them a day. But, with the availability of serious desktop computing power and SPC/analysis software, there is no gain in using approximations.



Assessing Stability

Using our sample data, an X Bar R chart, with CL, UCL and LCL would then look like the following:



In assessing stability, you want to examine 4 primary rules. There are others, but these are the most common. These stability rules are:

1. A single point outside three sigma
2. At least two out of three successive values outside two sigma
3. At least four of five successive points outside one sigma
4. At least eight successive points on the same side of the centerline

From the graph, you can see that none of the values exceed the UCL or LCL, so the data does not violate the first rule. You can use the single sigma value (above) to check stability rules 2 through 4.

Determining Subgroups

I want to revisit a subtle assumption within the sample data above, namely the subgroups that were used as the basis for the equations. In our case, we have a subgroup consisting of 4 data points at each measurement period. You must ensure that the subgroup members are indeed representative of the same basic business process. For example, data from a development



activity and data from a maintenance activity should not be used in the same subgroup unless the process (policy, procedure, tools, etc) are identical.

In essence, when you put several data points in a subgroup, you are assuming that the same or equivalent process is being used to generate the products being measured.

Summary

We looked at a control chart that can be used when there is a set of data values at each measurement period. The X Bar R chart uses the average of the data values at each period, along with the average of the range of values, to describe stability in a standard control chart.

6. Using SPC for Analyzing Defect Data

Statistical process control (SPC) can be used in assessing systems or software defects. To examine the variation that occurs in defect data, the C chart and U chart are used. The U chart has become a common technique for assessing the significance of defect counts and the stability of activities that find defects.

A defect is an item that has one or more nonconforming quality attributes. For example, a requirement that is ambiguous or a logical software statement that has a complexity level that exceeds an organizational limit. Strictly speaking, you cannot have a defect unless you have a standard that describes each quality attribute. Defect data is generated as a result of an inspection or a peer review of a life cycle product such as a document or software module. In many cases, the work product is sampled, such that only part of the work product is checked for defects, rather than the entire document.

Description of the U chart

To calculate the control limits for defects, we need to collect and then analyze defect data. Unlike periodic sampling where we measure a software size or page count each month, defects are found as a result of a-periodic activities. That is, you find defects as inspections and peer reviews are held, and they occur as needed. Unlike other control charts (and measurement graphs), you will not use dates across the horizontal, the X axis. Instead, you will use a sample set number, event number, or inspection identifier.

The discriminator between whether you would use the U chart or C chart is the size of the sample. If the sample size is constant, then you use a C chart. If the sample size changes (or could change) for each sample, then you use the U chart. The C chart monitors the total number of defects per unit, for a constant sample size. The U chart is used to control the average number of defects, and is used when the sample size is not constant.

As an example, consider the inspection of systems documentation in a systems engineering program. The program will generate several documents, each having a different page count. As the documents move through the systems engineering activities, an inspection is performed on them to identify defects as early in the life cycle as possible (most cost effective). Inspection



data will be generated on each document, at sporadic intervals, with the number of defects changing as (initially) more draft content is generated, and then (eventually) as the document is finalized and approved. A manager attempting to control and monitor the defect counts would need to use the U chart to make sense of the data being provided.

Sample U chart Data

In our example scenario, assume that the following table contains defect data for one of the systems documents being produced. Each row in the table represents an inspection (or peer review). The columns in the table contain a sequential inspection id, the number of pages inspected, the number of defects found and the defects per page.

<i>Event No.</i>	<i>Sample Size (pages)</i>	<i>Defects</i>	<i>Defects/Page</i>
1	15	3	.200
2	11	4	.364
3	11	5	.455
4	15	3	.200
5	15	6	.400
6	11	8	.727
7	15	10	.667
8	11	3	.273
9	11	2	.182
10	15	3	.200
	130 total pages	47 total defects	

To start our analysis, we need to draw a centerline (CL) that is the average number of defects:

$$\begin{aligned} \text{Centerline} &= \text{sum of defects} / \text{sum of pages} \\ &= 47 / 130 \\ &= .3615 \end{aligned}$$

From the previous sections, we want to calculate the upper control limit (UCL) and lower control limit (LCL) such that we have an indication that a change in data is significant. Unlike the X Bar R and X Moving R chart (that has a constant LCL and UCL), the LCL and UCL for the U chart go up and down at each event. (When the LCL and UCL are plotted using a square step function, they look like a city skyline.)

The equations for the LCL and UCL and the lower control limit are given by:

$$\begin{aligned} \text{UCL (i)} &= \text{CL} + (3 * \text{Square Root (CL / Sample Size (i))}) \\ \text{LCL (i)} &= \text{CL} - (3 * \text{Square Root (CL / Sample Size (i))}) \end{aligned}$$

Where (i) represents the event number

For example, the LCL and UCL for the first period are calculated as shown below:

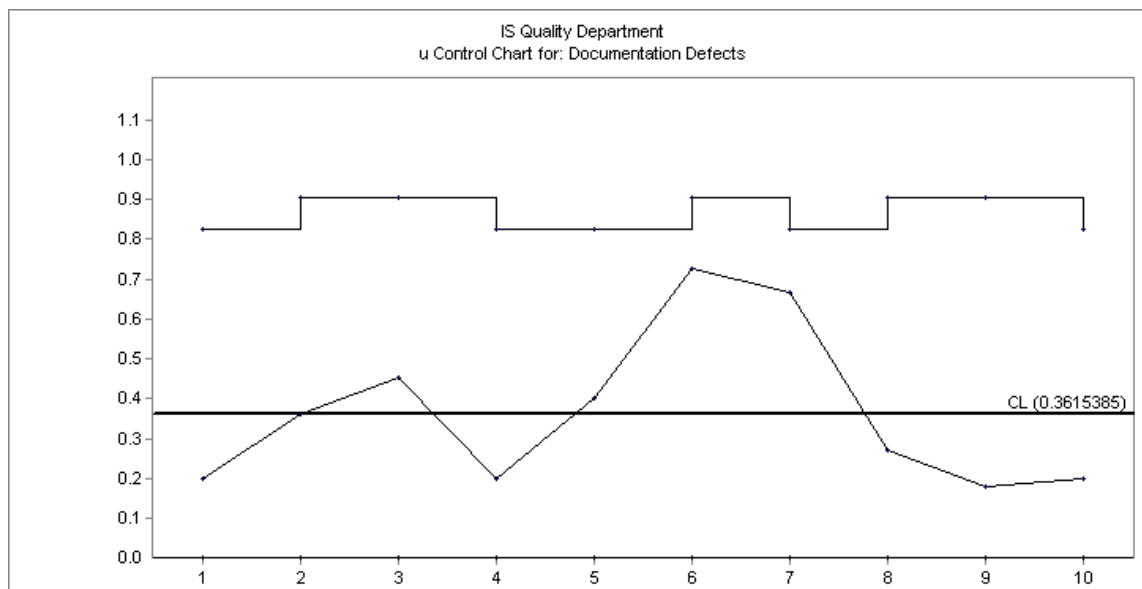


$$\begin{aligned} \text{UCL (1)} &= \text{CL} + (3 * \text{Square Root (CL / Sample Size (i)))} \\ &= .3615 + (3 * \text{Square Root (.3615 / 15)}) \\ &= .3615 + (3 * \text{Square Root (0.0241)}) \\ &= .3615 + (3 * .15524) \\ &= .3615 + 0.46572 \\ &= 0.82722 \end{aligned}$$

$$\begin{aligned} \text{LCL (i)} &= \text{CL} - (3 * \text{Square Root (CL / Sample Size (i)))} \\ &= .3615 - (3 * \text{Square Root (.3615 / 15)}) \\ &= .3615 - .46572 \\ &= -0.10422 \text{ since we cannot have less than 0 defects} \\ &= 0 \end{aligned}$$

In the above calculation of the LCL, we do not let the LCL go below zero. This is because we should not ever have the case where we find less than 0 defects. So, when the number for LCL is negative, we set it to 0.

The control chart in the figure below shows the U chart for the data in our example.



The centerline, labeled CL, is a constant value plotted for all samples. The UCL is a stepped function that changes for each event. The LCL value for each event is zero.

Note that by examining the control chart, a manager could quickly see that the average number of defects seems to be in control. You would apply the tests for stability covered in a previous section, namely:

- One point outside three sigma
- Two out of three successive points outside two sigma
- Four out of five successive points outside one sigma



- Seven of eight successive points on the same side of the centerline

Summary

The U chart technique described in this section is very effective at analyzing the defect data generated during inspections and peer review activities. The U chart provides technique for analyzing defects where 1) data is not generated periodically, and 2) where the sample size is not constant. Unlike other control charts, the UCL and LCL of the U chart are calculated for each event. Additionally, the LCL is bounded at zero, meaning that the LCL cannot be negative, to reflect the fact that you cannot find fewer than 0 defects.

7. The Z chart

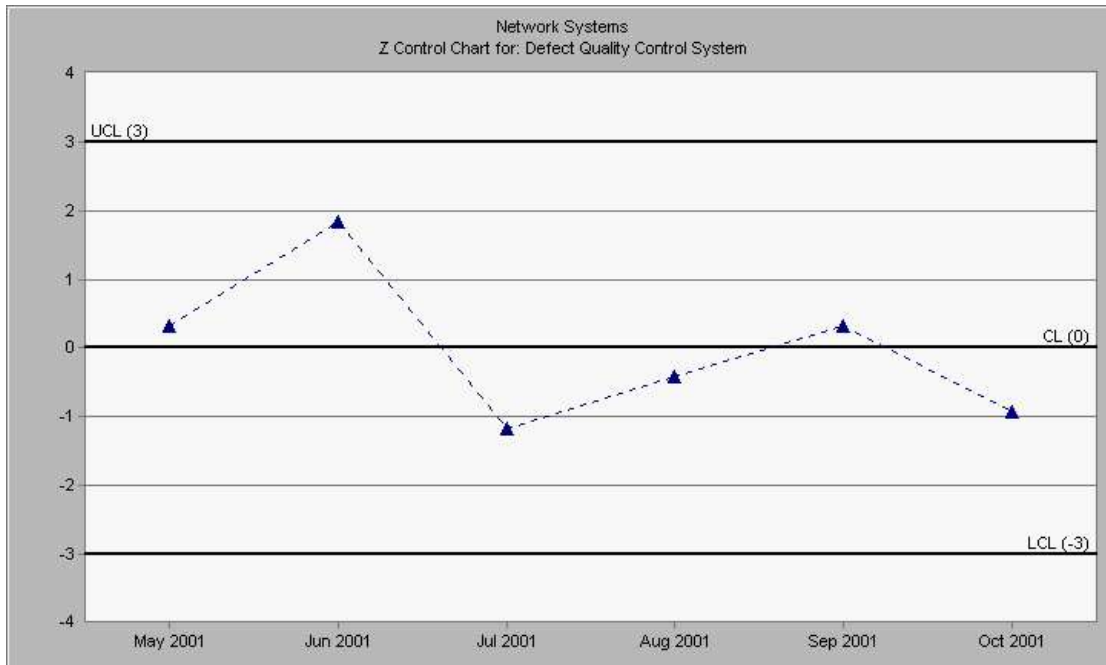
In the last section, the SPC technique for using a U chart was presented as a means for making quantitative decisions about defect data. In this section, another technique, known as the Z chart, provides a slightly different mechanism for evaluating measures that have variable control limits.

Remember that the U chart was used to evaluate data that is typically a-periodic (that is, not on a regular schedule) and is collected from a variable sample size. An example within systems engineering might be a set of systems requirement specification peer reviews where the number of pages inspected changes each time, and the meetings are held as the document evolves.

Unlike other control charts, the control limits for the U chart are calculated for each measurement sample – when plotted as a stepped straight line, the control limit resembles a city skyline, rather than a straight line of the XmR or X Bar R chart. In the U chart, the scale along the Y-axis is the defects or observations per unit being sampled.

The Z chart

The Z chart is a plot of the individual sigma values for each sample in the set. A typical Z chart is shown in the following figure.



From the above chart, the scale on the Y-axis is in sigma units. The “highest” scale unit is three sigma, allowing the quick identification of data samples that are outside of three sigma. The value for each point in the graph is calculated by:

$$Z(i) = (u(i) - \bar{u}) / \sigma_u(i)$$

Where $u(i)$ is the u chart value

\bar{u} is the average of u for the set

$\sigma_u(i)$ is one standard deviation of the “ith” sample

The equation above is basically stating that the z value is the difference between the sample value (given by “ $u(i)$ ”) and the mean converted to the number of standard deviations (by dividing by the ith points sigma value).

If you compared a U chart to a Z chart, you would see that the “shape” of the u and z charts is identical, only the scale is different. The value of the Z chart (over the U) is that you can quickly see points that fall outside the upper or lower sigma limits, as well as those that violate any of the other stability checks.

The Z chart descriptions sometimes refer to each data point as the “z score” of the sample. This score takes into account the centerline, since we are assuming that 0 sigma is at the center of our distribution. In addition, the z score takes into account the dispersion of the data, since it is directly based on the variance from the mean. The Z chart then depicts how far from the mean each sample score is from the centerline.

Summary



Similar to the U chart, the Z chart technique described in this section is very effective at displaying the variability of data generated during inspections and peer review activities. The Z chart provides technique for analyzing defects where 1) data is not generated periodically, and 2) where the sample size is not constant. Unlike the U chart, where the LCL and UCL values are calculated for each sample, the Z chart UCL and LCL are calculated as an average across all data sets. Additionally, the Z chart LCL is not “bounded” at 0 sigma (like the U chart). The Z chart is not the most popular SPC in the arsenal but is effective at presenting the relative variance of your data.

8. Selecting the Correct Control Chart

Background on Variable and Attribute Data

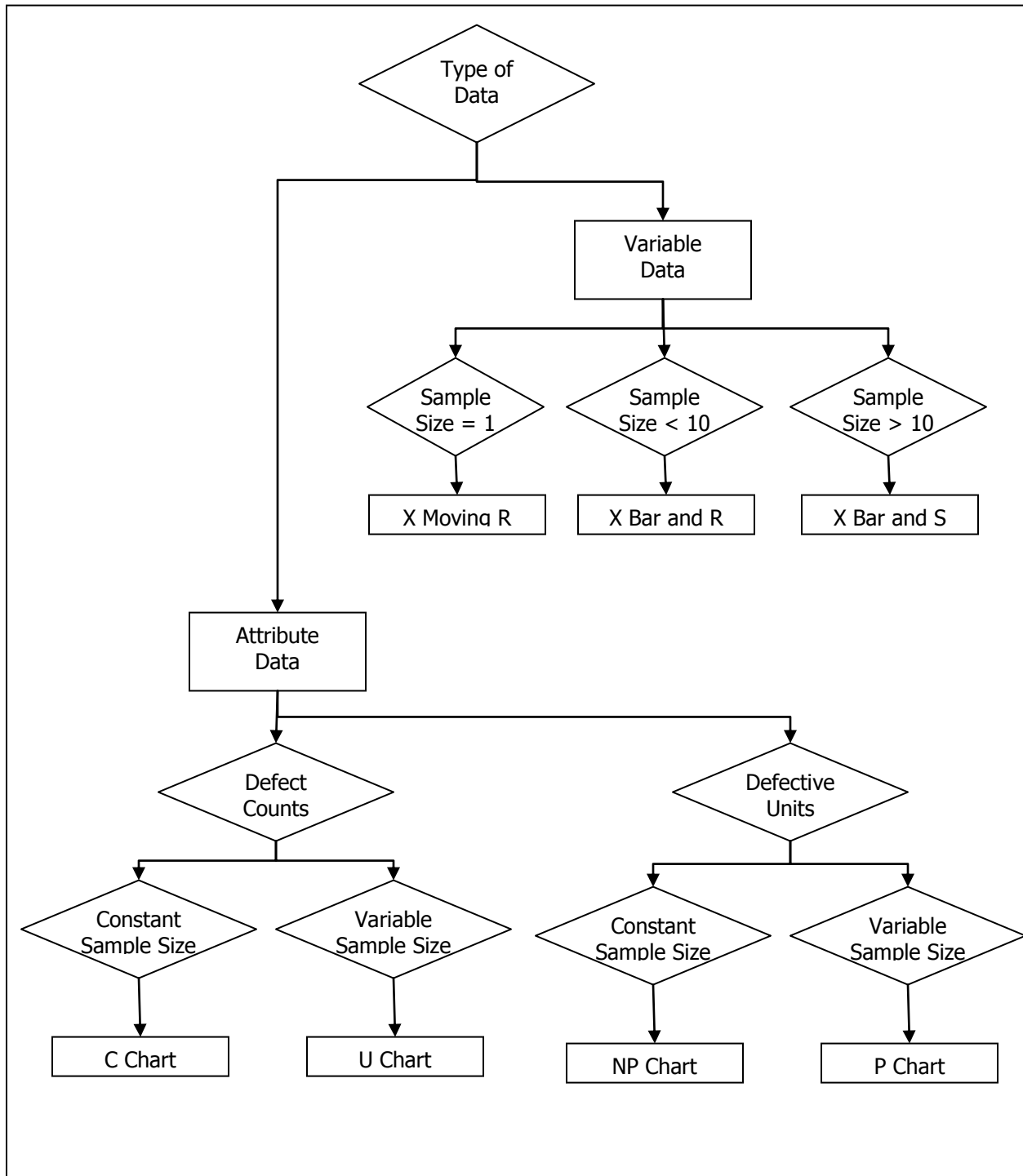
The first step in the selection flow chart is to characterize your data as one of two types: either variable or attribute data. Variable data is data that can be measured and plotted on a continuous scale. Attribute data is data that is measured from discrete events. As the terms are not as descriptive as they could be, an explanation and examples of each follows.

Variable data encompasses measurements of a process that is occurring nearly all the time. A common example of variable data is temperature, where you can walk outside and take the air temperature any time. In other words, temperature is a constantly **variable** quantity that can always be measured. With respect to systems engineering, variable data could be the number of requirements at any time during system analysis phase. That is, at any time during the phase, you would expect activity to be creating, modifying and deleting the requirements of your system.

Attribute data is used to describe data that happens at discrete events during the project life-cycle. That is, when you examine a single discrete event, the measurement data is a count of one of the **attributes** of the event. You are not counting the number of events, instead you are counting something, called an attribute, about that event. In software engineering, defect counts produced during inspections are considered attribute data. The number of defects found is an attribute of the inspection event. Further, while it may seem like inspections happen all the time (or maybe this is a programmer myth!), they are held sporadically as work products are produced.

The Selection Flowchart

With the classification of data as either variable or attribute, the selection flow chart is shown on the next page.



In the above figure, the “Defect Counts” and “Defective Units” address common types of attribute data found in software and systems engineering. “Defect Counts” represent a count of



the number defects found during an inspection, peer review or other defect discovery activity. “Defective Units” represent the number of items produced that failed to meet acceptance criteria. In the case where, your organization performs both type of quality activities (i.e. inspection and acceptance tests) you may need to use both parts of the decision tree for attribute data.

Conclusion

The selection flow chart can be used as a starting point in selecting which type of control chart you should use for your process data. To make sure you are aware of the tradeoffs and all the factors involved in selecting a control chart, you should consult one of the references listed at the end of this section or a measurement analyst who has experience with statistical process control.

Additional Reading on These Topics

Measuring the Software Process: Statistical Process Control for Software Process Improvement by William A. Florac and Anita D. Carleton (SEI Series of Addison Wesley)

Statistical Methods for Software Quality: Using Metrics for Process Improvement by Adrian Burr and Mal Owens (Thomson Computer Press/ITP)

Dr. Deming: The American Who Taught The Japanese About Quality by Rafeal Aguayo (Fireside Book by Simon and Schuster)

Evans, James R. Statistical Process Control for Quality Improvement: A Training Guide to Learning SPC. Englewood Cliffs, New Jersey: Prentice Hall, 1991.

Florac, William A., and Carleton, Anita D. Measuring the Software Process: Statistical Process Control for Software Process Improvement. Addison-Wesley, 1999.