

Making statistical process control user friendly

Charles E. Geisel

Statistical process control, although based on sophisticated mathematics, can be understood and used by anyone who can use grade-school arithmetic.

Statistical process control. The very name can evoke fear into the minds of some workers and management people, particularly those who do not consider themselves adept at anything involving mathematics. They fear that their jobs or careers are in jeopardy if they cannot use this new powerful tool. They have heard of the courses on SPC (statistical process control), which cover days or weeks, and believe that it must be very complicated and difficult to understand and use. This is not true. Anyone who understands grade-school arithmetic can be taught the principles of SPC and how to use it in six hours, with the help of a few worksheets (cookbook statistics) and a small 34-page handbook that gives all the necessary notes, equations, and tables.

Definition

By definition, SPC is "a mathematical means of determining the variation expected of a process from random samples taken during the process."

Although SPC is based on some fairly sophisticated mathematics, you don't have to know or understand this math to use the tool. Your automobile is governed by the mathematical laws of thermodynamics, but you don't have to know them to operate your car. The same is true of SPC.

Understanding SPC

To understand and use SPC, you only need to know about quality characteristics, random samples, and variability.

Quality characteristics

Quality characteristics are the distinguishing traits, qualities, or properties of the output of a system, be it a product or information, that are important to the user or customer. The "customer," incidentally, may be the person running the next operation.

Quality characteristics fall into only two groups:

1. Variables. Characteristics that can be measured, such as weight, length, density, and temperature.

2. Attributes. Characteristics that are either good or bad, such as a product that is glued or not glued, torn or not torn, fits or doesn't fit.

Random sample

A random sample is one item or a subgroup of items taken from a large group called a population. A population could be all the items in an order, or all the items run in a shift, or all the data produced in a day. The random sample must be taken in such a way that every item in the population has an equal chance of being taken. If the sample is selected for some reason, if it is not random, SPC will not work. It will produce false information.

It is important not to use SPC to intimidate the operator. If SPC is used in that way, he or she will not take random samples but instead will select samples to please his or her supervisor. Then none of the SPC data is usable.

Random sampling is as effective as 100% inspection and a lot cheaper. One-hundred percent inspection is usually only 70-90% accurate when done by human beings. Inspection does not build quality; the quality is already made. You don't get ahead by making a product and separating the good from the bad. If bad product is made, 100% inspection will usually not catch all of it. Some will get to the customer. It's better to make it right in the first place. SPC helps you do that.

Variability

Variability exists in all processes and systems. Everything in the universe varies. If you measure something over and over and your measuring instrument shows it never varies, that merely tells you the measuring instrument is not precise enough or sensitive enough to pick up the variation. After you have established a sampling plan and start to take random samples and record the variation of the characteristics that you are measuring, you can then use SPC. It tells you what variation to expect from a process.

Variability falls into only two groups:

1. Chance or common causes brought about by chance behavior in a random manner. It is stable within limits. SPC will establish limits of variation that can be expected consistently when only chance variation is in operation (called upper and lower control limits).
2. Assignable causes brought about by a special assigna-

Geisel is president of Simplified Systems, Inc., 1713 Towpath Court, Naperville, Ill. 60565.

ble cause not normally expected. This variation from assignable causes falls outside the limits of variation expected from chance causes alone.

Variability of any characteristic can be described graphically with a histogram or frequency diagram. This is a bar graph showing how many samples fall into each measuring unit.

Variability can also be described mathematically with measures of central tendency and dispersion.

The measure of central tendency usually used in SPC is the average. Average or mean is the sum of a group of measurements divided by the number of measurements. An average is expressed as a letter representing the measurement with a bar over it. Mathematically, an average of measurement X is written $\bar{X} = (X_1 + X_2 + X_3) / n$ or $(\Sigma X) / n$ where Σ means "sum of" and n equals the number of measurements. Averages, however, are not enough. A man tried to wade across a river where the average depth was noted as 1 ft. He drowned because although the average was 1 foot and most of the river was only 6 in. deep, it had a 10-ft hole into which he fell.

The measures of dispersion fall into two categories:

1. Range is the largest measurement in the group minus the smallest measurement. Mathematically, it is written $R = X_{\max} - X_{\min}$. If the man wading across the river knew the range of the river depth, which was 1 in. to 10 ft, he would have known there was a 10-ft hole somewhere in the river and would not have tried to wade across.
2. Standard deviation is the measure of dispersion expressed as a small sigma (σ) or a lower case s. To calculate the standard deviation, the average of the measurements $\bar{X}$ is subtracted from each measurement X . Each difference is then squared (multiplied by itself). These are then added up and divided by the number of measurements minus 1. Then the square root is taken of the resulting quotient, and you have the standard deviation. You can also purchase an inexpensive calculator, put in all the numbers, and press the $\sigma n - 1$ key to get the standard deviation of the measurements.

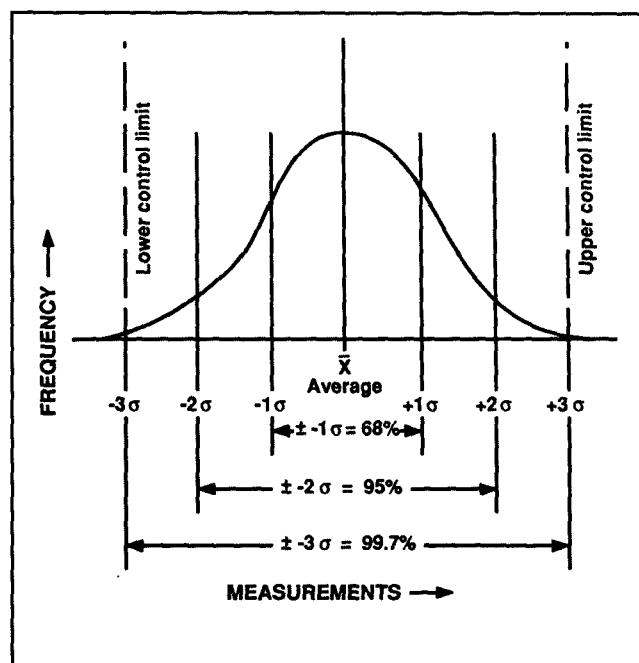
The equation for a standard deviation is:

$$\sigma = [\Sigma(X_i - \bar{X})^2 / n - 1]^{1/2}$$

Most measurements vary about their average in a "normal" fashion, giving a frequency diagram that forms the normal or bell curve. This means that 68% of the measurements will lie within ± 1 standard deviation (σ) from the average, 95% between $\pm 2\sigma$ and 99.7% within $\pm 3\sigma$. The $\bar{X} + 3\sigma$ is the upper control limit (UCL) and $\bar{X} - 3\sigma$ is the lower control limit (LCL) on SPC control charts and is the variation expected by common chance causes alone. Any measurements outside these limits have assignable causes. Figure 1 shows the normal curve with the $\pm 1, 2$, and 3σ limits.

For example, if the average of the measurements $\bar{X}$ was 10 and the standard deviation σ was 2, then 68% of the measurements should fall between $10 - 2$ and $10 + 2$, or 8 and 12. Ninety-five percent should fall between $10 - 2 \times 2$ and $10 + 2 \times 2$, or 6 and 14. Likewise, 99.7% should fall between $10 - 3 \times 2$ and $10 + 3 \times 2$, or 4 and 16. Any

1. The normal distribution curve

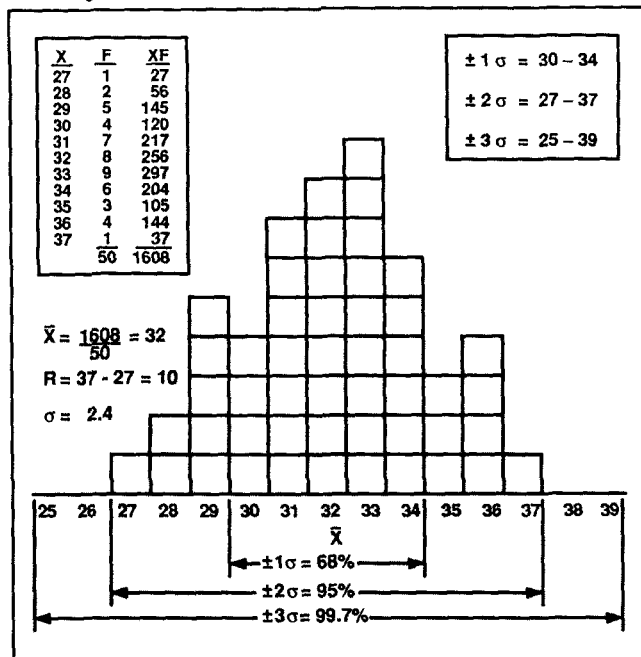


measurement outside the three limits, such as 3 or 17, would have occurred due to an assignable cause.

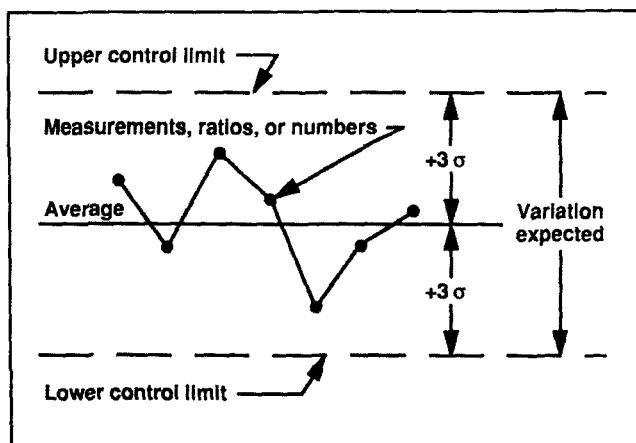
In the game of dice or craps, the numbers appearing on each throw will vary, but they will follow a distribution very close to a normal distribution, with the average $\bar{X}$ being 7 and the standard deviation being about 2. Therefore, about 68% of the time the player should roll between a 5 and 9, 95% of the time between 3 and 11, and in this case, of course, 100% of the time between 2 and 12. At a crap table in a gambling house, if they do not get this kind of variation at the end of the evening, they know that there is an assignable cause; someone has slipped in loaded dice. This is what SPC will do for you. It will show what variation to expect from your process, and you will know when you have an assignable cause (loaded dice). Remember, in a gambling house they know what variation to expect. They are not gambling; you are. In our processes, we don't want to gamble; we want to know what variation to expect so we can react quickly when an assignable cause shows up.

For example, we took a sample of 50 quality characteristic measurements from a process and made a frequency diagram or histogram of this data. A histogram is like a bar graph showing the distribution or variation of the measurements (Fig. 2). The total of the measurements is 1608. Therefore, the average $\bar{X}$ ($1608/50$) equals 32. The range R is $37 - 27 = 10$. The standard deviation σ calculated from a hand-held calculator is $\sigma = 2.4$. Therefore, the process measurements from this system should average 32. We should expect 68% of the measurements to fall between $\bar{X} + 1\sigma = 32 + 1(2.4)$ and $\bar{X} - 1\sigma = 32 - 1(2.4)$, or between 30 and 34; 95% should fall between $\bar{X} \pm 2\sigma = 32 \pm 2(2.4)$, or between 27 and 37; 99.7% should fall between $\bar{X} \pm 3\sigma = 32 \pm 3(2.4)$, or between 25 and 39. This would represent the upper control limit (39) and lower control limit (25) on an SPC control chart. This means that even though there are no 25, 26, 38, or 39s

2. Histogram



3. The control chart



in our sample, we can expect this process to produce some 4.7% of the time (99.7% - 95%). If the customer does not want any 25, 26, 38 or 39s, then we will have to change the process, not yell at the operator. However, if we found some measurements over 39 or under 25, this indicates something has changed, and we should look for an assignable cause. This is one of the powers of SPC. It not only tells what is happening—it tells what is likely to happen.

Applying SPC

Statistical process control involves the operator. It is the operator who makes the quality, and it is the operator who should check the quality as he makes the product.

SPC charts will tell him:

1. How much variation is expected, and when his process is stable and predictable

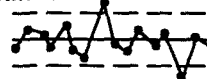
4. Control chart patterns

Normal

UCL
Avg.
LCL

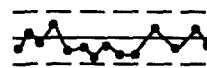
68% of points one half of the distance from average to control limits
27% of points in middle third
5% of points in last third

Outliers



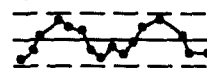
Points outside control limits
Look for assignable cause

Runs



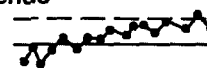
Sequence of points above or below average
7 in a row
10 out of 11
12 out of 14

Cycles



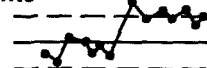
Variation that is repetitive, usually by shift or order

Trends



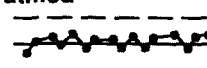
Gradual successive movement of the points up or down
Shows something gradually changing in wearing

Shifts



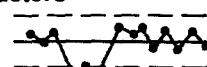
A sudden movement of points
Look for a change of material of crews

Stratified



Points hugging the average line usually indicate false reporting or insufficient accuracy of measuring equipment

Clusters



Small group of points
Look for an unusual batch of material

2. When to look for assignable causes of abnormal variation
3. If his process is capable of meeting specifications, and, if not, the percentage anticipated to be out of spec.

The operators in over 100 plants, with very few exceptions, have enthusiastically accepted the task of periodically checking their work and plotting it on a control chart. They like the idea of knowing what variation to expect, so they are not continually adjusting the process, which introduces more variation. With SPC, they will know when to adjust the process and when to leave it alone. They can work smarter, not harder. Operators want to produce good quality, and SPC promotes pride in workmanship. Most of the resistance comes from the front-line supervisors, superintendents, and occasionally managers, who sense some loss of authority and power.

To apply SPC, you first must select one or more characteristics produced by a process that you wish to

1. Plotting and calculating control limits

$\bar{X}$ and R charts

Used for quality characteristics which are measurable (variables), such as caliper, score bend ratio, opening force, basis weight, etc.

Steps to construct $\bar{X}$ and R charts:

A. Data collection

1. Select subgroup size (n), usually 3, 4, or 5.
2. Select frequency at which subgroup size will be taken often enough to catch nonconformities and representative of the process (shorter periods give faster feedback).
3. Take number of samples at random required for subgroup.
4. Measure quality characteristics of each sample in subgroup and record.
5. 100 or more measurements required to calculate control limits.

B. Calculations and plotting

1. Calculate average ($\bar{X}$) by adding all measurements in subgroup and dividing by number of measurements.
2. Plot $\bar{X}$ on $\bar{X}$ chart.
3. Calculate range (R) by subtracting smallest measurement from largest measurement in subgroup.
4. Plot R on R chart.
5. Connect all points with straight lines.

C. Control limits and central line

1. For $\bar{X}$ chart:

Calculate average of averages ($\bar{\bar{X}}$) by adding subgroup averages $\bar{X}$ and dividing by number of subgroups.

Calculate average range ($\bar{R}$) by adding subgroup ranges and dividing by the number of subgroups.

Calculate control limits:

$$\text{Upper control limit UCL} = (\bar{\bar{X}}) + A_2 \bar{R}$$

$$\text{Lower control limit LCL} = (\bar{\bar{X}}) - A_2 \bar{R}$$

$A_2 \bar{R} = 3$ standard deviations

2. For R chart:

Calculate control limits:

$$\text{Upper control limit UCL} = D_4 \bar{R}$$

$$\text{Lower control limit LCL} = D_3 \bar{R}$$

3. A_2, D_3, D_4 are constants dependent on subgroup size (n)

$\bar{X}$ and MR (moving range) charts for individuals

Used for quality characteristics which are measurable (variables), but only one measurement at a time is practical.

A. Select frequency of sample (same as $\bar{X}$ and R except subgroup size is 1).

B. Calculations and plotting

1. Plot measurements (X).
2. Calculate moving range by subtracting each measurement from the next and record (more than 10 readings required).

C. Control limits and central line

1. For $\bar{X}$ chart:

Calculate average ($\bar{X}$) by adding measurements and dividing by the number of measurements.

Calculate average range ($\bar{R}$) by adding moving ranges and dividing by the number of ranges (one less than the number of measurements).

Control limits:

$$\text{Upper control limit UCL} = \bar{X} + 2.66 \bar{R}$$

$$\text{Lower control limit LCL} = \bar{X} - 2.66 \bar{R}$$

$2.66 \bar{R} = 3$ standard deviations

2. For R chart:

$$\text{Upper control limit UCL} = 3.27 \bar{R}$$

$$\text{Lower control limit LCL} = 0$$

p and np charts

Used for quality characteristics which are either good or bad

(attributes) making the item good or bad.

p chart. Plot fraction or percent defective (p) in a sample.

n = Sample size which may vary.

$$\bar{p} = \frac{\text{Total number of defective items}}{\text{Total number of items inspected}}$$

$$\bar{n} = \frac{\text{Total number of items inspected}}{\text{Total number of samples}}$$

$$\text{UCL} = \bar{p} + 3 \sqrt{\frac{\bar{p}(1-\bar{p})}{\bar{n}}}$$

$$\text{LCL} = \bar{p} - 3 \sqrt{\frac{\bar{p}(1-\bar{p})}{\bar{n}}}$$

$$\text{Standard deviation} = \sqrt{\frac{\bar{p}(1-\bar{p})}{\bar{n}}}$$

np chart. Plot number of defective items (np) in a sample.

n = Sample size which must be constant.

$$\bar{np} = \frac{\text{Total number of defective items}}{\text{Total number of samples}}$$

$$\bar{p} = \frac{\bar{np}}{n}$$

$$\text{UCL} = \bar{np} + 3 \sqrt{\bar{np}(1-\bar{p})}$$

$$\text{LCL} = \bar{np} - 3 \sqrt{\bar{np}(1-\bar{p})}$$

$$\text{Standard deviation} = \sqrt{\bar{np}(1-\bar{p})}$$

c and u charts

Used for quality characteristics which are either good or bad (attributes) that occur in a sample.

c chart. Plot number of defects (c) in a sample, which may be one item, period of time, or group of items such as spots/yard² downtime/shift, etc.

n = Sample size, which must be constant.

$$\bar{c} = \frac{\text{Total number of defects}}{\text{Total number of samples}}$$

$$\text{UCL} = \bar{c} + 3 \sqrt{\bar{c}}$$

$$\text{LCL} = \bar{c} - 3 \sqrt{\bar{c}}$$

$$\text{Standard deviation} = \sqrt{\bar{c}}$$

u chart. Plot the number of defects per unit (u) such as downtime/run, errors per order, etc.

n = Sample size, which may vary.

$$\bar{u} = \frac{\text{Total number of defects}}{\text{Total number of units}}$$

$$\bar{n} = \frac{\text{Total number of units}}{\text{Total number of samples}}$$

$$\text{UCL} = \bar{u} + 3 \sqrt{\frac{\bar{u}}{\bar{n}}}$$

$$\text{LCL} = \bar{u} - 3 \sqrt{\frac{\bar{u}}{\bar{n}}}$$

$$\text{Standard deviation} = \sqrt{\frac{\bar{u}}{\bar{n}}}$$

Constants for formulas

n	A ₂	D ₃	D ₄	d ₂
2	1.88	0	3.27	1.13
3	1.02	0	2.57	1.69
4	0.73	0	2.28	2.06
5	0.58	0	2.11	2.33
6	0.48	0	2.00	2.53
7	0.42	0.08	1.92	2.70
8	0.37	0.14	1.86	2.85
9	0.34	0.18	1.82	2.97
10	0.31	0.22	1.78	3.06

n = sample or subgroup size

control. This characteristic may be determined by the customer, or you may rank a group of characteristics by their cost of poor quality, highest cost to lowest. This is called a Pareto chart. Then begin by selecting the highest-cost characteristic. You can analyze a problem by a cause-and-effect diagram, where you write all the causes of a problem (the effect) in a Fishbone diagram classified by methods, machines, materials, control, etc. Then select characteristics back in the process that, if controlled, would have the best chance of solving or reducing the problem. The further back in the process you can control characteristics the better.

Once a characteristic is chosen, we must select the proper SPC control chart. A control chart is a continuous picture of the variation of a process characteristic (Fig. 3).

You must determine if the characteristic is a variable or attribute.

For variables (measurable characteristics) use $\bar{X}$ and R, and $\bar{X}$ and MR.

For attributes (good or bad characteristics) use:

- Percent or fraction of defective items
Variable sample size, p chart
Constant sample size, np chart
- Number of defects per unit
Constant area or number of items, c chart
Variable area or number of items, u chart.

Control charts will:

1. Show how the process is varying and how it is likely to vary, not how you want it to vary
2. Show if assignable causes are acting on the process
3. Show if the process is stable
4. Show if the process is capable
5. Control limits show where 99.7% of the points should fall by chance alone.
6. Control limits are not related to specifications.

Once the chart has been selected, the size and frequency of the random samples must be determined. The sample must be random; that is, each item being produced by the process must have an equal chance of being taken for the sample. The sample must be representative of the process. That is, if the product is run on three shifts, all three must be sampled. Don't take all the samples in the morning. Take them throughout the day. The sample size should be large enough and taken often enough to catch nonconformities in the process. Larger samples provide more accuracy, but smaller ones are cheaper. More frequent samples give faster feedback.

The operator should take the samples, check or measure the characteristic, and plot the results on the chart. Under an SPC program, the operator is responsible for the quality, and inspectors do not inspect items in production. Inspectors calculate control limits, train and assist operators in charting and understanding customer requirements, calibrate and maintain test instruments, and periodically audit operator inspections. To start with, the operators should eliminate or minimize adjustments to the process while measurements are being gathered to calculate the control limits. Leave the machine alone and see what it can do. The chart will then be the machine talking to you, telling you how it will vary normally.

Table I provides an outline of how to plot and calculate control limits for the charts selected.

After 25–100 samples are taken, plotted, and the average and upper and lower control limits calculated and drawn on the control chart, we look for the following patterns. These patterns will indicate whether or not an assignable cause of variation is present (Fig. 4).

If a process is in control, the process capability can then be determined to see if it is capable of meeting the specifications.

For an $\bar{X}$ and R chart, calculate the upper and lower process control limits as follows:

$$UPCL = \bar{\bar{X}} + 3 (\bar{R}/d_2)$$

$$LPCL = \bar{\bar{X}} - 3 (\bar{R}/d_2)$$

For the other charts, the upper and lower process control limits are the same as the upper and lower control limits.

If the process control limits are outside the specification limits, then the process is not capable of meeting the specifications.

To calculate the anticipated percent of output of the process that will be out of specification, calculate a Z score for the upper specification limit (USL) and lower specification limit (LSL) as follows:

$$Z_{USL} = (USL - \bar{\bar{X}})/\sigma$$

and

$$Z_{LSL} = (LSL - \bar{\bar{X}})/\sigma$$

where

$$\sigma = \bar{R}/d_2$$

Then look up the value of Z in a table of the area under the normal curve, which will be the percent of output out of specification.

The process should be allowed to run to establish statistical control, which is a stable state of randomness in which limits of variation are predictable and will remain, for practical purposes, fixed until the system is altered.

If control limits are outside specification limits, the system must usually be changed by management, with operator assistance.

If points are outside control limits, or show abnormal patterns, the operator should look for assignable causes and make changes.

If all points are in control, with normal variation, the operator should leave the process alone.

If the chart shows lack of control, take action regardless of where the specifications are.

Operators should strive for control; management should deal with specifications.

With SPC, the operator inspects for the quality characteristics he or she is producing and plots the results on a chart with averages and control limits. He or she then runs to the control limits, not the specification limits, and reacts only when a point gets out of control. This means the operator will now react before the process gets out of specification and produces unacceptable product. He or she will know what variation to expect and, therefore, will

not adjust the system needlessly, wasting effort and introducing unwanted variations.

Conclusion

Every process varies. SPC tells what variation to expect from a process when it is running normally, which means it is stable and predictable. It will also show what variations are caused by assignable causes, which can be identified and corrected. It will tell if the process is capable, or if it is not, the percentage of output which will be outside of specifications. It tells what is happening and what is likely to happen, not what you want to happen.

With SPC, operators can control their quality, have pride in their work, reduce waste, increase productivity, and produce a product that meets the customer's perception of quality.□

Received for review May 28, 1989.

Accepted Feb. 12, 1990.

Presented at the TAPPI 1989 Polymers, Laminations and Coatings Conference.