Performance Excellence in the Wood Products Industry
Statistical Process Control

Part 7: Variables Control Charts

EM 9109 • May 2015

Scott Leavengood and James E. Reeb

Our focus for the prior publications in this series has been on introducing you to Statistical Process Control (SPC)—what it is, how and why it works, and how to use various tools to determine where to focus initial efforts to use SPC in your company.

SPC is most effective when focused on a few key areas as opposed to measuring anything and everything. With that in mind, we described how to:

- Use Pareto analysis and check sheets to select projects (Part 3)
- Construct flowcharts to build consensus on the steps involved and help define where quality problems might be occurring (Part 4)
- Create cause-and-effect diagrams to identify potential causes of a problem (Part 5)
- Design experiments to hone in on the true cause of the problem (Part 6)

Now, in Part 7, we describe the primary SPC tool: control charts. Assuming our experiment helped identify target values for key processes, we are ready to focus on day-to-day control and monitoring to ensure the process remains stable and predictable over time.

It is important, however, to not lose sight of the primary goal: Improve quality, and in so doing, improve customer satisfaction and the company's profitability.

How can we be sure our process stays stable through time?

In an example that continues throughout this series, a quality improvement team from XYZ Forest Products Inc. (a fictional company) determined that size out of specification for wooden handles (hereafter called out-of-spec handles) was the most frequent and costly quality problem. The team identified the process steps where problems may occur, brainstormed potential causes, and conducted an experiment to determine how specific process variables (wood moisture content, species, and tooling) influenced the problem.

Assuming the team makes a process change, they next need to (1) verify that the change actually produces the desired beneficial result and (2) ensure the process remains stable through time. In other words, they need to sustain the gains they've made.

Scott Leavengood, director, Oregon Wood Innovation Center and associate professor, Wood Science and Engineering; **James E. Reeb**, Extension forester and associate professor, Forest Engineering, Resources, and Management, Lincoln County; both of Oregon State University.

Oregon State University OSU Extension Service

The team's experiment revealed that moisture content as well as an interaction between wood species and tooling affect the number of out-of-spec handles. They identified two options for process change:

1. If the company can tightly control and monitor moisture content and prefers not to change tooling each time they switch between birch and poplar, both species should be machined at 6% moisture content using the new tooling.
2. If the company can't tightly control moisture content and changing tooling between species is feasible, then the existing tooling should be used with birch (regardless of moisture content) and the new tooling should be used for poplar (regardless of moisture content).

Let's assume the team chooses the first option, which requires tightly controlling and monitoring moisture content. Simply checking moisture content periodically might do the trick. However, if they see variation (and given the natural variability of wood, we know they will), how do they know how much variability is normal and acceptable? At what point should they stop the process and look for a problem? When should they let things run? A control chart is precisely the tool to use to answer these questions.

Control charts: A quick introduction

Control charts are graphs that display the value of a process variable over time. For example, we might measure the moisture content of five items at 8:00 a.m. and plot the average on a chart. We would then repeat the process at regular time intervals. In addition to a series of points, control charts also include a centerline that represents the overall average of the variable being monitored and upper and lower limits known as control limits. Figure 1 shows an example control chart for average moisture content. We describe how to construct and interpret control charts later in this publication.

Variability is inherent in all processes. There are two broad categories of variability:

- Variability that is **inherent** to the process (also called random or common cause)
- Variability that is due to some **assignable cause** (also called special cause)

We discuss these categories in more detail below. For now, know that the primary objective of control charts is to answer this question: **Is the variability within the expected range for this process (inherent), or has something changed in the process to affect where the process is centered and/or the amount of variability (assignable cause)?**

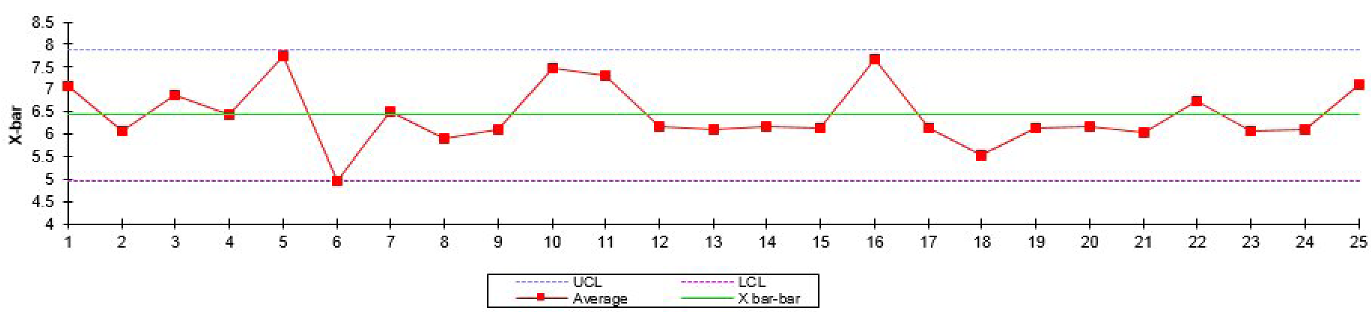

Figure 1. Example X-bar control chart for average wood moisture content.

Points above the upper control limit or below the lower control limit as well as trends in the pattern of the points on the chart help us determine when variability is within the expected range. In SPC terminology, this is called being **in control**. Although control charts have a foundation in statistics, the primary driver behind the invention of these charts was economics. Chasing "false alarms" wastes time and resources. When a control chart indicates a process is in control, it's likely a waste of time to go searching for problems. In contrast a control chart that indicates a process is **out of control** indicates we should take action to return stability to the process.

There are two types of control charts: charts for **variables** and charts for **attributes**. The primary difference between the two is the type of data being collected. Variables data are typically measured values such as moisture content, thickness, weight, time, and adhesive viscosity. Attributes are usually counts or tallies of some factor such as number or the percentage of nonconforming items in a sample. In this publication, we focus on variables charts. We will address attributes charts in Part 8.

Anatomy of a control chart

To understand how control charts work, it's helpful to examine their components. As shown in Figure 1, a control chart has points, a centerline, and control limits. To effectively monitor a process, we need to track process **centering** and **variability**. Therefore, we have two control charts: one for centering and another for variability. Figure 1 is an example of a chart used to monitor process centering.

Chart for process centering

This chart is known as the **X-bar** chart and is used to monitor sample averages. The individual measurement values are referred to as the x values, and the average of these values is referred to as x-bar, shown with the symbol $\bar{x}$. The points on the chart are the averages of the individual values in each sample group. The centerline is the overall, or grand, average and is referred to as x double bar or x bar-bar and shown as $\bar{\bar{x}}$. The upper and lower horizontal lines are known as the **upper control limit** and **lower control limit**, respectively. We discuss these limits in a later example (Calculate control limits, page 12).

Chart for process variability

This chart is known as either the **R chart** (if the range is used) or **s chart** (if the standard deviation is used)[1]. For this discussion, we will focus on the sample **range** (maximum value in the sample minus the minimum value). The points on the chart are the ranges of the individual values in each sample group. As with the X-bar chart, the centerline on the R chart is the overall average of all sample ranges, shown as $\bar{R}$. The R chart also has upper and lower control limits.

Don't worry if you aren't clear on the details at this point. Later, we will walk through an example to show how to construct and interpret X-bar and R charts for our XYZ example. But first, we need to review some basic statistics principles because the points, centerline, and limits on control charts are derived using statistics.

[1] As we discuss below, the choice of chart to use depends on sample size. If the sample size is three to five items, use the R chart. If the sample size is variable or has more than 10 items, use the s chart.

Overview of statistics

When evaluating SPC courses and workshops, participants sometimes comment that there is "too much math," even if we show only one formula. There is some complex arithmetic involved in SPC. However, in this publication series and our related workshops, we always aim to focus on the critical information and provide a brief review of basic concepts when needed. If you're interested in learning more about the statistics and math involved in SPC, see the resources listed in the "For more information" section at the end of this publication. Also, there are many computer programs that will handle the math for you. For reference, formulas used in this publication are listed on page 18.

For now, let's focus on key statistics concepts and values: **population**, **sample**, **average**, **range**, and **standard deviation**. We'll also review different categories of **variability** (inherent vs. assignable cause) and conclude by discussing the **normal distribution**, which is very important in SPC.

Population vs. sample

When we need to estimate the centering and variability of a process, we usually take a sample rather than measure every single item in the population.

For example, if we wanted to know the average height of adults in the United States, we could measure the height of every single adult (the population) and divide by the number of adults. That would give us the population average. But that is not feasible. It would be too expensive and take too much time to collect that much data. Instead, we could measure the height of 100,000 adults from across the United States (a sample), calculate the average height of the sample, and use that as an estimate of the population average.

Use common sense when choosing a sample. If you want a reliable, reasonable estimate, you need to use samples that represent the range and variety of values found in the population. Think about our example of adult average height. If we choose a sample of adults only from people who shop in the "big and tall" section of clothing stores, our average would probably be skewed.

Average: Measuring process centering

To estimate process centering, we use the average (also called the mean). The average gives us an indication of the middle of a group of data. To calculate the average, simply add the value of each sample and divide by the number of samples (the sample size). In statistics, the symbol for sample size is a lowercase letter n. Table 1 shows five moisture content readings and the sample average.

Table 1. Sample values and sample average

Measurement	Moisture content (%)
X_1	8.8
X_2	8.0
X_3	5.7
X_4	6.1
X_5	6.8
Sum	35.4
Average ($\bar{x}$) (35.4/5)	7.08

Range and standard deviation: Measuring variability

Because a primary focus of quality control is to minimize and control variability, we also need to measure the variation of a sample. The range is a common measure of variation. To calculate the range, simply subtract the minimum value from the maximum value. The range of the sample in Table 1 is 3.1 ($X_1 - X_3$, or 8.8 – 5.7 = 3.1).

The range is easy to calculate and understand. Most people who construct a control chart by hand with paper, pen, and a calculator use the range. However, the range uses only two data points from a sample. If we want a more accurate measure of variation, we need to use all the sample data. That's what the standard deviation does.

The standard deviation is a better measure of variation, but the calculations are more complicated. In SPC, we simply use table values to estimate the standard deviation for a given range. But because understanding standard deviation is critically important for understanding the control limits on control charts, we provide an example to explain this concept further.

The standard deviation answers the question: On average, how far away are the values in a sample from their average? Table 2 shows the same data as Table 1 with two more columns. The third column shows the average subtracted from each measurement. Notice that the average of these deviations is zero—because they sum to zero. That is a key principle of the average; for any set of data, deviations from the average will always sum to zero. But a zero value here would indicate zero variation. And we can see that's not the case.

To remove the issue of positive and negative values for the deviations, we square the values (the fourth column). The sum of those values is a number in squared units. For this example, the units are "percent squared," which is meaningless. If we divide the squared sum of deviations (6.587) by one less than the sample size (n – 1), we get what is known in statistics as the sample variance (1.647), again in units of percent squared. The last step is to calculate the square root of the variance to get the sample standard deviation (1.283).

What does this value mean? In this example, the standard deviation indicates that, on average, these measurements are 1.283 percentage points away from their average.

Table 2. Calculation of standard deviation

Measurement	Moisture content (%)	Deviation from average	Squared deviation from average
X_1	8.8	8.8 – 7.08 = 1.70	2.904
X_2	8.0	8.0 – 7.08 = 0.89	0.799
X_3	5.7	5.7 – 7.08= -1.37	1.866
X_4	6.1	6.1 – 7.08 = -0.98	0.953
X_5	6.8	6.8 – 7.08 = -0.26	0.066
Sum	35.4	0	6.587
Average ($\bar{x}$)	7.08 (average)	0	1.647 (variance)
			1.283 (standard deviation)

Why divide by n – 1 rather than simply n, as when calculating an average? The standard deviation is a good, but not perfect, estimator of the real value of interest: the population standard deviation. Statistical theory shows that the standard deviation of small samples taken from a population is **biased**; specifically, it tends to be less than the true value. So we inflate the value a bit by dividing by n – 1.

Let's review:

For this sample of five moisture content readings, the sample average (measure of process centering) is 7.08%. The measures of process variability are the sample range at 3.1% and the sample standard deviation at about 1.28%.

We have two values to estimate variability; which should we use? Quality control experts recommend using the range if sample size (n) is small (e.g., three to five items), and the standard deviation if sample size varies or is 10 or more. If the sample size is six to nine items, you're free to choose either measure. If you do the calculations by hand, range has obvious advantages. If you use a programmable calculator or software program, calculating the standard deviation is straightforward.

Categories of variability

Are you still wondering why we need to do all this math for quality control? When manufacturers realize that variability of their processes is excessive, perhaps via internal rejects or even customer complaints, it is often very time consuming and expensive to troubleshoot a process to determine and eliminate the root causes of the variability. Troubleshooting often requires shutting down a process. Again, it's about economics and efficient use of resources. If you are going to invest resources in troubleshooting, make sure there really is a problem.

Also remember that the primary objective of control charts is to determine if the variability in a process is within the expected range or due to some special circumstance, such as a dull cutterhead, incorrect operating procedures, or pockets of high moisture content. Using statistics allows us to evaluate the patterns on a control chart and determine if the variability is inherent or due to an assignable cause.

Making this determination is important because the approach to troubleshooting is different for inherent and assignable-cause variability. We distinguish between the two with statistical probabilities. The further away a value is from its average, the less likely it is to occur. But how far away is enough to take action? The normal distribution helps answer that question.

Normal distribution

Statistical distributions provide a way to link estimates of mean and variability. Part 2 in this series (*How and Why SPC Works*) described how to summarize data with a histogram—a bar chart of the frequency of values within specific ranges. Figure 2 shows the histogram from Part 2 overlaid with the smooth, bell-shaped curve of the normal distribution. The curve isn't a perfect fit for the histogram, but it's reasonably close.

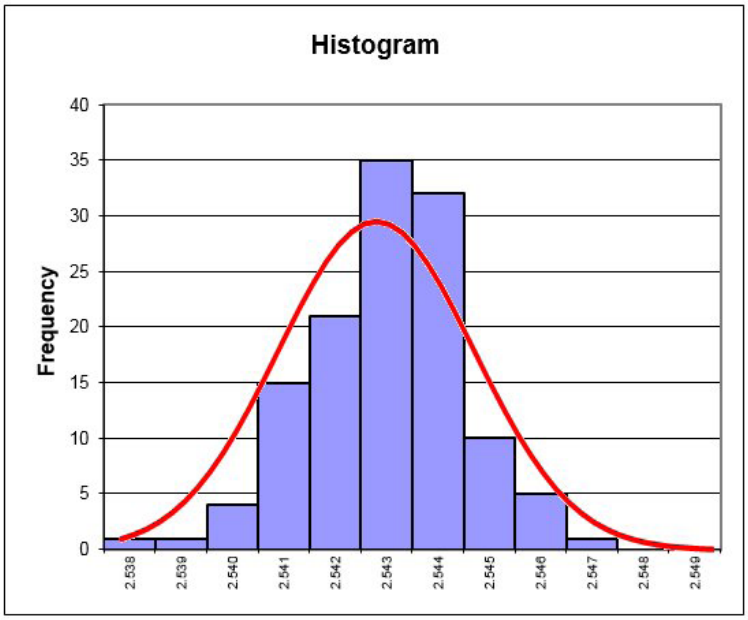

Figure 2. Histogram from Part 2 (*How and Why SPC Works*) with normal curve.

The normal distribution represents many natural phenomena well—the height of adults is a good example. It can even work with processes that are not normally distributed. According to statistical theory (specifically, the central limit theorem), sample means are approximately normally distributed even when the individual values themselves come from a non-normally distributed process (see sidebar, page 8).

A normal curve indicates that observations (measurements) closer to the average are more common than those farther away from the average. The width (spread) of the curve represents the standard deviation. More precisely, the average plus and minus three times the standard deviation accounts for well over 99% of the distribution if the data are normally distributed.

Let's return to our adult height example. If the average height of adults in our sample is 5′10″ with a standard deviation of 2½″, the normal curve that fit these data would have the left tail at 5′2½″ (5′10″ – (3 × 2½″)), the peak at 5′10″, and the right tail at 6′5½″.

Consider what you know about people's heights. Seeing a person who is within a few inches of 5′10″ is pretty common. Seeing an adult who is 5′3″ is a bit less common, but not rare. On the other end, an adult who is 6′5″ is not common, but not unusual either. The normal curve tells us that we can expect more than 99% of the population to fall within plus and minus three standard deviations of the average.

The tails of the distribution do not end; in theory, they go to infinity in both directions. Although we know infinity is not possible for height (no one has negative height or infinite height), we do know that heights above 6′8″ do occasionally occur. For example, former professional basketball player Manute Bol is 7′7″—12.4 standard deviations above the average!

The normal distribution works the same for height as it does for out-of-spec wooden handles. It also serves as the foundation for the control limits on control charts. The control limits are the primary means we use to distinguish between inherent and assignable-cause variability.

Let's return to our XYZ example and put these statistical principles into practice.

The top graph in Figure 3 shows a population that is exponentially distributed. This looks nothing like the bell-shaped curve of a normal distribution. One example of an exponential distribution is human age. There are no negative values and very few people who are very old. (The example in Figure 3 shows a simple exponential distribution that varies from zero to four, rather than human ages from zero to 100+.)

The middle graph shows the distribution of a typical sample of 10 from this population. The bottom graph shows the distribution of the average of 200 samples of size 10 from this population. The population is very non-normally distributed, but the distribution of sample averages from this population is reasonably normally distributed.

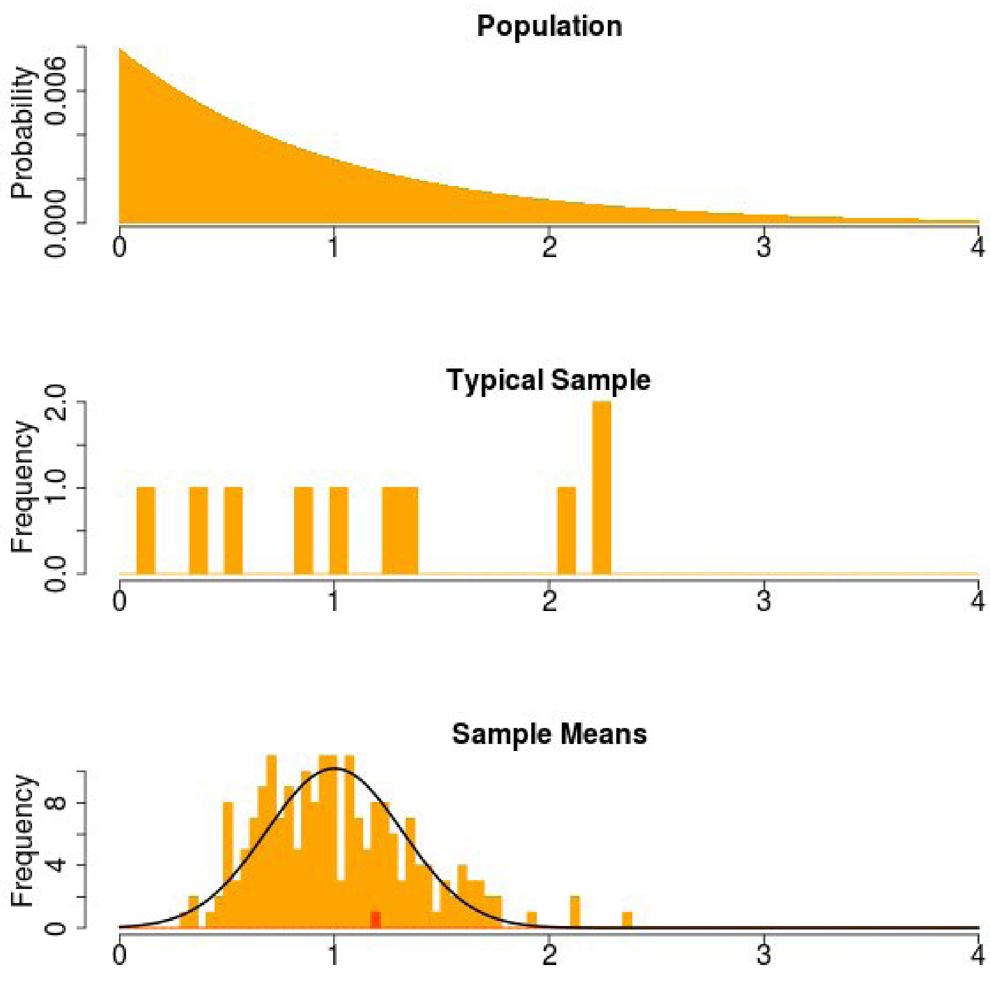

Figure 3. Graphs showing an exponentially distributed population (top), typical sample (n = 10) from this population (middle), and the distribution of sample averages (bottom).

Source: *https://ihstevenson.shinyapps.io/sample_means/*

Example: XYZ Forest Products Inc. uses SPC to monitor moisture content

In Part 6, the quality improvement team at XYZ Forest Products Inc. designed an experiment and discovered that moisture content as well as an interaction between wood species and tooling affected the number of out-of-spec handles. For this example, we assume the team chooses to begin by controlling and monitoring moisture content so they can machine both poplar and birch at 6% moisture content using the new tooling.

Given the natural variability of wood, we know the team will see some variation in moisture content. How much variability is acceptable? At what point should the team stop the process and look for a problem? A control chart can help answer these questions.

Collect data

The first step is to collect data. We also decide how much data to collect, as well as where and when to collect it. In SPC, the approach to data collection is known as **rational subgrouping**[2]. If money and time weren't an issue, we would just measure everything. But since money and time are always an issue, we need an efficient approach—one that gives the maximum amount of information with the minimum amount of effort.

We should use a sampling approach that minimizes the opportunity for variation to occur *within* a sample but provides much greater opportunity for variation *between* samples. To do this, we collect a small sample (3 to 10 items) in rapid succession, in the order they were produced, and then wait some amount of time. We then collect another sample, and repeat.

When collecting items in rapid succession, there is minimal opportunity for variation to occur within the sample. If we wait 30 minutes or more to collect the next sample, there is a greater likelihood the process will change between samples. Contrast this approach to measuring one item at 8:00 a.m., another at 8:30 a.m., and so on until we get a sample of five. If the process is prone to fluctuate (most are), that variation will occur within the sample, and we will have less ability to detect changes between samples. This concept will make more sense when we later discuss how the sample range is used to calculate the control limits on control charts.

If we have a stable process, we should expect the variability within groups to be the same as the variability between groups. For stable processes, it doesn't matter if we collect five items at 8:00 a.m. or five items over a course of 5 hours. However, we have to collect samples in such a way that we can test this assumption.

Because control charts are time-series data, we measure items in the order they were produced. As we scan a control chart from left to right, we are watching the process through time. People sometimes want to collect data at a more convenient location (e.g., the warehouse). Unless there is a way to know for certain the order the products were produced (barcoding, perhaps?), collecting data in the warehouse doesn't work for SPC—unless, of course, you happen to be monitoring one of the warehouse operations.

[2] The terms *sample* and *subgroup* are often used interchangeably in SPC.

To reap the most benefits from SPC, your goal should be to collect and plot data in real time in the real environment and then make decisions about the process as soon as possible. Think of it this way: If you collect data but don't plot it until the next day, and the charts show an unstable process, how much out-of-spec product was produced (and how much money was wasted) before you detected the problem?

Let's return to our XYZ example. The team collects five moisture content readings every 20 minutes beginning at 8:00 a.m. Table 3 shows the first 12 samples.

Analyze data

Average and range

To construct a control chart, we now need to calculate the average ($\bar{x}$) and range (R) for each sample. We then need to calculate the overall average ($\bar{\bar{x}}$) and overall range ($\bar{R}$). Table 4 shows those results.

Table 3. Moisture content sample data

	Time											
Measurement	**8:00**	**8:20**	**8:40**	**9:00**	**9:20**	**9:40**	**10:00**	**10:20**	**10:40**	**11:00**	**11:20**	**11:40**
X_1	8.8	6.9	7.0	6.7	6.9	5.1	5.0	7.0	6.8	7.9	7.0	7.0
X_2	8.0	7.9	8.0	7.6	7.8	4.7	6.9	4.8	4.9	7.9	7.7	4.8
X_3	5.7	5.6	5.8	8.1	8.8	5.0	7.0	5.8	5.8	6.7	6.8	5.8
X_4	6.1	5.0	9.0	5.0	7.1	5.1	7.6	5.0	5.0	7.1	7.0	5.0
X_5	6.8	4.9	4.7	4.8	8.0	5.3	6.1	6.9	8.0	7.9	8.1	8.2

Table 4. Moisture content sample data showing sample averages and ranges and overall average and range

	Time												
Measurement	**8:00**	**8:20**	**8:40**	**9:00**	**9:20**	**9:40**	**10:00**	**10:20**	**10:40**	**11:00**	**11:20**	**11:40**	
X_1	8.8	6.9	7.0	6.7	6.9	5.1	5.0	7.0	6.8	7.9	7.0	7.0	
X_2	8.0	7.9	8.0	7.6	7.8	4.7	6.9	4.8	4.9	7.9	7.7	4.8	
X_3	5.7	5.6	5.8	8.1	8.8	5.0	7.0	5.8	5.8	6.7	6.8	5.8	
X_4	6.1	5.0	9.0	5.0	7.1	5.1	7.6	5.0	5.0	7.1	7.0	5.0	
X_5	6.8	4.9	4.7	4.8	8.0	5.3	6.1	6.9	8.0	7.9	8.1	8.2	
$\bar{x}$	7.08	6.07	6.88	6.44	7.74	5.04	6.51	5.91	6.09	7.48	7.30	6.16	6.56 ($\bar{\bar{x}}$)
R	3.1	3.0	4.3	3.4	1.9	0.6	2.6	2.1	3.1	1.2	1.3	3.4	2.5 ($\bar{R}$)

We now have enough information to plot the points on X-bar and R charts and draw the centerlines. We still need control limits. These limits are what allow us to examine the chart to determine if the is variability inherent, or if there is some assignable cause influencing the process. To calculate the limits, we again turn to statistics.

Remember that in the normal distribution, more than 99% of the population data will fall within plus and minus three standard deviations of the mean. We use this principle to draw control limits on the chart. For the X-bar chart, the upper control limit is the centerline ($\bar{\bar{x}}$) plus three standard deviations of the average; the lower control limit is the centerline minus three standard deviations of the average. The phrase *of the average* is important.

Standard deviation

On the X-bar chart, we are monitoring the averages of samples, not the individual measurements. Therefore, our control limits cannot simply be based on our estimate of the standard deviation. They must reflect the fact that the variability of the averages is different (much less, in fact) than the variability of the individual measurements. Figure 4 presents this concept using number lines.

The vertical lines on Figure 4 show the overall variability of the individual measurements from the four samples, which range from 4.7 to 9.0. The bottom row (below the solid horizontal line) shows that variability of the averages ranges only from 6.07 to 7.08. Notice that the variability of the averages is significantly less than the variability of the individual measurements.

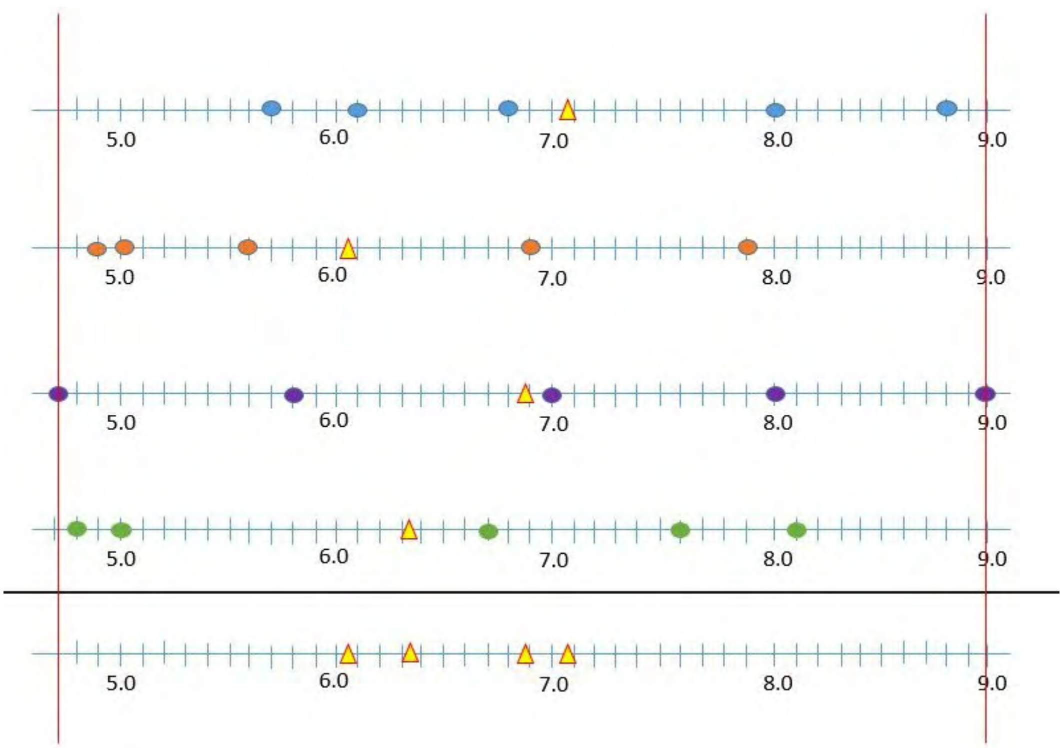

Figure 4. Variability of individual measurements vs. variability of averages.

Ovals show individual measurements for the first four samples in Table 4 (from 8:00 to 9:00 a.m.).
Triangles show the averages for each of the four samples.
The bottom row (below the solid horizontal line) shows the distribution of the averages.

Statistical theory allows us to take this one step further. We know that the standard deviation of the averages is equal to the standard deviation of the samples divided by the square root of the sample size ($\sqrt{n}$). In this case (samples of 4), the standard deviation of the averages is ½ the standard deviation of the individual measurements ($1/\sqrt{4}$). If we had samples of 100, the standard deviation of the averages would be $1/10$ the standard deviation of the individuals ($1/\sqrt{100}$). This is important because control charts are based on three standard deviations of the average.

The final step before calculating control limits is to estimate the standard deviation of the process from our samples. Recall that we collected samples, calculated the range for each, and calculated the overall range ($\bar{R}$). Again, statistical theory tells us how to estimate process standard deviation from the sample range. In fact, the relationship between the range of a sample from a normal distribution and the standard deviation of that distribution is well known. In SPC, this relationship has been simplified to a table value that is based on sample size. The table value is known as d_2, and it's available in textbooks[3]. The formula is: standard deviation = $\bar{R}/d_2$. Table 5 shows a few d_2 values.

Table 5. Table values for d_2

Sample size	d_2
2	1.128
3	1.693
4	2.059
5	2.326
6	2.534

Calculate control limits

Remember: the estimate of the standard deviation of the averages is the standard deviation of the samples ($\bar{R}/d_2$) divided by the square root of the sample size ($\sqrt{n}$).

The formula for the upper control limit of the $\bar{x}$ chart ($UCL_{\bar{x}}$) is the overall average ($\bar{\bar{x}}$) plus three standard deviations of the average:

$$UCL_{\bar{X}} = \bar{\bar{X}} + 3\left[\frac{\bar{R}/d_2}{\sqrt{n}}\right]$$

The formula for the lower control limit ($LCL_{\bar{x}}$) s identical, except the values are subtracted rather than added. Notice that only two of the values in this formula come from our sampling: $\bar{\bar{x}}$ and $\bar{R}$. Three of the values are constants: 3, d_2, and $\sqrt{n}$. So we can simplify the formulas by using another table value. The value A_2 incorporates all constants in the formula and simplifies the equations:

$$Centerline = \bar{\bar{X}}$$

$$UCL_{\bar{X}} = \bar{\bar{X}} + A_2\bar{R}$$

$$LCL_{\bar{X}} = \bar{\bar{X}} - A_2\bar{R}$$

[3] The Oregon Wood Innovation Center website also provides common table values for SPC: *http://owic.oregonstate.edu/spc*

Now we have control limits for the X-bar chart. Next we need control limits for the R chart (UCL_R and LCL_R). There are table values for calculating these limits as well, specifically D_3 and D_4.

$$Centerline = \bar{R}$$

$$UCL_R = D_4\bar{R}$$

$$LCL_R = D_3\bar{R}$$

Table 6 lists several table values for calculating the limits on X-bar and R charts. For small samples, D_3 is 0. Therefore, there is no lower control limit for the R chart when the sample size is six or fewer.

Formulas are provided at the end of this publication for calculating control limits for X-bar and s charts (control charts that use sample standard deviation instead of sample range).

Table 6. Table values for calculating control limits on $\bar{x}$ and R charts

Sample size	A_2	D_3	D_4
2	1.880	0	3.267
3	1.023	0	2.574
4	0.729	0	2.282
5	0.577	0	2.114
6	0.483	0	2.004

Construct control charts

Let's return to our XYZ example and construct control charts using the data from Table 4. These 12 samples provide all necessary information: n = 5, $\bar{\bar{x}}$ = 6.56, and $\bar{R}$ = 2.5.

Use the simplified equations and the A_2 table value from Table 6 to calculate upper and lower control limits for the X-bar chart:

$$UCL_{\bar{X}} = \bar{\bar{X}} + A_2\bar{R} = 6.56 + 0.577(2.5) = 8.00$$

$$LCL_{\bar{X}} = \bar{\bar{X}} - A_2\bar{R} = 6.56 - 0.577(2.5) = 5.12$$

Next calculate the upper control limit for the R chart (the sample size is five, so there is no lower control limit):

$$UCL_R = D_4\bar{R} = 2.114(2.5) = 5.29$$

Figure 5 shows completed control charts for the data in Table 4.

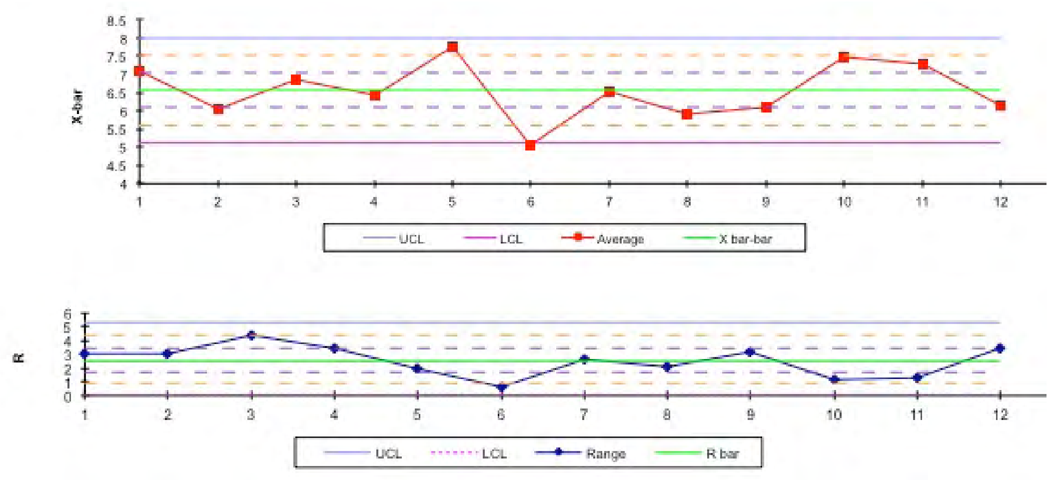

Figure 5. X-bar and R control charts for data in Table 4.

Dashed lines between the solid centerline and control limit lines represent one and two standard deviations from the centerline. These intermediate lines are not required for control charts but are included here to assist with interpretation.

Interpret control charts

Now we need to decide what these charts mean for the XYZ team. Our primary interest is the stability and consistency of the process. The question is not whether the customer will be satisfied or even if the process is on target (at 6% moisture content).

Determine what type of variability is present

Remember that the primary objective of control charts is to determine if the variability in a process is within the expected range (inherent) or due to some special circumstance (assignable cause). If the charts show evidence of only inherent variability, we may conclude that the process is in control (stable). If there is evidence that assignable-cause variability is present, we conclude that the process is out of control (unstable). In SPC, out of control does not mean complete mayhem. It simply means the process is unpredictable—we cannot yet predict with any certainty where it is centered or the magnitude of the variability. If the process is out of control, we need to take action to return stability to the process.

What are the indicators of assignable-cause variability? The centerline and upper and lower control limit lines are key. Remember that in a normal distribution, more than 99% of the data in a normal distribution will fall within plus and minus three standard deviations of the mean—exactly where we've drawn the upper and lower control limit lines on the control charts. If a point falls outside these limits, we conclude that (1) we have witnessed a very rare event or (2) the process has changed.

There is a chance of a false alarm. That is, a point could fall outside the limits and the process would still be in control. But the odds of that happening are quite slim. So as a first-level indicator, a point outside the control limits is a sign that assignable-cause variability is present and we should take action. Other indicators of an out-of-control process are known as the **Western Electric Rules** because they were developed by the Western Electric Company and published in their *Statistical Quality Control Handbook* (1956). These rules are summarized in Table 7. There are other versions of these rules as well.

Table 7. Western Electric Rules for interpreting control charts

Rule	Description
1	Any point outside of the control limits
2	Two of three consecutive points in outer one-thirds (beyond two standard deviations)
3	Four of five consecutive points beyond one standard deviation and on same side of centerline
4	Seven consecutive points above or below centerline
5	Five consecutive steps upward or downward

These rules are based on statistical principles. For example, Rule 4 is like flipping a coin seven times and getting heads each time; it's not impossible, but the likelihood of this happening is pretty low. The more rules you apply, the higher the chance of false alarms. As always, there are trade-offs. For processes that are critical to quality, it makes sense to apply all the rules at the expense of a few false alarms. For processes that are less critical, perhaps only Rule 1 might be used.

Examine X-bar and R charts

When examining X-bar and R charts, always start with the R chart. Remember that limits on the X-bar chart are based on R. If the range isn't in control (stable), our estimate of the range isn't reliable and, therefore, we can't be confident that the limits on either chart are accurate. Look at the R chart in Figure 5. There are no indications that the process is out of control. All points are within the limits, and none of the rules listed in Table 7 are violated. Now we can examine the X-bar chart.

The X-bar chart is not in control. Point six is just below the lower control limit. We could conclude that the process is stable and we just witnessed a very rare event. If that's the case, we proceed as if things were ok. Alternatively, we could conclude that something has changed and explore further to see what the problem might be. In this case, we see that Sample 6 is unusual. All of the values are low. Perhaps more important, the range is suspiciously low. Could this have been a measurement error? Or perhaps the same handle was inadvertently measured more than once (which would explain the very low range)? Such a situation is cause for exploration to determine the root cause.

For this example, we'll assume that the XYZ team investigated and discovered that the meter used at the time of Sample 6 was faulty. Now what? The correct approach is to exclude this sample; recalculate $\bar{\bar{x}}$, $\bar{R}$, and the control limits; and then determine if the charts are in control. The team found that without Sample 6, the process is in control.

This is a basic, initial assessment of the process. To be confident that you've thoroughly assessed the stability of a process, take at least 25 samples. Figure 6 shows the control charts for the process after the XYZ team collected 13 more samples. Note that Sample 6 was eliminated because of the measurement error. These charts show that the process is in control; there are no indications on either chart that assignable-cause variability is influencing the process. Therefore, we conclude that the process is in control at $\bar{\bar{x}}$ = 6.49% and $\bar{R}$ = 2.6.

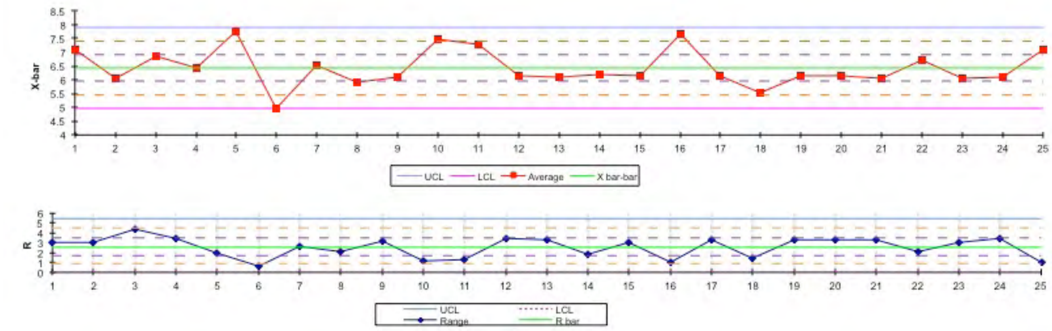

Figure 6. X-bar and R control charts for the XYZ example process with 25 samples.

This example described control chart interpretation at a basic level. The patterns on control charts can actually reveal much more about a process. With practice and experience, you will get better at interpreting and using the charts to improve processes. For example:

- If the R chart indicates the process is out of control because points are above the upper control limit but the X-bar chart is in control, this indicates that variability has increased. Experience with the process may give you an idea of what typically leads to this situation. For part dimensions, this could be an indication of dull tooling.
- If the R chart indicates the process is in control but data points on the X-bar chart are gradually trending downward or outside the lower control limit, this is an indication that the process average has decreased. For moisture content, this could be a sign that relative humidity in the building has decreased significantly (perhaps someone turned on the air conditioning?).
- If both charts indicate the process is out of control, there may be several issues to explore and address.

Again, experience using control charts combined with knowledge of your processes will help you reap the greatest benefits from SPC and guide your continuous improvement efforts.

Next steps: Monitor and take action

Now that the XYZ team has some assurance the process is in control and has reliable estimates of centering and variability, they need to monitor it for several more days to be sure it remains stable.

The estimates of centering and variability ($\bar{\bar{x}}$ = 6.49% and $\bar{R}$ = 2.6 for this example) are critical for monitoring. We use these values to plot control limits on the next series of charts. Once we have established control (stability), we have **trial control limits** for the X-bar and R charts. Next, we draw a new control chart using the trial control limits and collect another set of 25 samples. If that chart and several more (perhaps 3 or 4 days' worth) indicate the process is in control at these limits, we have even greater confidence in where the process is centered and the magnitude of the variability.

What if the next control chart does not show the process is in control at these values? Then we need to identify and address the problems. We repeat this process until the charts indicate an in-control process.

Remember: Establish control first. Then monitor to ensure the process remains in control.

If the example process exhibits control for a period of time, can the XYZ team be confident that the size-out-of-spec problem is solved? Unfortunately, it's not that simple. Recall that the team's goal was to control moisture content at 6%. If the process stays in control at 6.5% (as shown in Figure 6), is that OK? And is an average range of 2.6 acceptable? Control charts can't help answer these questions. For this, we need to do a process capability analysis (PCA)—a technique to compare the variability of an in-control process to specifications, expectations, or requirements. We cover PCA in Part 9 in this series.

A note about specifications

How should the XYZ team establish these specifications? What level of variability is acceptable?

We now know the XYZ team can control moisture content at 6.5% with a range of 2.6. The d_2 table value for samples of five is 2.326. Using these values and the $\overline{R}/d_2$ formula, we can estimate the process standard deviation:

$$1.12\% = (2.6/2.326)$$

And given what we know about the normal distribution, we can expect that more than 99% of our moisture content readings should fall between about 3% and 10% (mean plus and minus three standard deviations):

$$6.5 \pm (3 \times 1.12)$$

So we could conduct another experiment with those moisture content values for birch and poplar and measure the number of out-of-spec handles. If the number of out-of-spec handles is acceptable, we would recommend that the XYZ team continue to monitor the process to ensure it stays at these values ($\overline{\overline{x}} = 6.49\%$ and $\overline{R} = 2.6$).

If, however, the experiment indicates that this target value is too high or that there's too much variability, we would recommend that the XYZ team engage in continuous improvement to bring the process on target (6% moisture content), reduce variability, or both. Then, they would again use control charts to monitor the process and ensure it stays at these new values.

Summary

This publication described variables control charts. These are the most powerful SPC tools and provide the most information for troubleshooting and process improvement. But along with this power comes some complexity. The next publication in this series will focus on attributes control charts, which are used to control and monitor things such as counts or fractions of nonconformities in a sample. Because you now have a solid understanding of variables control charts, you will find the discussion of attributes control charts to be more straightforward.

Formulas used in this publication

Average (x):

Add (sum, Σ) all the individual values (X_i) from the first (i = 1) to the last (the sample size, n). Then divide by the sample size (n).

$$\bar{X} = \sum_{i=1}^{n} \frac{X_i}{n}$$

Range (R):

Subtract the minimum sample value (X_{min}) from the maximum sample value (X_{max}).

$$R = X_{max} - X_{min}$$

Sample standard deviation (s):

In some respects, the standard deviation formula is similar to the average. Subtract the average from each individual value (to calculate the deviation from average), square that value, sum all of the squared deviations, divide by one less than the sample size (n − 1), and then take the square root of the result.

$$s = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}}$$

Formulas for $\bar{x}$ charts using the sample range (R):
Note: A_2 is a table value.

$$Centerline = \bar{\bar{X}}$$

$$UCL_{\bar{X}} = \bar{\bar{X}} + A_2\bar{R}$$

$$LCL_{\bar{X}} = \bar{\bar{X}} - A_2\bar{R}$$

Formulas for R charts:
Note: D_3 and D_4 are table values.

$$Centerline = \bar{R}$$

$$UCL_R = D_4\bar{R}$$

$$LCL_R = D_3\bar{R}$$

Formulas for $\bar{x}$ charts using the sample standard deviation (s):
Note: A_3 is a table value.

$$Centerline = \bar{\bar{X}}$$

$$UCL_{\bar{X}} = \bar{\bar{X}} + A_3\bar{s}$$

$$LCL_{\bar{X}} = \bar{\bar{X}} - A_3\bar{s}$$

Formulas for s charts:
Note: B_3 and B_4 are table values.

$$Centerline = \bar{R}$$

$$UCL_R = B_4\bar{s}$$

$$LCL_R = B_3\bar{s}$$

For more information

The Oregon Wood Innovation Center website provides common table values for SPC: *http://owic.oregonstate.edu/spc*

The listing for this publication in the OSU Extension Catalog also includes a supplemental spreadsheet file that includes all data and charts from this publication: *https://catalog.extension.oregonstate.edu/em9109*

Brassard, M., and D. Ritter. 2010. *The Memory Jogger II: A Pocket Guide of Tools for Continuous Improvement & Effective Planning*. Methuen, MA: Goal/QPC. *http://www.goalqpc.com*

Grant, E.L., and R.S. Leavenworth. 1996. *Statistical Quality Control* (7th edition). New York, NY: McGraw-Hill.

Montgomery, D.C. 2012. *An Introduction to Statistical Quality Control* (7th edition). New York, NY: John Wiley & Sons.

Western Electric Company Inc. 1956. *Statistical Quality Control Handbook*. Milwaukee, WI: Quality Press.

About this series

Publications in the *Performance Excellence in the Wood Products Industry* series address topics related to wood technology, marketing and business management, production management, quality and process control, and operations research.

For a complete list of titles, visit the Oregon State University Extension Catalog and search for “performance excellence”: *https://catalog.extension.oregonstate.edu/*

Published May 2015.