

\bar{x} and s Charts

The development and use of statistical tools known as control charts for monitoring and managing a wide variety of processes has historically been credited to Dr. Walter Shewart. During the 1920s, Dr. Shewart identified two types of variation that prevailed among all processes: chance cause variation and assignable cause variation (Swift, pp 135). Any process will inherently contain chance cause variation, as this is normal, random variation present in all systems. However, a process that contains anything other than chance cause or random variation is exhibiting assignable cause variation and is referred to as “out of statistical control.” Control charts may be used to monitor a process to determine *whether or not* it is in statistical control, to evaluate a process and determine normal statistical control parameters, and to identify areas of improvement in processes (Swift, pp 135 - 136). Control charts may be used to monitor variables that are continuous in nature, or non-continuous attributes. Continuous variables include quantifiable or measurable values that can be calculated across a continuous range, such as averages, dimensions (length, height, width), weight, and temperature. Attributes are not inherently quantifiable, but can be counted, such as number of parts defected, or number of defects per unit inspected (Smith, pp 158 -159). The control chart discussed in this paper is the union of both the \bar{x} (or x-bar) and s charts, and is used to monitor processes with continuous variables.

Rationale for Statistical Process Control Charts

The statistical argument for control charts relies upon the *Central Limit Theorem*. The central limit theorem suggests that even if the underlying population from which a series of observations are gathered is *not* normally distributed, the resulting distribution of averages from the sample observations *will* be normally distributed around an average value, and that 99.73 percent of the values will be contained within three standard deviations of that average value. Using this theorem, a control chart can be constructed that utilizes the statistical mean

as the reference value or centerline on the chart, and an upper and lower control limit equivalent to $\pm 3\sigma$ (three standard deviations) from the statistical mean. If an observed value falls beyond an upper or lower control limit, it can be concluded that this value is in an extreme tail of the distribution (only 0.27 percent of values should fall in this region), and is therefore out of control. Additionally, a series of values above or below the reference line can be evaluated to determine process stability. This will be discussed in more detail at a later point.

Uses of the \bar{x} and s Charts

The \bar{x} and s charts are typically implemented together, as the two charts offer different, but closely related forms of information about a process. The \bar{x} chart relays information about the *central tendency* of the data, or the tendency of the measurements to accumulate in a normal distribution around an average value, or \bar{x} (x-bar). While this is an important statistic to know and understand, it is also critical that the amount of *variation* around a particular average is also known. The s chart, recording the standard deviation from the average value in a given sample, is used for measuring and monitoring variation, and together \bar{x} and s charts can provide insight into the stability of a process or system. For example, a sample of five measurements might have an average value (\bar{x}) of 100 with a standard deviation (s) of 10; while another sample of five measurements might have an average value (\bar{x}) of 100 with a standard deviation (s) of 30. Although the average values are the same in both samples, the sample with s equal to 30 exhibits much more variation. In some processes, variation within broad control limits might be acceptable, but in others, this kind of variation might be undesirable. In some applications, such as in high-tech manufacturing or machine part manufacturing, tolerances for variation are often very limited. The use of the \bar{x} and s charts can be valuable in monitoring these types of processes.

Gathering a Sample and Preparing the Data

Frequently, due to the expensive nature of processes (capital or labor) in which the \bar{x} and s charts are utilized, only small sample sizes are available. As with most statistical analysis, larger sample sizes are *preferred*, but in this case, not absolutely required. For purposes of simplification and demonstration, a sample size, n , equal to five and the number of samples taken, k , equal to ten, will be used. In other words, in each of ten samples ($k = 10$), there will be five observations ($n = 5$). The observations are recorded, and average values (\bar{x}) and standard deviations are calculated for *each sample*. Additionally, average s and \bar{x} values are calculated for the all of the sample averages and standard deviations. The resulting measurements and calculations can be recorded in a grid, such as in *Table 1* (Page 3).

		Sample _k									
		1	2	3	4	5	6	7	8	9	10
Observation _n	1	0.015	0.018	0.017	0.018	0.014	0.017	0.012	0.014	0.015	0.015
	2	0.019	0.017	0.018	0.016	0.013	0.015	0.013	0.015	0.018	0.015
	3	0.022	0.013	0.019	0.012	0.014	0.018	0.015	0.014	0.016	0.018
	4	0.016	0.014	0.014	0.020	0.015	0.019	0.015	0.013	0.014	0.016
	5	0.013	0.015	0.015	0.019	0.017	0.014	0.016	0.012	0.013	0.017
	\bar{x} -bar	0.017	0.015	0.017	0.017	0.015	0.017	0.014	0.014	0.015	0.016
	s	0.004	0.002	0.002	0.003	0.002	0.002	0.002	0.001	0.002	0.001

Average \bar{x} -bar 0.016

Average s 0.002

Table 1: Completed grid with measurements and calculations

The calculation of \bar{x} and s in each sample can be accomplished by using formulas in a spreadsheet package, such as Microsoft Excel™ [\bar{x} : =AVERAGE(cell range:cell range), and s : =SDEVA(cell range:cell range)] or by utilizing common statistical equations for these values (not discussed here).

Calculating Upper and Lower Control Limits

The calculated *average x-bar* and *average s* values are used as the centerlines, or reference values on each chart, against which the average values from each sample are plotted. The last step in preparation for configuring the control charts is the calculation of the upper and lower control limits for both the \bar{x} and s charts. The first set of control limits calculated is the set for the s chart. The reason these are calculated in advance of the \bar{x} control limits is twofold: first, to determine if the process is in statistical control with regard to variance; and second, to provide insight as to why the process is out of control. Once the process has been brought under control with regard to variance, the \bar{x} chart can be constructed (Smith, pp 237). The formulas for calculating the control limits for each chart are derived through a lengthy statistical proof that is based, as previously referenced, on the application of the *Central Limit Theorem*. The end result of the proof demonstrates that calculated statistical constants, based on the number of observations in each sample, may be substituted for more complicated formulas when determining the Upper Control Limit (*UCL*) and Lower Control Limit (*LCL*). The statistical proof and equations are not discussed here, as this documentation emphasizes the applied techniques,

rather than the details of the theory. The constants are reproduced in Table 2 (Foster, pp .365), through n values of six. Using the constants in

n	B_3	B_4	C_4	A_3
2	0.000	3.267	0.7979	2.659
3	0.000	2.568	0.8862	1.954
4	0.000	2.666	0.9213	1.628
5	0.030	2.089	0.9400	1.427
6	0.118	1.970	0.9515	1.287

Table 2, the formulas for calculating the UCL and LCL for the s chart are described below:

Table 2: \bar{x} -bar and s chart values

$$UCL = B_4(\bar{s}) \quad \text{Calculate the Upper Control Limit for the } s \text{ chart} \quad (1)$$

$$LCL = B_3(\bar{s}) \quad \text{Calculate the Lower Control Limit for the } s \text{ chart} \quad (2)$$

Using the values calculated in Table 1, and the constants from Table 2, the following s chart control limits can be calculated:

$$UCL = (2.089)(0.002) = 0.00418 \quad (1a)$$

$$LCL = (0.030)(0.002) = 0.00006 \quad (2a)$$

As previously stated, after the control values for the s chart are calculated, the control values for the \bar{x} chart are determined, using the following equations:

$$UCL_{\bar{x}} = \bar{\bar{x}} + A_3(\bar{s}) \quad (3)$$

$$LCL_{\bar{x}} = \bar{\bar{x}} - A_3(\bar{s}) \quad (4)$$

Again, using the values calculated in Table 1, and the constants from Table 2, the following \bar{x} chart control limits can be calculated:

$$UCL = 0.016 + [(1.427)(0.002)] = 0.0189 \quad (3a)$$

$$LCL = 0.016 - [(1.427)(0.002)] = 0.0131 \quad (4a)$$

With the centerline values, upper control limits, and lower control limits all calculated, the \bar{x} and s charts can be constructed, and sample values plotted.

Constructing the \bar{x} and s charts

Typically, the construction of the \bar{x} and s charts is done together, so that the plotted values can be compared at even intervals. Below, the two charts are illustrated in Figures 1 and 2:

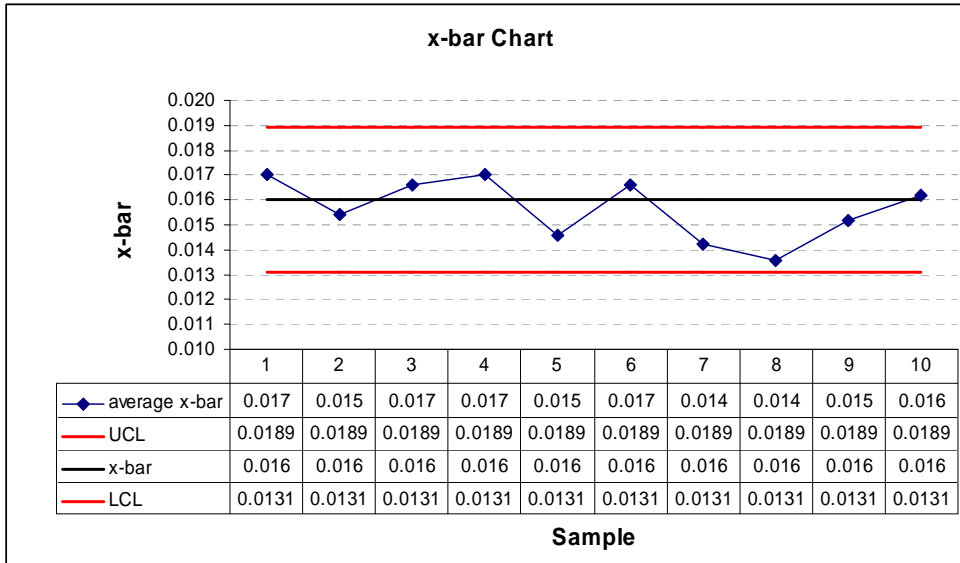


Figure 1: x-bar Chart

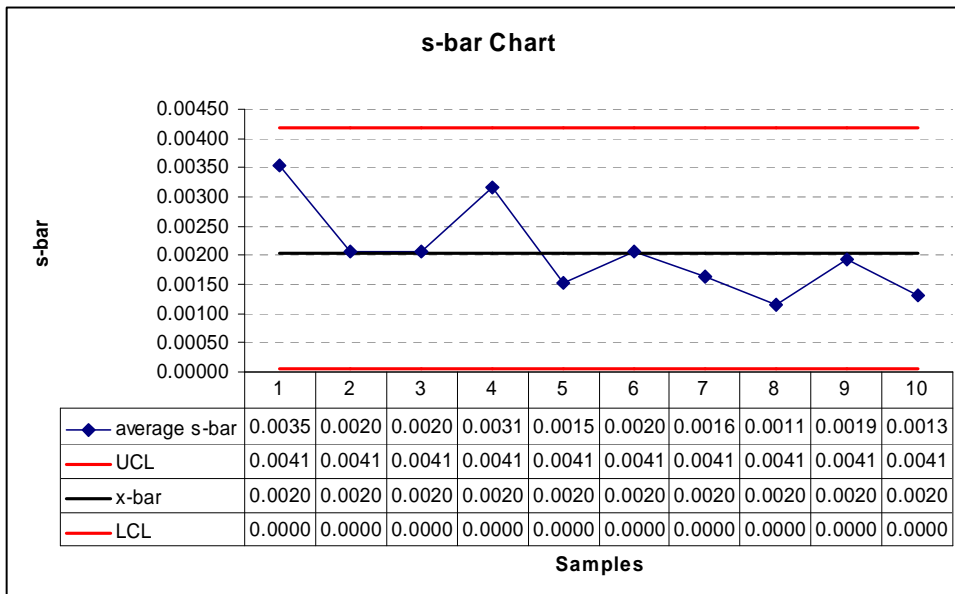
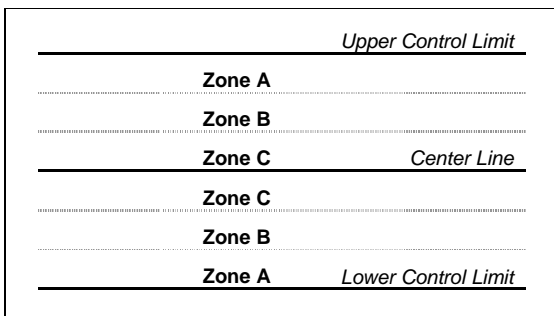


Figure 2: s-bar Chart

Assessing the Results

With respect to control charts, recording the observations, calculating the control values, and plotting the observed values on the \bar{x} and s charts is only the beginning of the process. Once the chart is constructed, and data points are plotted, the chart needs to be evaluated to determine whether or not the process is operating in statistical control. There are several guidelines for evaluating control charts, all with a prevailing philosophy: the plotted points should only exhibit random variation that cannot be fitted with an identifiable or quantifiable pattern. Additionally, the following features, in accordance with the parameters of a normal distribution and the *Central Limit Theorem*, should also apply (Swift, pp 160):

1. 68 percent of the points should be within $\pm 1\sigma$ of the reference line.
2. 4.27 percent of the points should be between $\pm 2\sigma$ and $\pm 3\sigma$ of the reference line.
3. No more than 0.27 percent of the points should exceed $\pm 3\sigma$ of the reference line.



In addition to the above general philosophy of evaluating the control chart, an additional set of guidelines, described in Swift (pp 161-163) are adapted from the *AT&T Rules*, as described in the *AT&T Statistical Quality*

Figure 3: AT&T Zone Chart
Control Handbook. The AT&T Rules divide the control chart into three zones, mirrored across the center line. This is illustrated in Figure 3. According to these guidelines, a process can be ruled “out of statistical control” if any of the following apply:

1. Any point falls outside of the upper or lower control limit (beyond $\pm 3\sigma$).
2. Two of three successive points in zone A or beyond..
3. Four out of five successive points on zone B or beyond.
4. Eight successive points fall in zone C or beyond, on one side of the center line.

(list continues on page 8)

5. *Trends*: a series of points without any appreciable or random change in direction, or values moving continuously up and down, or across the centerline in pattern.
6. *Cycles*: short trends in which the data may repeat in a pattern.
7. *Shifts*: a sudden change in level in one direction or another.
8. *Stratification*: a pattern of “unnatural consistency” within a single zone, or near the centerline.
9. *Systematic Variables*: a predictable pattern, where a high point is always followed by a low point, or a low point by a high point.

Evaluating the Example Case

As the rules and guidelines are applied to the example provided in this document, it is apparent that the s chart is exhibiting signs of instability, or that the underlying process is out of statistical control. No values ever exceed control limits, however, there appears to be a sustained trend in variation from the positive side of the centerline to the negative or lower side of the control line. When checked against the \bar{x} chart, it is less apparent that the system is out of control. The \bar{x} and s chart together suggest that while the process output is consistently clustered around the mean value, the variation is out of control and needs to be addressed.

Finding More Information About \bar{x} and s Charts

Most basic statistics or statistical process control reference books or textbooks contain a section relating to control charts and how to implement them. The internet can also provide an excellent resource for obtaining information about statistical process control charts. Due to the rapidly changing nature of the internet, a website is not listed here, however, a search on a popular search site for “x-bar and s charts” yielded several thousand possible references. Additionally, in the attached reference section are several textbooks that were valuable in compiling this document.

References

- AT&T Statistical Quality Control Handbook*. Charlotte, NC: Delmar, 1985
- Foster, S. Thomas (2004). *Managing Quality: An Integrative Approach, 2nd Edition*. Upper Saddle River, NJ: Prentice-Hall
- Groebner, David F., Shannon, Patrick W., Fry, Phillip C., Smith, Kent D. (2000). *Business Statistics: A Decision Making Approach*. Upper Saddle River, NJ: Prentice-Hall
- Montgomer, Douglas C. (1991). *Introduction to Statistical Quality Control, 2nd Edition*. New York, NY: John Wiley & Sons
- Smith, Gerald M. (1998). *Statistical Process Control and Quality Improvement, 3rd Edition*. Upper Saddle River, NJ: Prentice-Hall
- Swift, J.A. (1995). *Introduction to Modern Statistical Quality Control and Management*. Delray Beach, FL: St. Lucie Press