DAVID S. SCHWARZ & ASSOCIATES

Statistical Consultants

4261 Americana Drive
Suite 229
Stow, OH 44224

Phone: (330) 945-4733
Fax: (330) 945-4733
E-Mail: david.schwarz@goodstats.biz

Calculating The Process Standard Deviation For Individual (I) SPC Charts

Individual or I Charts are a commonly used chart in manufacturing environments. In such situatations, the sample size for a given run is one (1) so it is not possible to calculate the sample standard deviation (as it is for the X-bar ($\bar{X}$) chart where multiple samples are taken). This leaves one with the problem of how to calculate σ and the 6σ interval.

One approach would be to use the standard, independent sample, calculation; or what is sometimes called the 'Excel' method.

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

where: $\bar{X}$ is the sample mean, and n is the sample size.

This calculation method assumes that each data value is independent of every other value in the set. In charting methodology, this translates to having a process that is in statistical control and all variability observed is just random. Unfortunately, that is not the case in most practical situations. Most processes tend to drift, trend, or shift to some degree over time. If one uses the above calculation with data of that sort, then the standard deviation will be inflated because the shifts, etc. will be included as part of the natural variability, producing 6σ intervals that are too large.

To overcome this problem, charting routines use the MR(2) method to calculate the standard deviation. MR(2) stands for moving range (of 2) and is computed by first calculating the successive ranges of each data point with the data point that precedes it. The absolute value of the range is calculated followed by additional calculations that get us to the process standard deviation, σ. This number is an unbiased estimate of the process standard deviation. The table and formulas below demonstrates this calculation. Column 1 is just the raw data example. Column 2 is the same data lagged by one cell. Column 3 is the moving range for each point calculated as the absolute difference between Columns 1 and 2.

x_i	x_{i+1}	$MR_i = \lvert(x_i - x_{i+1})\rvert$
1371	-	-
1394	1371	23
1369	1394	25
1327	1369	42
1309	1327	18
1420	1309	111

$$\overline{MR} = \frac{\sum_{i=1}^{n-1} MR_i}{n-1} \quad \text{and} \quad \sigma = \frac{\overline{MR}}{1.128}$$

For our example data in the table above:

$$\overline{MR} = \frac{(23+25+42+18+111)}{5} = 43.8 \quad \text{and} \quad \sigma = \frac{43.8}{1.128} = 38.8$$

An Excel calculation produces: $\sigma = 42.2$, a somewhat higher value (The value 1.128 above is called Hartley's constant and is used in all such MR(2) calculations. Some texts refer to it as d_2).

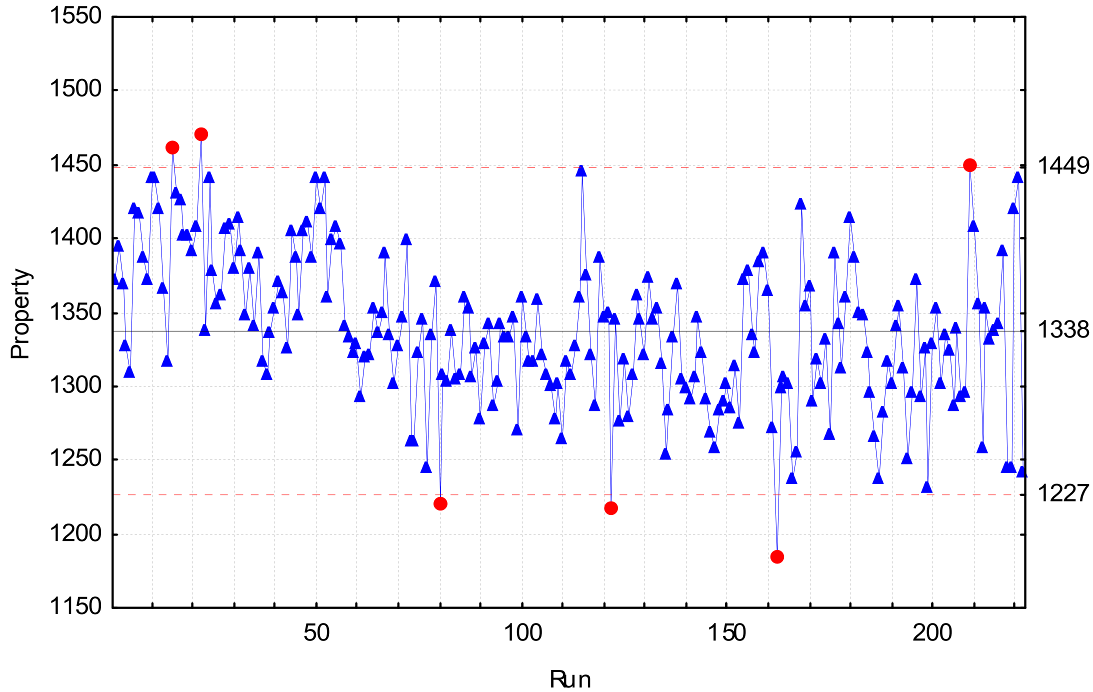
Applied to a quality control data set, this calculation would produce a value of sigma that is unbiased and devoid of the inflationary effects drift, trend, or shifts in the data. It is, in effect, the purely random component of the process variation; sometimes called the 'white noise' variability. This is the value of σ that must be used to calculate the 6σ interval on all Individual or I charts.

This explanation can be further motivated by examining two Individual charts. Graph 1 below, shows a typical industrial process plotted on an individual chart. Note the downward shift in the process at about run 60. This results in some points above the upper statistical limit before run 60 and some below the lower statistical limit after run 60. The value of sigma on the chart is 37.0. This was calculated using the MR(2) method above and again, represents only the random variation in the data (not the shift).

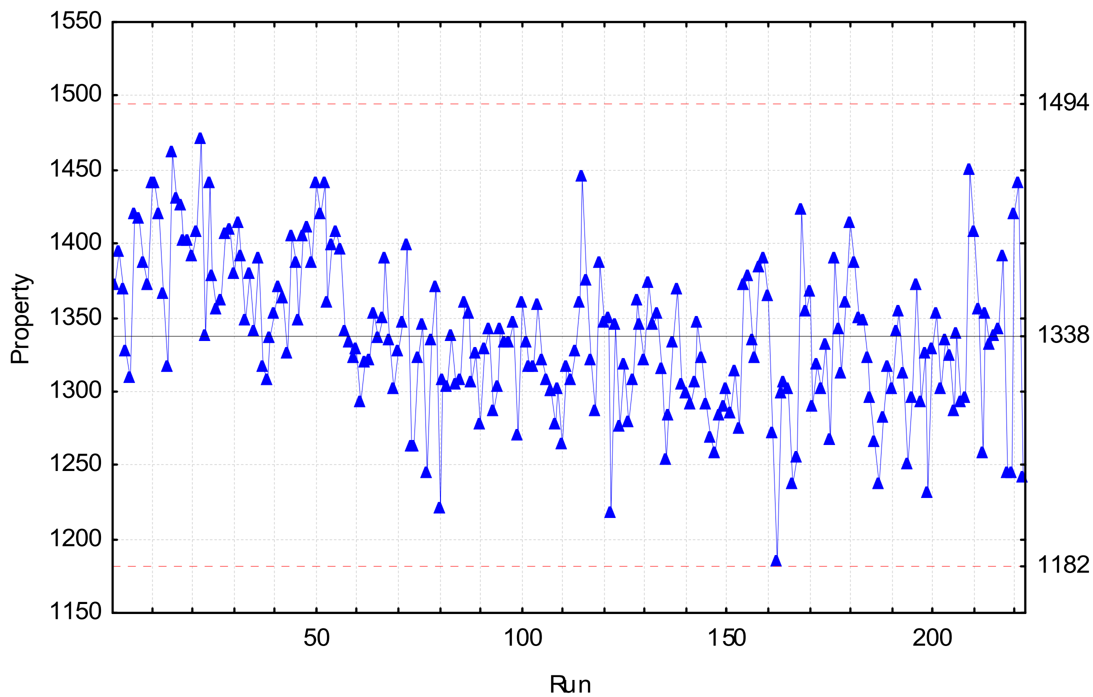
Graph 2 shows the same process except that a sigma value of 52.1 was used. This is the 'Excel' value. Note that all of the data is inside of the statistical limits which would lead one to believe that there are no problems with the data. However, we can observe that problematic shift at around point 60. The 'Excel limits' will almost always encompass all of the data, including any abnormalities, and hence, are of little diagnostic value.

Further insight is provided by Graph 3. This is an I Chart of the process in Graph 1 except that it is simulated so that all of the drifts, trends, and shifts are removed. This is what the process should look like if it were in perfect statistical control. Note that there are no points outside the statistical limits and no abnormal patterns in the data. This is the 'white noise only' of the process.

Graph 1
Process Charted Using the MR(2) Calculation For Sigma
sigma = 37.0



Graph 2
Process Charted Using the 'Excel Method' For Calculation Of Sigma
sigma = 52.1



Graph 3
A Simulated 'In Control' Chart of the Process in Graph 1
sigma = 37.0