

## BaseLine© Examples

### H. Interpreting "special causes" of variation – triggered by RED FLAGS

Whenever you wish to remind yourself of what rules are being used to trigger Red Flags, just click on **HELP** and then the "Introduction" tab.

All 4 of the following Western Electric Rules for detecting "Special Cause" variation are used in the **Ind** Chart, and Rules #1 and #4 are used in the **mR** Chart:

**Detection Rule #1** One point more than 3 sigma from the mean.

**Detection Rule #2** Two out of three points more than 2 sigma on the same side of the mean.

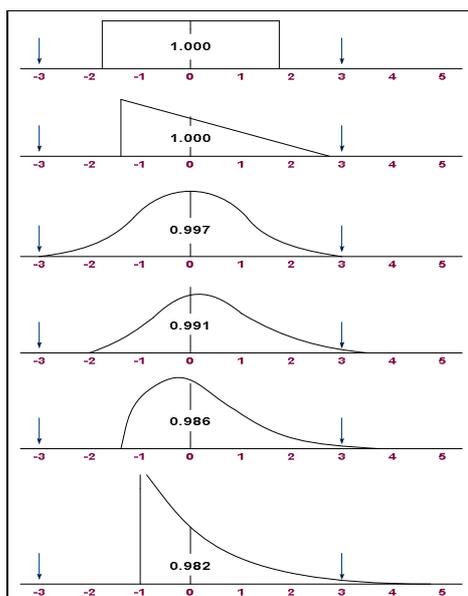
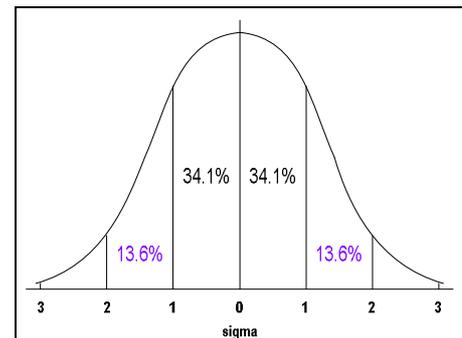
**Detection Rule #3** Four out of five points more than 1 sigma on the same side of the mean.

**Detection Rule #4** Nine points on the same side of the mean.

Additionally, an explanation should be sought any time a pattern seems to repeat itself eight times or more.

**The Empirical Rule** states that for a "normal distribution" more than 2/3 of the data will lie within  $\pm 1$  sigma, and that nearly all your data will lie within 3 sigma of the centre line.

The detection rules, individually and collectively aim to keep the chance of a false alarm to less than 1%.



The "normal distribution" is a theoretical model and doesn't exist in real life, and therefore every data set you come across will be to some extent skewed. However, the XmR Chart, together with the detection rules used in BaseLine©, work in almost all situations – even with extreme distributions like the ones shown opposite.

There are many different types of SPC chart, but the most ubiquitously robust is the Individuals and Moving Range (XmR) chart, the one used exclusively in BaseLine©.

**Note:** BaseLine© automatically flags Special (assignable) Causes of variation, according to the principles originally laid down by **Walter Shewhart** and codified in the Western Electric Company's Statistical Quality Control Handbook, first published in 1956.

There are many other rules used by expert SPC users, but it is all too easy to go overboard with detection rules, and forget the main objective which is to minimize the risks of the 2 mistakes – one is delusion and the other is oversight:

**mistake #1** is to falsely interpret noise as if it is a signal.

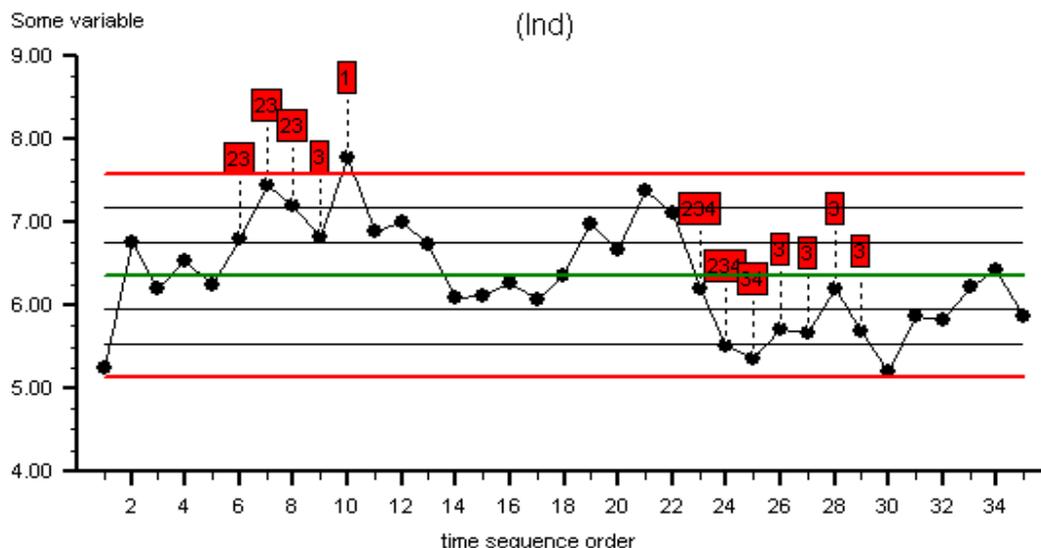
**mistake #2** is to miss a signal when one is present.

The more rules you use in your quest to avoid oversight, the more you increase the delusional risk of mistake #1. Of course no software can know the context, and the main thing is always to interpret your chart in the context of the system itself, which is why we emphasize that BaseLine© is designed to tee-up and guide insightful conversation.

The **Help** tab also gives examples of some of the other rules that some people use but BaseLine© avoids in order to achieve the best likely balance between the two mistakes.

Here is an **Ind** chart before the data has been split. The "show sigma levels"  option has been checked in order to illustrate each of the rules in operation. Each red flag contains within it the number of the rule that has been triggered. The first flag **23** at **point 6** indicates that both rules #2 and #3 have been triggered. This is because point 6 is the **beginning** of a run of three points (6,7 and 8) on the same side of the mean, and two of those points are more than 2 sigma from the centerline shown in green.

And point 6 is also the **beginning** of a run of five points (6,7,8,9 and 10) on the same side of the mean, and four of those points are more than 1 sigma from the centerline.



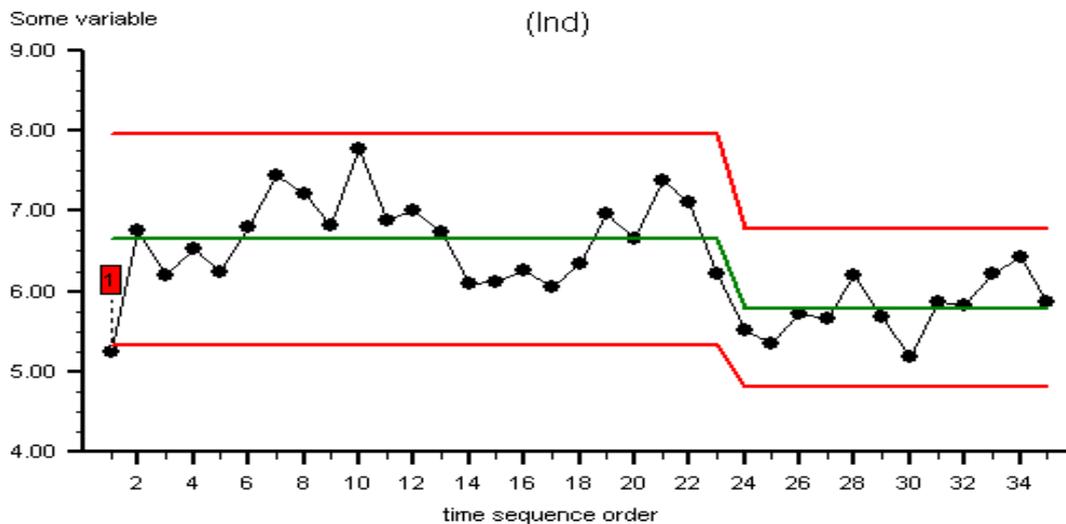
**Point 10** is flagged **1** to indicate that Rule #1 has been triggered.

**Point 23** is flagged **234** to indicate that Rules #2, #3 and #4 have all been triggered. This is because there are more than nine points (actually eleven) on one side of the centerline, and point 23 is at the **beginning** of that run. Point 23 is also signaled as a special cause because two of the next three points are beyond 2 sigma, and four of the next five points are beyond 1 sigma. If this data had been monitored in real time, point 23 would have been flagged as special by Rule#2 alone.

## How best to identify the point at which a systemic shift has probably occurred.

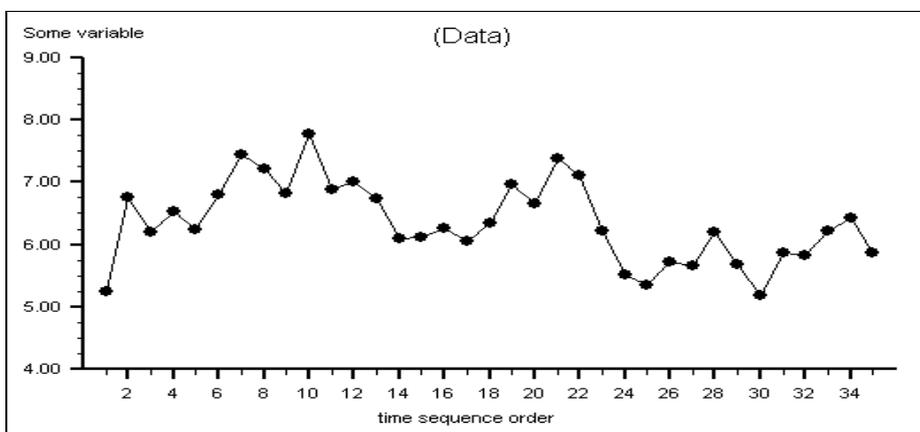
In **Example D** we saw how to detect systemic change by removing special cause FLAGS and eyeballing changes in both the centre-line and the level of variation. Here we want to go further using the example above to highlight just how to choose the precise moment at which your system has in all likelihood changed – at that this is often as much an art as a science, and a task best undertaken by the people who have most understanding of the particular context pertaining to the data.

Starting with the most recent data i.e. working Right to Left, I try splitting the data at **point 24**.

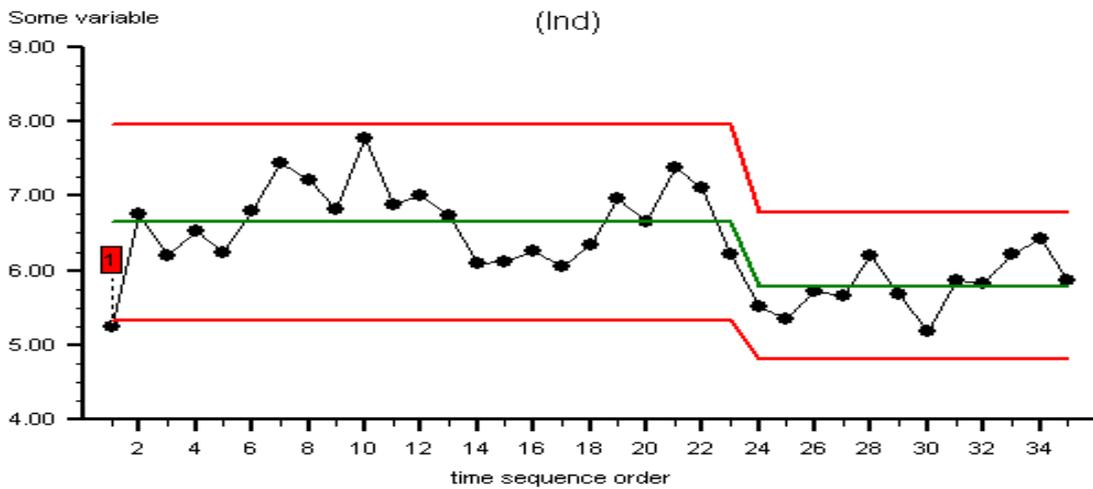


This has the effect of eliminating all the red flags that were bunched to the right of the chart, including the two preceding **points 22 and 23**. But why did I not choose either of those points as the best place to split the data?

I'm actually keeping an eye simultaneously on both the level of variation and the mean. I'm suspecting that the system changed and was reflected mostly in the average, so using my intuition I decide to split the data at where I believe the new system appears to have begun i.e. **point 24**. This is perhaps easier to see by clicking on the **Data** screen to view the data in unexpurgated form:



Having made the split at point 24, a **Rule #1 special cause** is then triggered at **point 1**.



This point does seem to be from a different system to the one following from point 2 onwards, so even without knowing the context, or being able to give a reason, it could well be sensible to exclude this first data point from the limits calculation.

BaseLine 1.00 from www.SAASoft.com Project: "Recipes XmR - showing flag use 10-04-15.bpf"

### Data

X Order: time sequence orde Y Unit Name: Some variable  
 X Format: Y Upper Spec: 0.000  
 Y Lower Spec: 0.000

Order	Date/Time	Group	Metric	mR	Special
12			7.010	0.120	
13			6.730	0.280	
14			6.090	0.640	3
15			6.120	0.030	
16			6.260	0.140	
17			6.060	0.200	
18			6.350	0.290	
19			6.970	0.620	
20			6.660	0.310	
21			7.380	0.720	
22			7.110	0.270	
23			6.210	0.900	
24			5.520		
25			5.350	0.170	
26			5.720	0.370	
27			5.660	0.060	
28			6.200	0.540	

Order: 26 Time: Time Group: Metric: 5.720

Add Set Ins Excl Incl Del Spl Mrg Lock L2S

Help User Guide Copy Paste Append Undo

### Individuals Chart

Some variable (Ind)

Segment

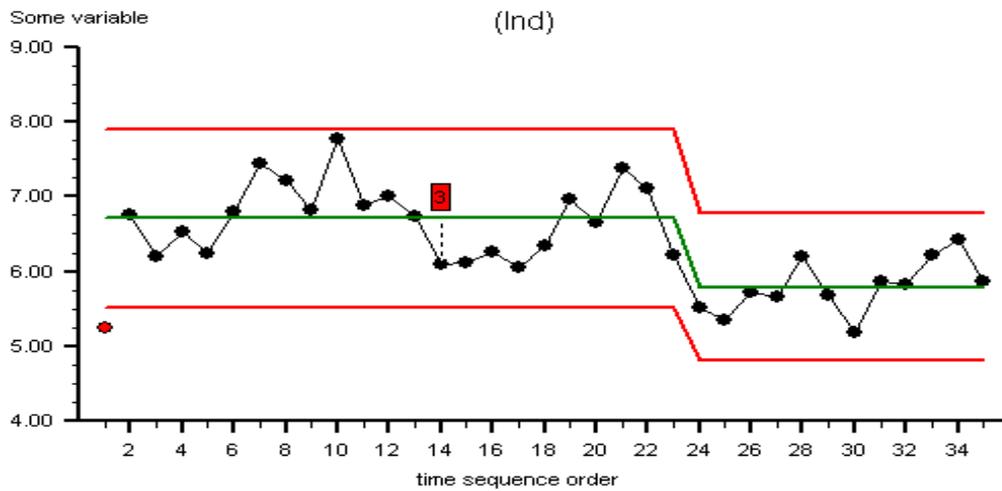
Segment	Start	Finish	Mean	Sigma	Count	LCL	UCL	Stable
1	1	23	6.710	0.396	22	5.524	7.897	No
2	24	35	5.793	0.326	12	4.817	6.770	Yes

Sigma Levels 
  Specification 
  Histogram 
  Groups 
  Flags

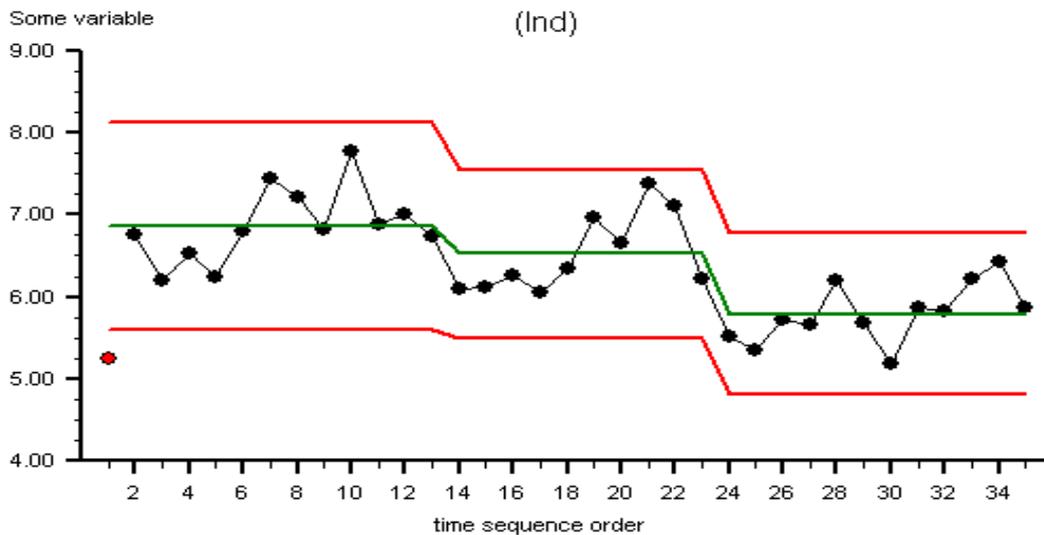
Data Run Ind mR Notes New Load Save Export Exit

And should you feel that an explanation is valid, a note can be recorded. This can be done by right clicking on the point and opting for **"Add Note"**. Simply type your note in the space indicated by the cursor, and click **"Save Note"**. The note will then appear when ever you hover over the point.

**Excluding point 1** has an interesting effect. The centreline and limits calculation changes sufficiently for **point 14** to emerge as a special cause, triggered by **Rule #3**.



Again, using a combination of your eyeballed intuition, and the flag, makes it possible to argue that the system has indeed experienced a change at **point 14**. Splitting the data at gives us the following chart:



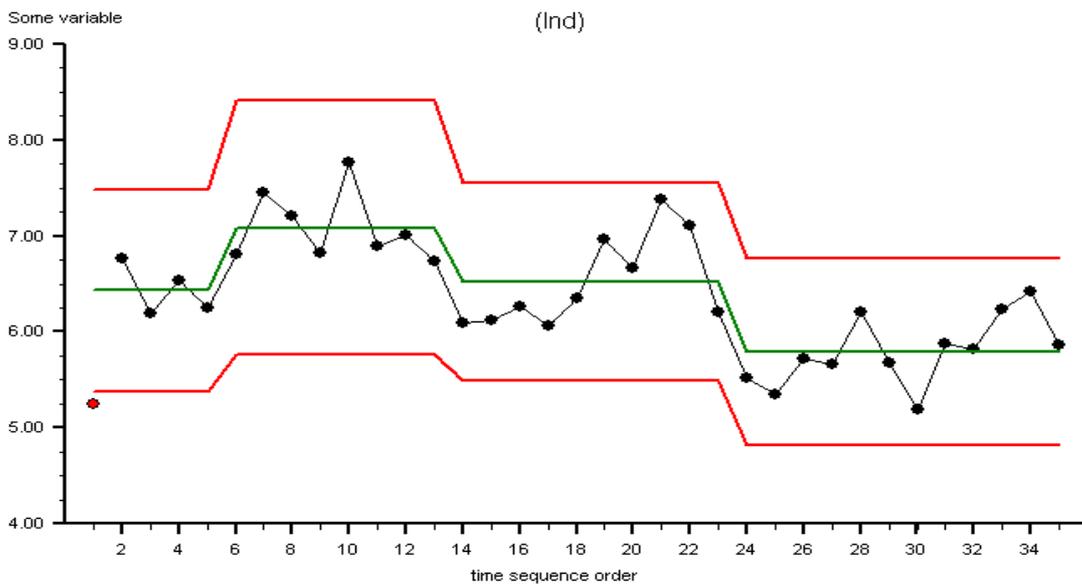
**Summary**

Segment	Start	Finish	Mean	Sigma	LCL	UCL	Stable
1	1	13	6.868	0.423	5.599	8.138	Yes
2	14	23	6.521	0.343	5.493	7.549	Yes
2	24	35	5.793	0.326	4.817	6.770	Yes

The chart is now ready for a wider conversation – with those who are connected to the system the chart represents, and who know something of the context. But wait.. what if someone argues that there was another shift at **point 6**, even though it's not one that has been flagged?

Again it's important to state that the flags should be treated as guides for your intuition. Without the flags, most people are liable to over-interpret the system's changes, and end-up deluding themselves. But without intuition, too much weight is awarded to the rules, that when applied pedantically can miss important aspects of the story that those who know the system's context are bringing to the party.

So, depending on how well you feel the argument has been made, you may legitimately split the data one more time at **point 6**. Here is what the chart would then look like:



**Summary**

Segment	Start	Finish	Mean	Sigma	LCL	UCL	Stable
1	2	5	6.433	0.352	5.378	7.487	Yes
2	6	13	7.086	0.443	5.756	8.416	Yes
3	14	23	6.521	0.343	5.493	7.549	Yes
2	24	35	5.793	0.326	4.817	6.770	Yes

**Specific Notes**

- 1 This point has been excluded because ..