

IE 361 Module 16

What if the Sample Size is 1?

Prof.s Stephen B. Vardeman and Max D. Morris

Reading: Section 4.4, *Statistical Quality Assurance Methods for Engineers*

(Section 3.2.3 Revised *SQAME*)

To call n observations a "sample" or a "rational subgroup" is to implicitly assume/guarantee that they were collected under essentially constant process conditions. But there are many process monitoring contexts where it is simply not safe to assume that even two successive process outcomes are necessarily generated under the same conditions. (This is particularly the case in low volume production contexts where there is a substantial time interval between observations.) That is, sometimes the only reasonable sample size in process monitoring is $n = 1$. This brief module concerns two issues that arise when sample sizes are 1

- the possibility of control charting moving ranges
- what can be done towards the estimation of σ

X/MR Charting (... A Generally Bad Idea)

Where rational subgroups are of size $n = 1$ there is really only one possible choice for a plotted statistic Q , namely x . One can "transform" the most natural measurement to some other scale, but ultimately it is $Q = x$ that is available for plotting. However, people used to making \bar{x} and R chart (or \bar{x} and s chart) pairs in cases where $n > 1$ sometimes reason that it might be useful to supplement an x chart with a chart on which one plots **moving ranges**

$$MR_i = |x_i - x_{i-1}|$$

The most common version of this is one where standards given control limits are

$$UCL_x = \mu + 3\sigma \quad \text{and} \quad LCL_x = \mu - 3\sigma$$

together with

$$UCL_{MR} = D_2\sigma$$

(for D_2 based on the pseudo-sample size of $n = 2$). This practice turns out to produce a large "false alarm rate." That is, it turns out that the all-OK ARL for this scheme when the process distribution is normal is only 105!

A possible "fix" of this initial objection to standard practice is to replace the constants 3 and D_2 with other (typically larger values), say \mathcal{M} and \mathcal{R} chosen to produce a better all-OK ARL based on control limits

$$UCL_x = \mu + \mathcal{M}\sigma \quad \text{and} \quad LCL_x = \mu - \mathcal{M}\sigma \quad \text{and} \quad UCL_{MR} = \mathcal{R}\sigma$$

If you are interested in this possibility, Section 4.4 of *SQAME* provides tables of \mathcal{M} and \mathcal{R} that will produce desired all-OK ARLs. (Also in that section are methods for evaluating (not-all-OK) ARLs (mean times to detection).)

However, we can **not** recommend that you use the " $X|MR$ " combination *even after correcting it so that it has reasonable all-OK properties*. That is because careful study of ARLs for these charts shows that once one restricts to a given all-OK ARL, " $X|MR$ " combinations do poorly in comparison to a competing x chart. (They have only marginally better ability to detect a change in process standard deviation, bought at the price of a serious degradation in ability to detect changes in process mean.) Adding a moving range chart to an individuals chart just turns out to be a bad idea that should disappear from practice.

There is, to our knowledge, only one circumstance in which adding the moving range chart to an individuals chart makes sense. That is the case where the departure from all-OK process behavior that one fears and needs to detect is one of *oscillation* in consecutive individuals. There, a moving range chart is (not surprisingly) more effective than the individuals chart at "seeing" the non-standard process behavior.

To be absolutely explicit, **in cases where $n = 1$, the best thing to do is typically to use only an x chart** (the $n = 1$ version of the \bar{x} chart) with corresponding control limits

$$UCL_x = \mu + 3\sigma \quad \text{and} \quad LCL_x = \mu - 3\sigma$$

Estimating Process Standard Deviation, σ

When $n = 1$, there are no sample ranges or standard deviations to use in estimating σ . In fact, there is no really honest way to estimate a process standard deviation unless one has a sample or samples with $n \geq 2$. But some dishonest methods are less dishonest than others, and the best known method

(least dishonest method) is based on an average of moving ranges of successive observations. (NOTICE that this is NOT a matter of process monitoring based on moving ranges, but rather using moving ranges to estimate process standard deviation.)

The rationale for using moving ranges of successive observations in estimating σ is this. If process conditions can change observation to observation, observations will vary not only because $\sigma \neq 0$, but because the process mean changes. However, it is not unreasonable to expect the variability in pairs of *successive observations* to be **less affected** by mean changes than the variability of any other type of group of observations that could be made up. It is thus reasonable to use moving ranges to make an estimate of process standard deviation, that while potentially inflated by variation in the process mean, can be expected to be *less so* than any other estimate one might make.

The exact form of estimator of σ we'll use (based on samples of size $n = 1$) is

$$\hat{\sigma} = \frac{\overline{MR}}{d_2}$$

where d_2 is for "sample size" 2 (as there are 2 observations represented in each moving range). This is a *conservative* estimator, as it will tend to over-estimate σ when μ is not constant. But it is the best one available.

Example 16-1 Consider the 8 successive observations in the table below and the corresponding 7 moving ranges.

Sample	1	2	3	4	5	6	7	8
x	5	3	9	10	17	4	6	2
MR		2	6	1	7	13	2	4

The values 5, 3, 9, 10, 17, 4, 6, 2 *certainly vary* because $\sigma \neq 0$. They *may vary* beyond what can be attributed to inherent process short term variability

if μ is not constant. That is, the 7 moving ranges should not be thought of as honest sample ranges, but as potentially over-representing σ . Nevertheless, the best available estimate of σ in this $n = 1$ context is

$$\begin{aligned}\hat{\sigma} &= \frac{\overline{MR}}{d_2} \\ &= \frac{(2 + 6 + 1 + 7 + 13 + 2 + 4) / 7}{1.128} \\ &= 4.43\end{aligned}$$