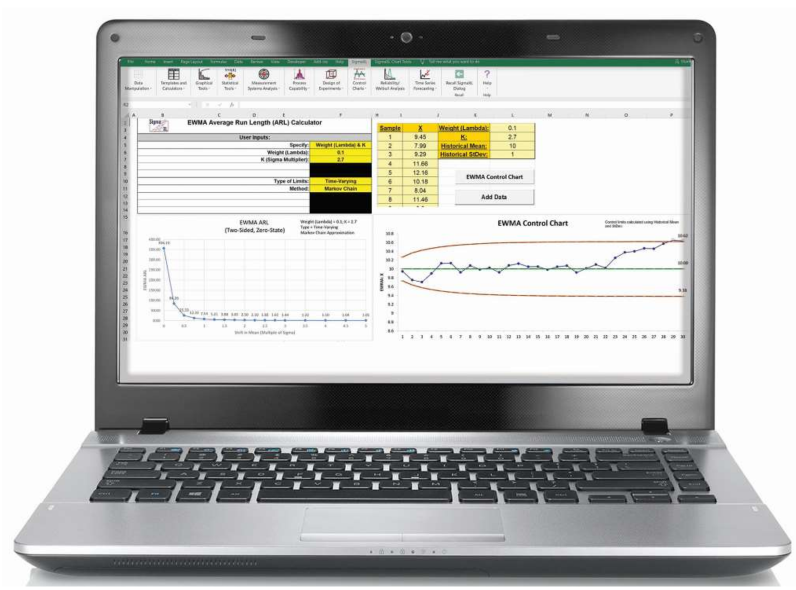
Lean Six Sigma Statistical Tools, Templates & Monte Carlo Simulation in Excel

What's New in SigmaXL® Version 9

Part 1 of 3: New and Improved Control Charts



John Noguera
CTO & Co-founder
SigmaXL, Inc.
www.SigmaXL.com

Webinar October 15, 2020

What's New in SigmaXL Version 9

Part 1: New and Improved Control Charts

- New Control Chart Templates:
  - Rare Events T, G and Probability-Based G
  - Trend/Tool Wear
  - Exponentially Weighted Moving Average (EWMA)
  - Tabular Cumulative Sum (CUSUM)
- Average Run Length (ARL) Calculators:
  - Shewhart with Tests for Special Causes
  - Attribute C & P
  - EWMA & CUSUM

What's New in SigmaXL Version 9

Part 1: New and Improved Control Charts

- Tests for Special Causes now supported for menu-based control charts:
  - MR/Range/StDev Charts (Tests 1-4)
  - Varying Subgroup Sizes (Moving Limits)
  - Historical Groups
- Tests for Special Causes Defaults dialog updated
- I-MR-R/S menu update

Rare Events T Chart Template

- The Rare Events T Chart (or time-between chart) is an alternative to a standard attribute chart when the adverse event of interest is relatively rare and a measurement of time between each occurrence can be obtained.
- If the rate of occurrence follows a Poisson distribution (the usual assumption for a C or U chart), then the times between occurrences will have an exponential distribution. The exponential distribution can be transformed to a symmetric Weibull distribution by raising the time measure to the (1/3.6) power.
- A Shewhart Individuals chart is calculated on the transformed data and then an inverse transformation is applied to the control limits in order to get back to the original time between units. This is an approximate model to the exponential and will result in asymmetric control limits.
- Note that an "out-of-control" signal above the UCL is desirable, indicating a significant increase in time between adverse rare events.
- The Provost and Murray book, “The Health Care Data Guide: Learning from Data for Improvement” is popular in health care, so SigmaXL uses the control limit calculations as given in the book:

Rare Events T Chart Template

t = time between incidents

z = transformed time $\left[z = t^{(1/3.6)}\right]$

Construct a Shewhart individuals control chart of z values:

$\overline{MR}_z$ = average moving range of z's

Remove moving range outliers (i.e., exceed UCL for moving range) for added robustness:

$\overline{MR}'_z$ = average moving range of z's with any MRs > $(3.27 * \overline{MR}_z)$ removed

$CL_z = \bar{Z}$ (average of transformed time)

$$UCL_z = \bar{Z} + 3 * \left(\frac{\overline{MR}'_z}{1.128}\right)$$

$$LCL_z = \bar{Z} - 3 * \left(\frac{\overline{MR}'_z}{1.128}\right)$$

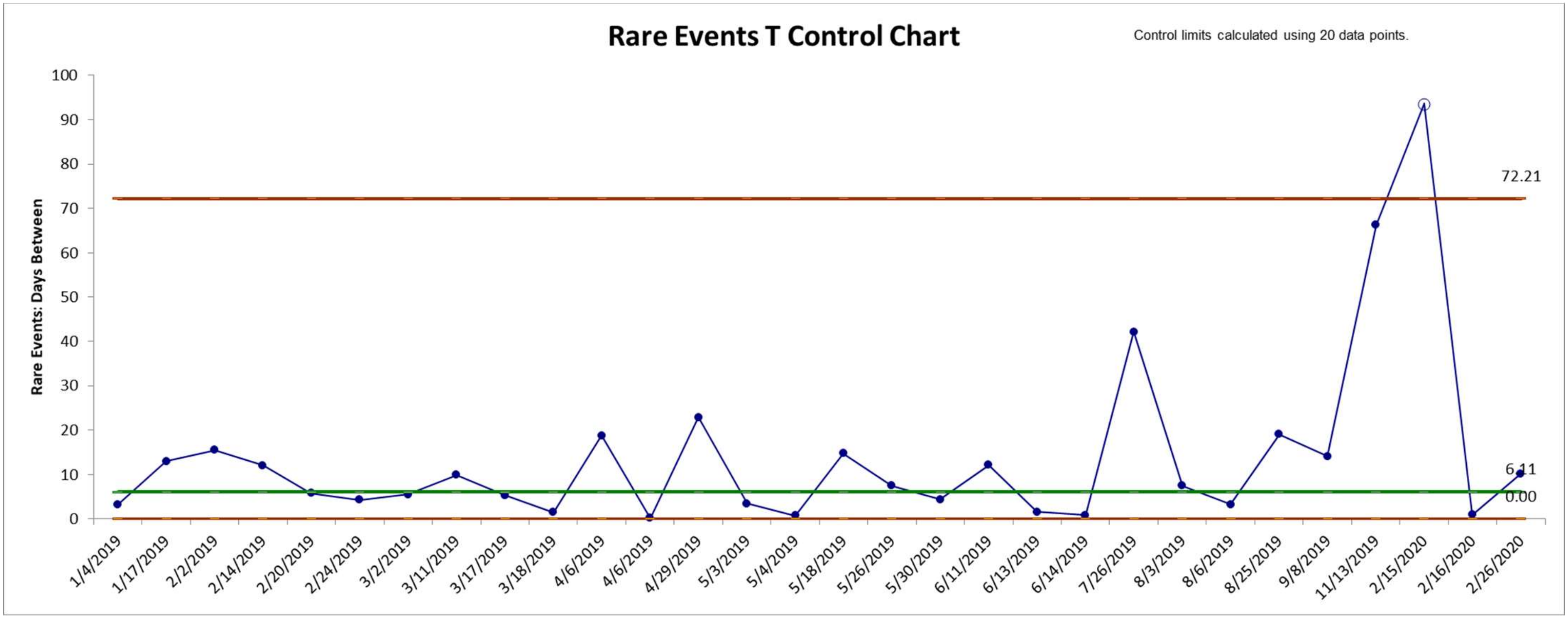
Transform the center line and the limits back to time scale by raising them to the 3.6 power:

$CL_t = {CL_z}^{3.6}, UCL_t = {UCL_z}^{3.6}, LCL_t = {LCL_z}^{3.6}$

Example: Rare Events T Chart Template

SigmaXL > Control Charts > Control Chart Templates > Rare Events > Rare Events T

Data: Days Between Surgical Site Infections.xlsx



Rare Events G Chart Template

- The G chart (or Geometric chart) is an alternative to a standard attribute chart when the adverse event of interest is rare and discrete opportunities between events are counted (e.g., number of units or days between).
- The calculation of control limits (Provost and Murray) is an approximation based on the geometric distribution. An "out-of-control" signal above the UCL is desirable, indicating a significant increase in units/opportunities or days between adverse rare events.

$\mathrm{g} = \text{number of opportunities or units between incidents}$

$\bar{g} = \text{average of g's}$

$CL = 0.693 * \bar{g}$ (CL is the theoretical median for a geometric distribution)

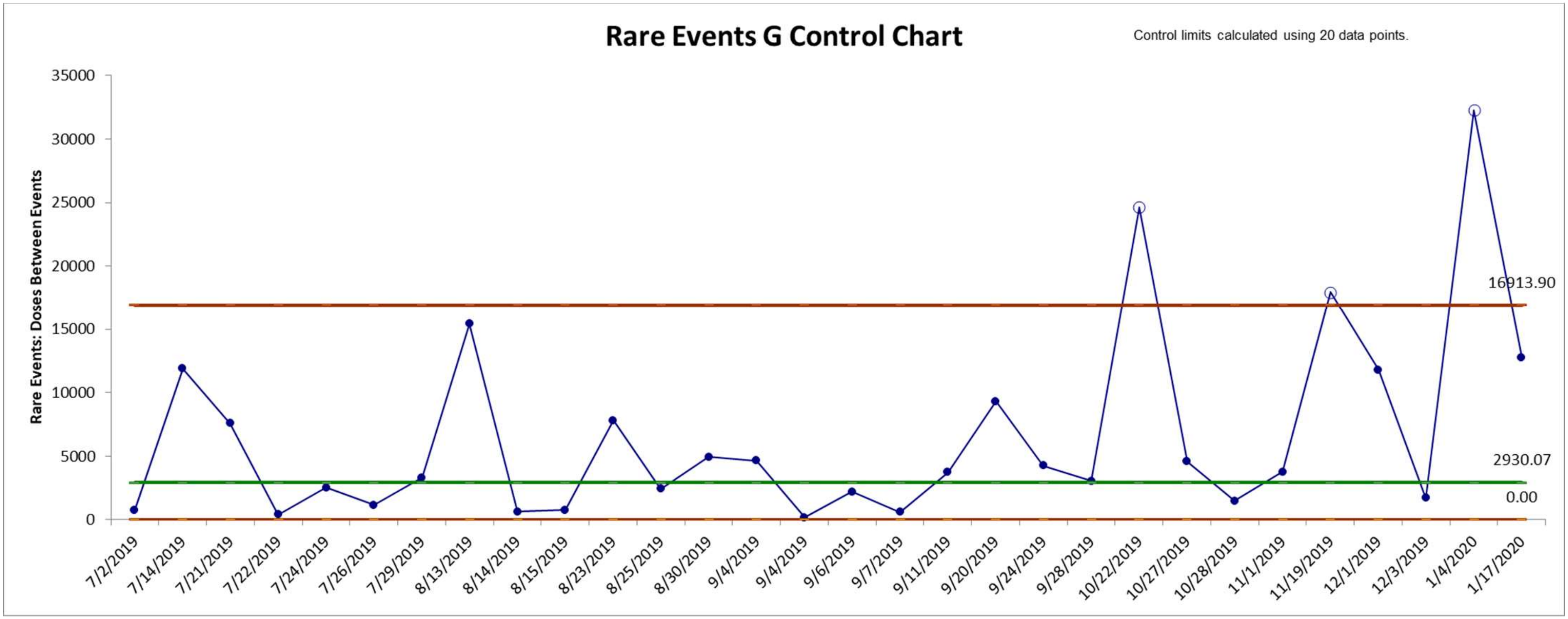
$$UCL = \bar{g} + 3 * \sqrt{\bar{g}(1 + \bar{g})}$$

$$LCL = 0$$

Example: Rare Events G Chart Template

SigmaXL > Control Charts > Control Chart Templates > Rare Events > Rare Events G

Data: Doses Dispensed Between Adverse Drug Events.xlsx



Rare Events Prob G Chart Template

- The Rare Events Prob (Probability-Based) G chart (or Geometric chart) is an alternative to a standard attribute chart when the adverse event of interest is rare and discrete opportunities between events are counted (e.g., number of units or days between). The use of probability-based control limits is recommended in order to properly control the Type I (false alarm) error rate.
- The calculation of control limits is based on the geometric distribution (Bennyan). Event probability and alpha are used to compute the non-symmetrical limits. An "out-of-control" signal above the UCL is desirable, indicating a significant increase in units/opportunities or days between adverse rare events.

g = number of opportunities or units between incidents

$\bar{g}$ = average of g's

The event probability is user specified or estimated as:

$$\hat{p} = \left(\frac{1}{\bar{g}+1}\right)\left(\frac{N-1}{N}\right)$$

This is the minimum variances unbiased estimator.

Rare Events Prob G Chart Template

The probability-based control limits are calculated as:

$$UCL = \frac{\ln(\alpha_{\mathrm{UCL}})}{\ln(1-p)} - 1$$

$$CL = \frac{\ln(0.5)}{\ln(1-p)} - 1$$

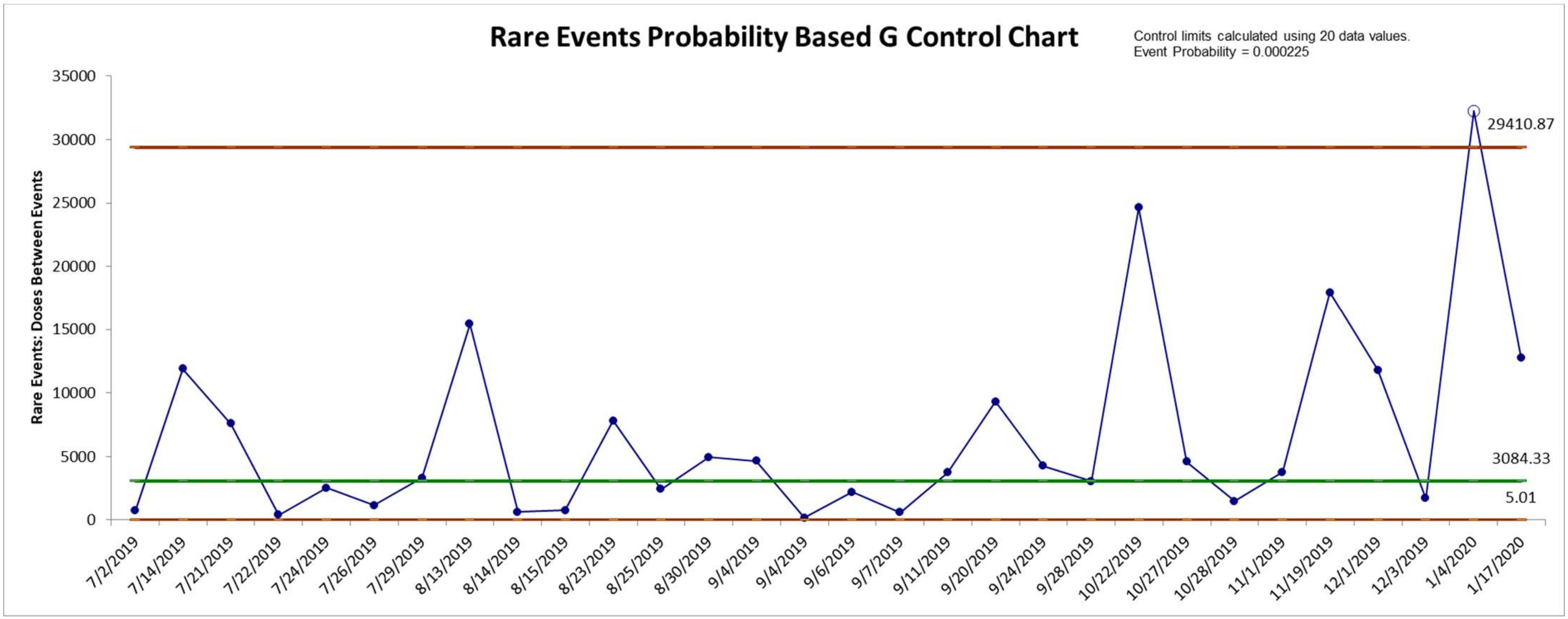
$$LCL = \max\left(0, \frac{\ln(1-\alpha_{\mathrm{UCL}})}{\ln(1-p)} - 1\right)$$

where α_{UCL} is user specified.

Example: Rare Events Prob G Chart Template

SigmaXL > Control Charts > Control Chart Templates > Rare Events > Rare Events Prob G

Data: Doses Dispensed Between Adverse Drug Events.xlsx



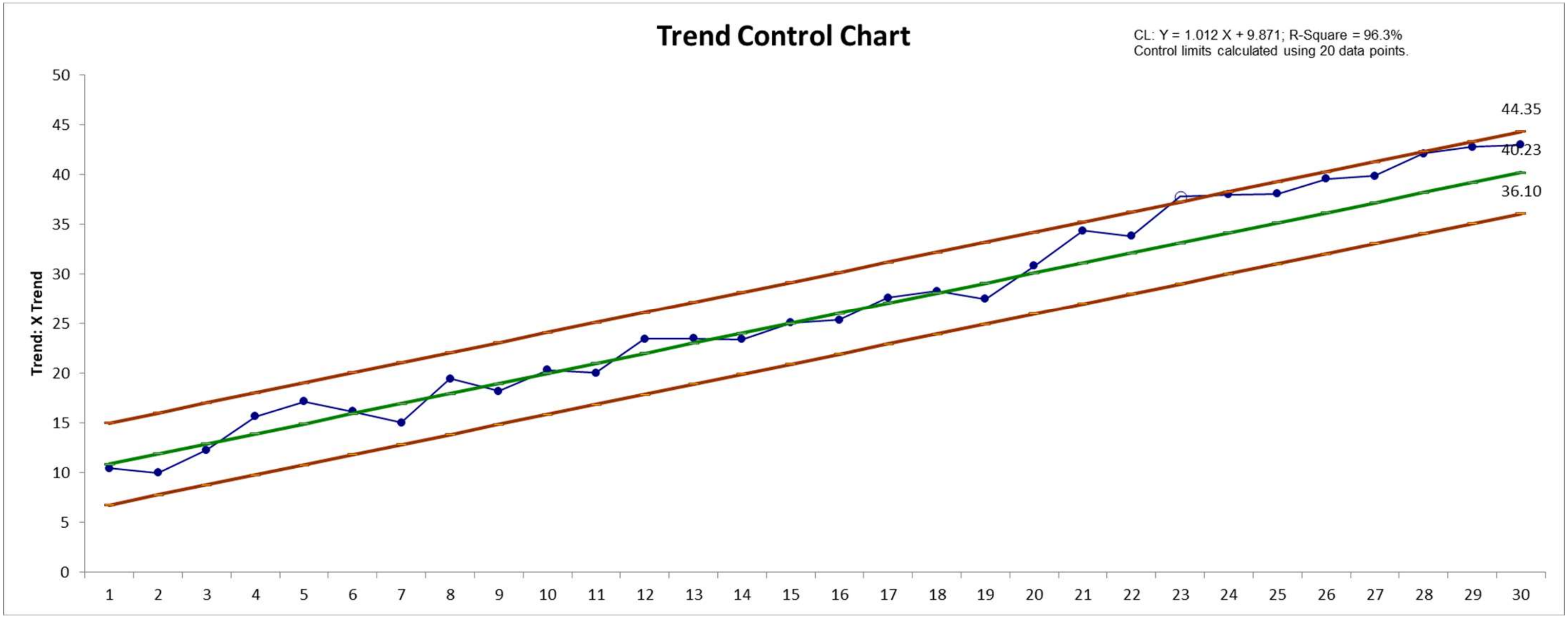
Trend/Toolwear Chart Template

- The Trend Control Chart template should be used with continuous data. The data must be in chronological time-sequence order and have a consistent positive or negative linear trend which is inherent to the process. This is also known as a Toolwear Control Chart.
- Center line (CL) is the linear regression equation. Stdev is estimated using MR-bar/1.128 of regression residuals.
- Note that the regression model estimation error is not included in the calculation of the control limits. R-Square should be at least 50%, preferably greater than 80%.
- Trend control limit formulas are given in Provost and Murray. Alternatively, a regular Individuals Chart may be constructed on the regression model residuals.

Example: Trend/Toolwear Chart Template

SigmaXL > Control Charts > Control Chart Templates > Trend

Data: Trend Chart Example.xlsx



Exponentially Weighted Moving Average (EWMA) Chart Template

- The EWMA control chart uses weighted averages, where the weights decrease exponentially as observations come from further in the past with the smallest weights associated with the oldest observations.

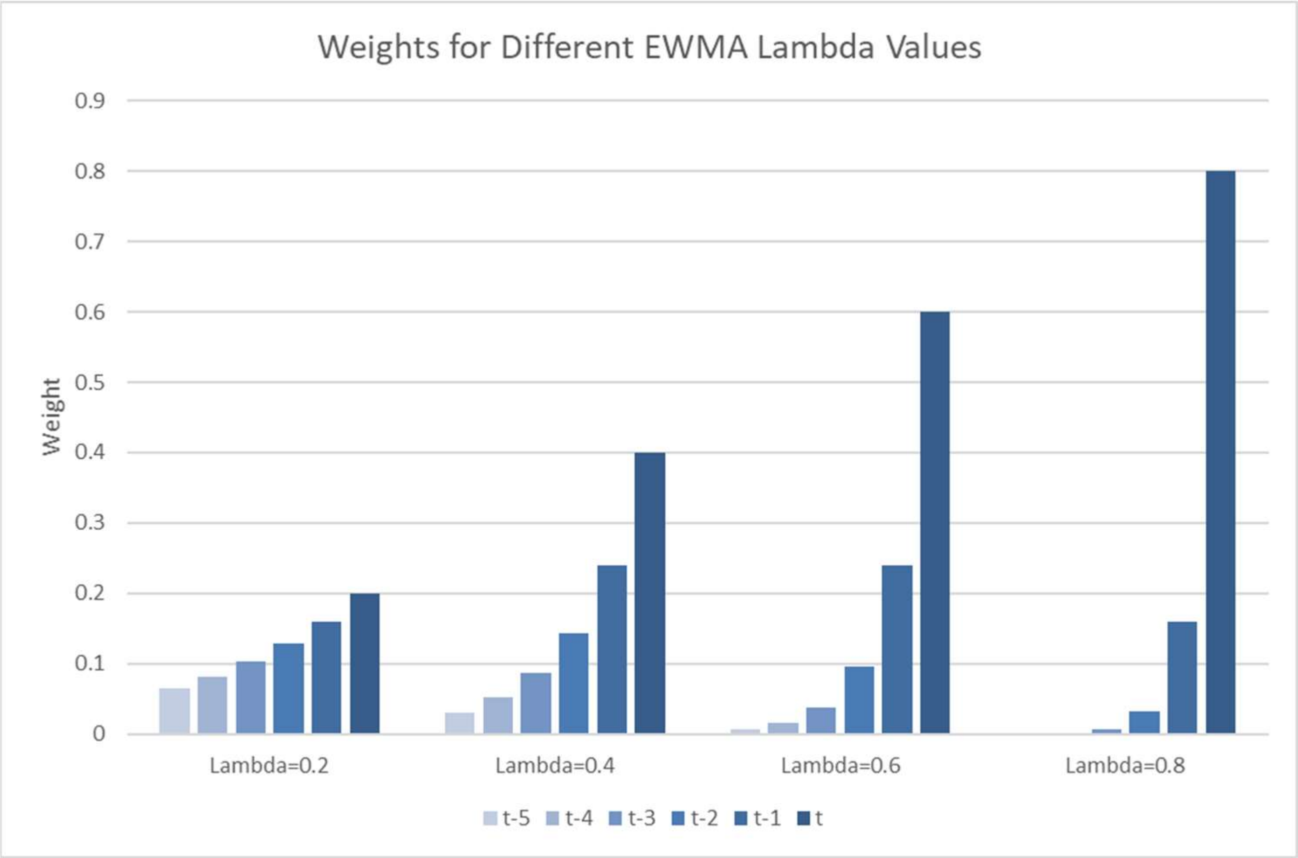
- The formula used for the EWMA statistic is given as:

$$\mathrm{EWMA}_t = \lambda X_t + (1 - \lambda)\mathrm{EWMA}_{t-1}$$

with the starting value EWMA_0 estimated as the data mean and the smoothing parameter λ specified according to desired average run length characteristics (see Average Run Length (ARL) Calculators).

Exponentially Weighted Moving Average (EWMA) Chart Template

Weights for different λ values are shown graphically:



Exponentially Weighted Moving Average (EWMA) Chart Template

The control limits for an EWMA chart are calculated as:

$$\text{UCL} = \mu_0 + K\sigma\sqrt{\frac{\lambda}{(2-\lambda)}[1-(1-\lambda)^{2t}]}$$

$$\text{CL} = \mu_0$$

$$\text{LCL} = \mu_0 - K\sigma\sqrt{\frac{\lambda}{(2-\lambda)}[1-(1-\lambda)^{2t}]}$$

where μ_0 is specified as the historical mean or estimated as:

$$\hat{\mu}_0 = \bar{x}$$

and σ is specified as the historical standard deviation or estimated as:

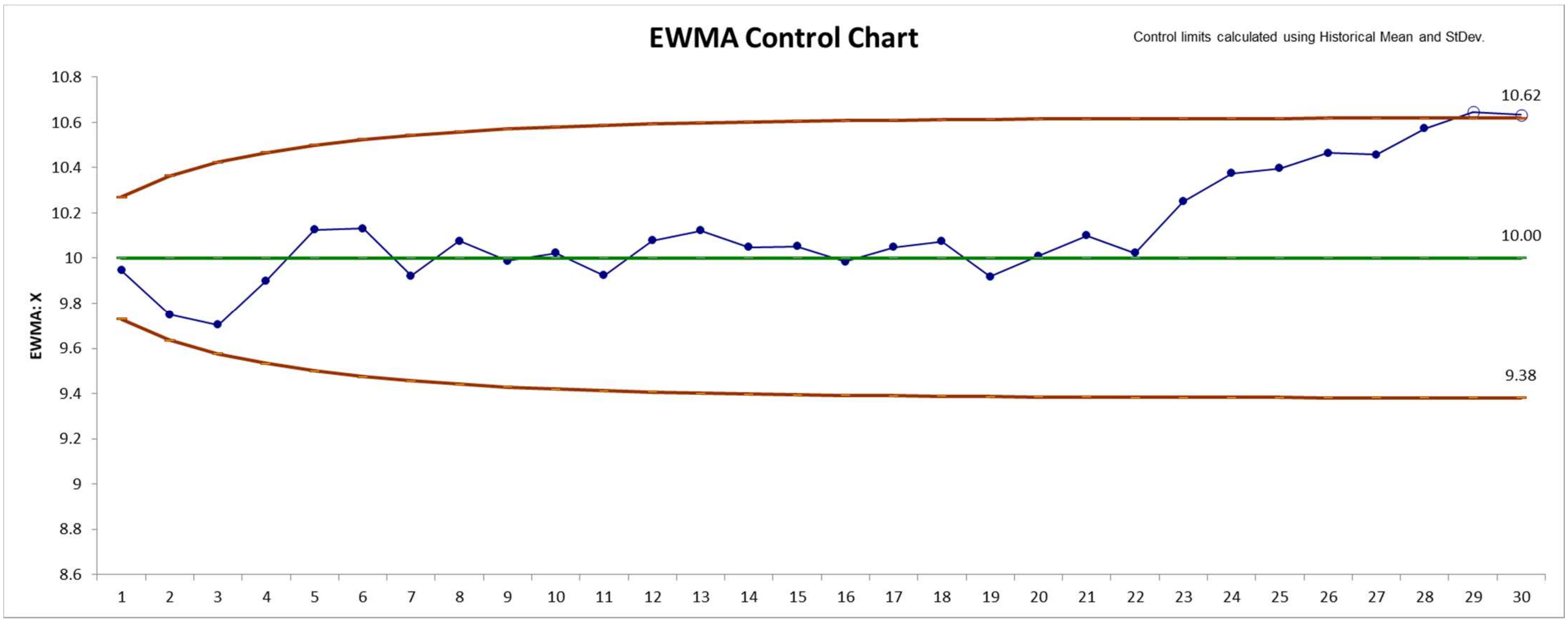
$$\hat{\sigma} = \overline{MR}/d_2\,.$$

Note, the K multiplier is called L in Montgomery (2013).

Example: Exponentially Weighted Moving Average (EWMA) Chart Template

SigmaXL > Control Charts > Control Chart Templates > Time Weighted > EWMA

Data: Montgomery Table 9.1.xlsx



Tabular Cumulative Sum (CUSUM) Chart Template

- The CUSUM chart plots the cumulative sums of deviations of sample values from a target value. Because they combine information from several samples, cumulative sum charts are more effective than Shewhart charts for detecting small process shifts.
- There are two ways to represent CUSUMs: the tabular (or algorithmic) CUSUM, and the V-mask form of the CUSUM (see Montgomery, 2013). SigmaXL utilizes the Tabular CUSUM.

Tabular Cumulative Sum (CUSUM) Chart Template

The formulas used for the Tabular CUSUM statistics are given as:

$$C_1^+ = \max[0, x_t - (\mu_0 + k\sigma) + FIR\sigma]$$
$$C_1^- = \min[0, x_t - (\mu_0 - k\sigma) + FIR\sigma]$$

$$C_t^+ = \max[0, x_t - (\mu_0 + k\sigma) + C_{t-1}^+]$$
$$C_t^- = \min[0, x_t - (\mu_0 - k\sigma) + C_{t-1}^-]$$

where μ_0 is specified as the Target or estimated as:

$$\hat{\mu}_0 = \bar{x}$$

and σ is specified as the historical standard deviation or estimated as:

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

and k and FIR are specified.

Tabular Cumulative Sum (CUSUM) Chart Template

k is the reference (or slack) value, typically set to 0.5. It sets the size of mean shift ($2k\sigma$) that you would like to detect quickly, so 0.5 denotes rapid detection of a shift in mean = 1 σ.

FIR is the optional fast initial response (or headstart) value. This sets the initial CUSUM statistic so that it improves the sensitivity to a mean shift at startup.

The Tabular CUSUM control limits are calculated as:

$$\text{UCL} = h\sigma$$

$$\text{CL} = 0$$

$$\text{LCL} = -h\sigma$$

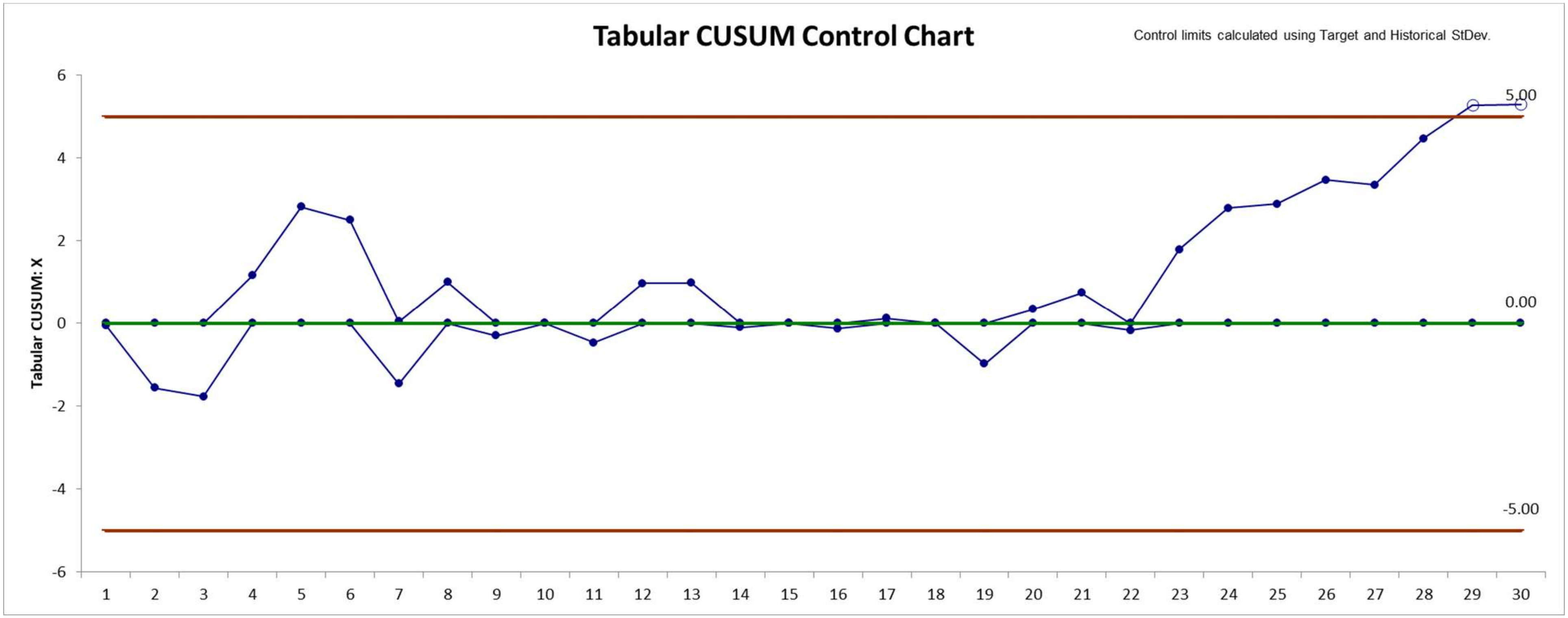
where h is the decision interval, typically set to 4 or 5.

If FIR is used, it is typically set to h /2.

Example: Exponentially Weighted Moving Average (EWMA) Chart Template

SigmaXL > Control Charts > Control Chart Templates > Time Weighted > CUSUM

Data: Montgomery Table 9.1.xlsx



Average Run Length (ARL) Calculators

- Average Run Length (ARL) characteristics are very useful to compare the performance of control charts and determine optimal parameter settings for EWMA and CUSUM time weighted control charts.
- The ARL value for a shift in mean = 0 sigma is the “in-control” average run length and is denoted as ARL0. ARL0 is $1/\alpha$, where α is the type I false alarm probability, so this should be as large as possible minimizing the likelihood that an out-of-control signal is a false alarm. In a Shewhart Individuals Control Chart, ARL0 = $1/\alpha = 1/(0.00135*2) = 370.4$. On average, we will see a false alarm once every 370 observations. Note that this is a mean of a geometric distribution, so in practice the actual ARL0 will vary widely with the standard deviation approximately equal to the mean value.
- When we have a sustained shift in mean > 0, the ARL value is the “out-of-control” run length and is denoted as ARL1. ARL1 = $1/(1-\beta)$, where β is the type II miss probability and $(1-\beta)$ is the power to detect. This should be as small as possible so that a shift in process mean is quickly detected.

Average Run Length (ARL) Calculators

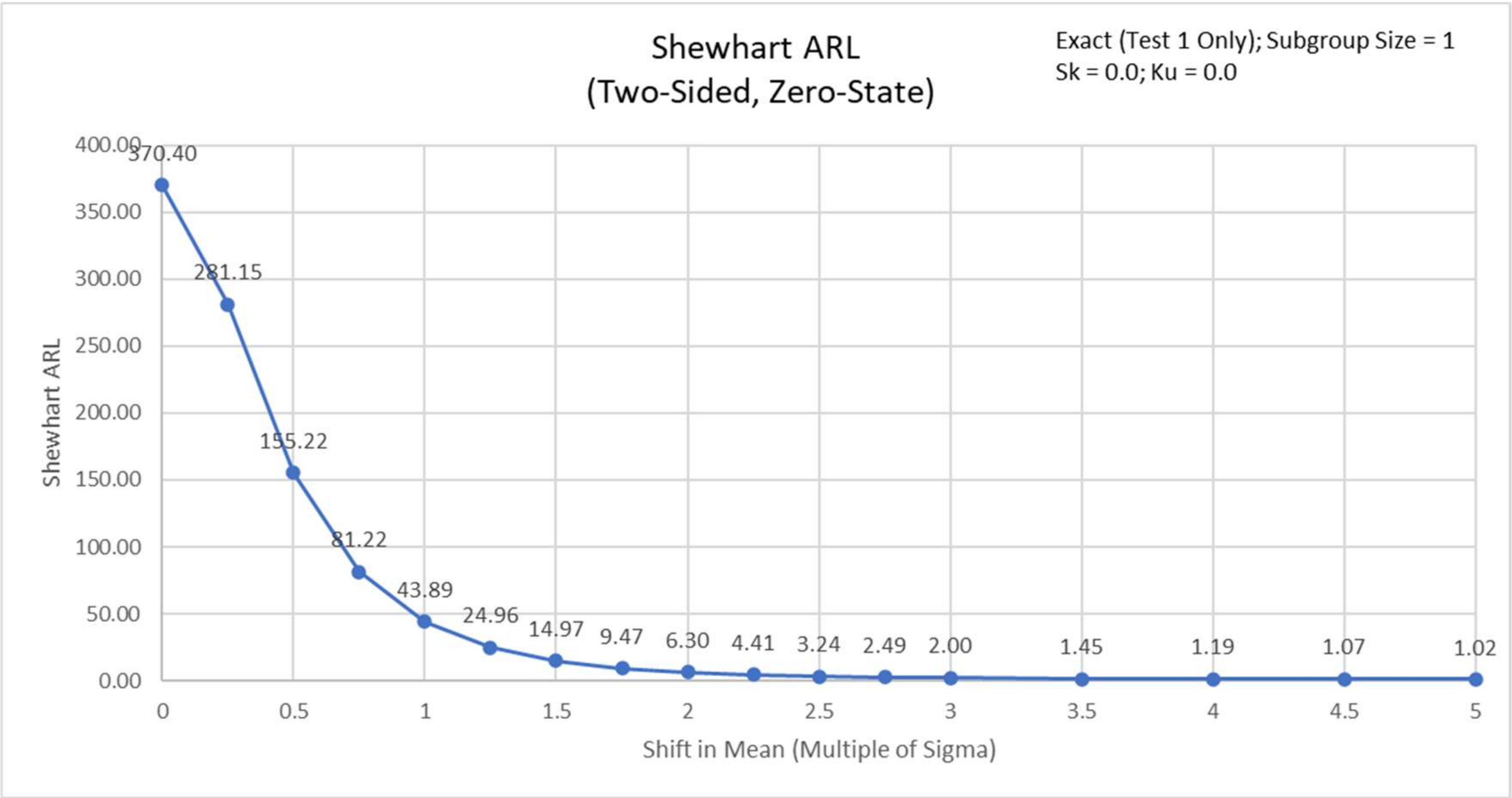
- The calculations for ARL are quite complex, involving either Exact, Markov Chain approximation or Monte Carlo simulation to solve. The SigmaXL ARL Templates take care of these calculations and are easy to use.
- If Monte Carlo simulation is used, additional Run Length standard deviation and percentile statistics are reported. Monte Carlo results will take some time and vary slightly every time they are run. With 1000 (1e3) replications it will be fast, approx. 10 seconds, but will have an ARL0 error of approximately +/- 10%; 10,000 (1e4) replications will take about a minute, with an ARL0 error of +/- 3.2%; 100,000 (1e5) replications will take about ten minutes, with an ARL0 error = +/- 1%. More than 1e5 replications will be very time consuming and may run into memory limitations, so is not recommended.
- When using EWMA or CUSUM charts, we typically set the parameters to minimize the ARL1 to give rapid detection for a small shift in mean of 1 sigma. Shewhart charts are typically used when trying to rapidly detect a large shift in mean of >= 3 sigma. Tests for special causes may be used with Shewhart to improve the small shift performance, but they give poor ARL0 performance resulting in frequent false alarms.

Average Run Length (ARL) Calculators

- Subgroups improve the small shift performance of a Shewhart chart without impacting the ARL0 rate and, if possible, should be used. The ARL for subgroup averages is adjusted by using the sigma of averages, sigma/$\sqrt{n}$. For example, with a subgroup size of 4, the ARL1 values at shift in mean = 1 will match the ARL performance of an Individuals chart with shift in mean = 2 sigma.
- Note that subgroups for CUSUM and EWMA are not available in SigmaXL.
- The problem of robustness to nonnormality can also be considered by using the Pearson family to specify any value of skewness and kurtosis and estimate the ARLs.
- The average run length calculators are for two-sided charts with zero-state, i.e., the shift is assumed to occur at the start. The parameters (mean, standard deviation and proportion) are also assumed to be known. This will not likely be the case in use, but is still useful for determining parameter settings and comparison of ARL across chart types.
- Due to the complexity of calculations, SigmaXL must be loaded and appear on the menu in order for the ARL templates to function.

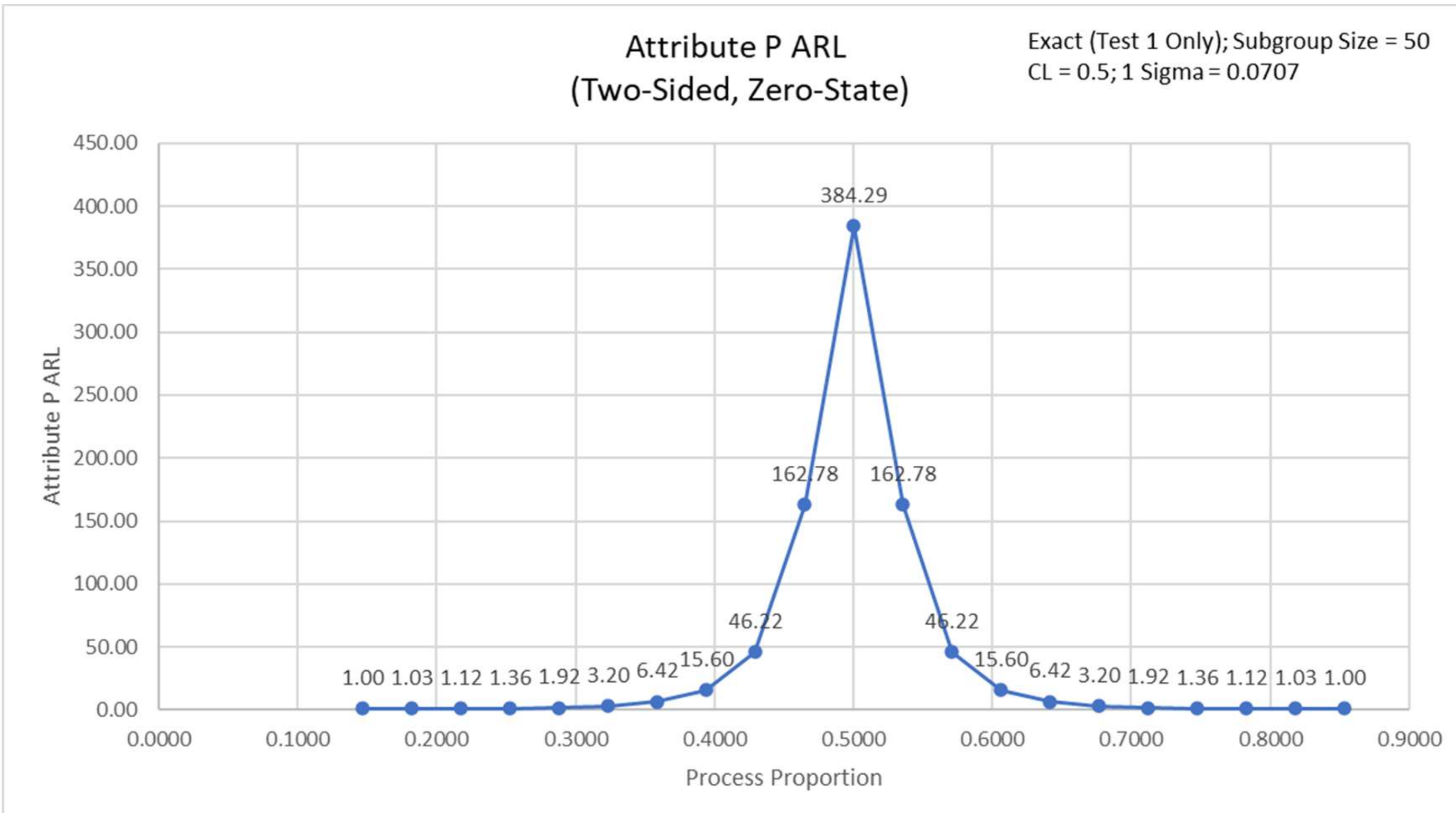
Example: Shewhart ARL Calculator

SigmaXL > Control Charts > Control Chart Templates > Average Run Length (ARL) Calculators > Shewhart ARL



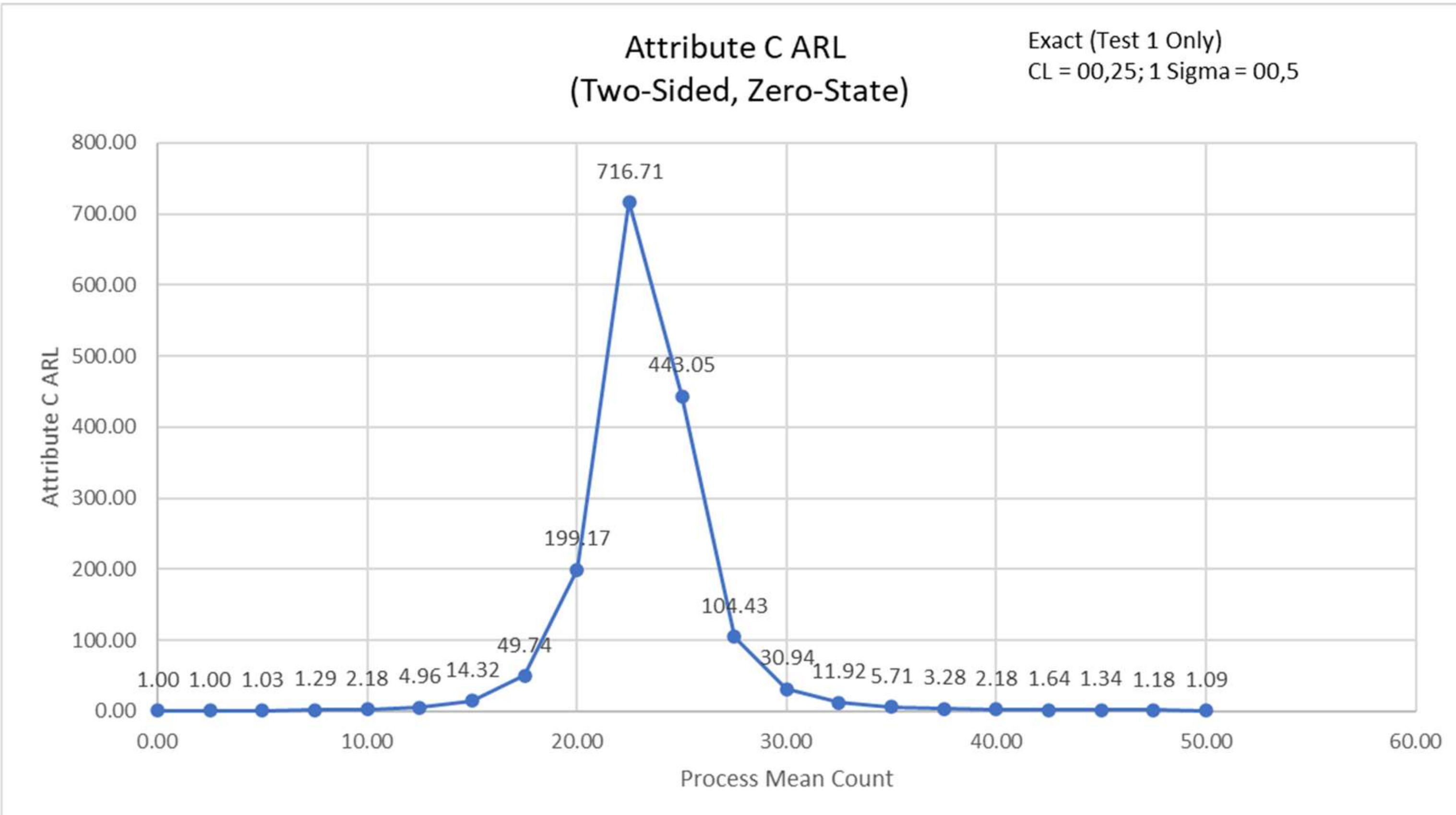
Example: Attribute P ARL Calculator

SigmaXL > Control Charts > Control Chart Templates > Average Run Length (ARL) Calculators > Attribute P ARL



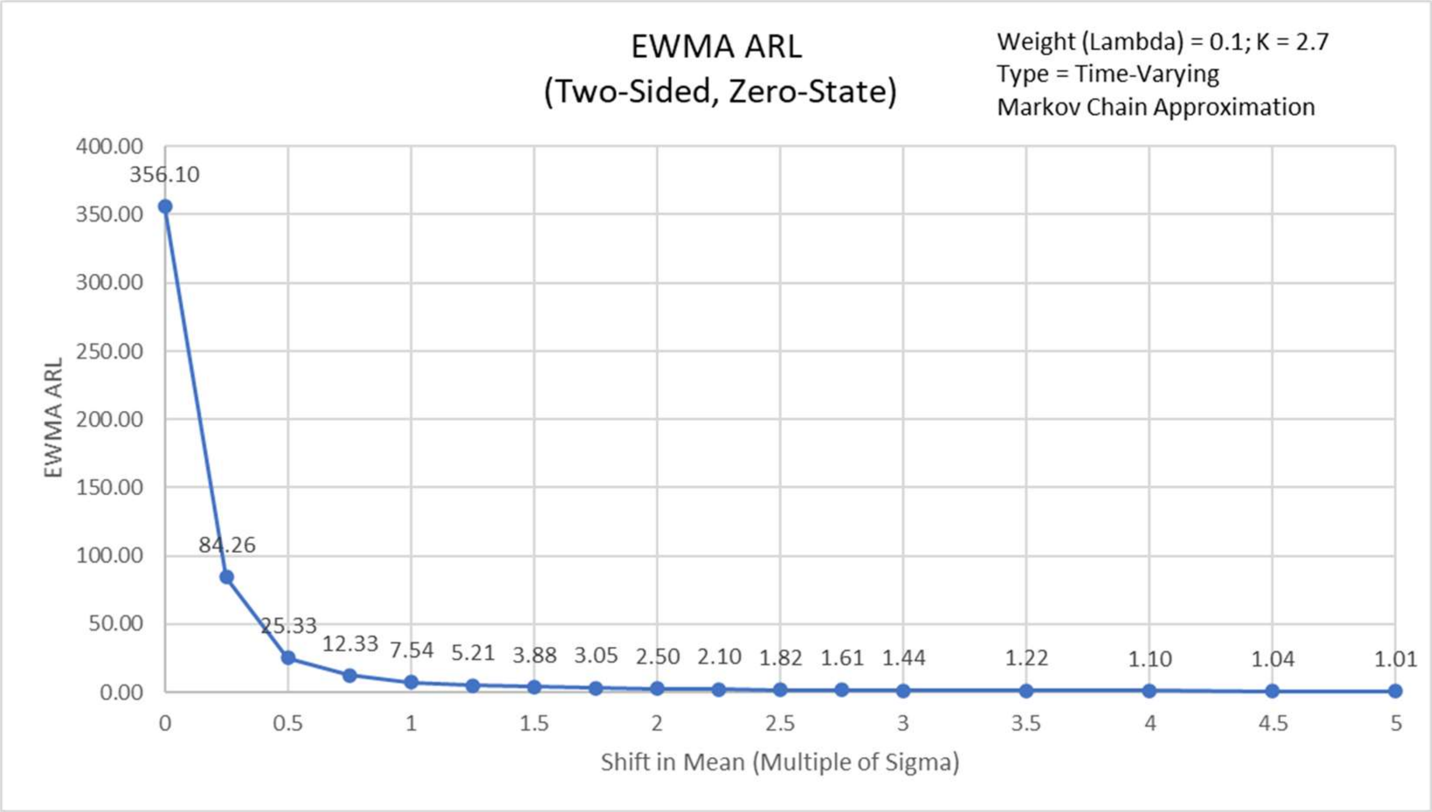
Example: Attribute C ARL Calculator

SigmaXL > Control Charts > Control Chart Templates > Average Run Length (ARL) Calculators > Attribute C ARL



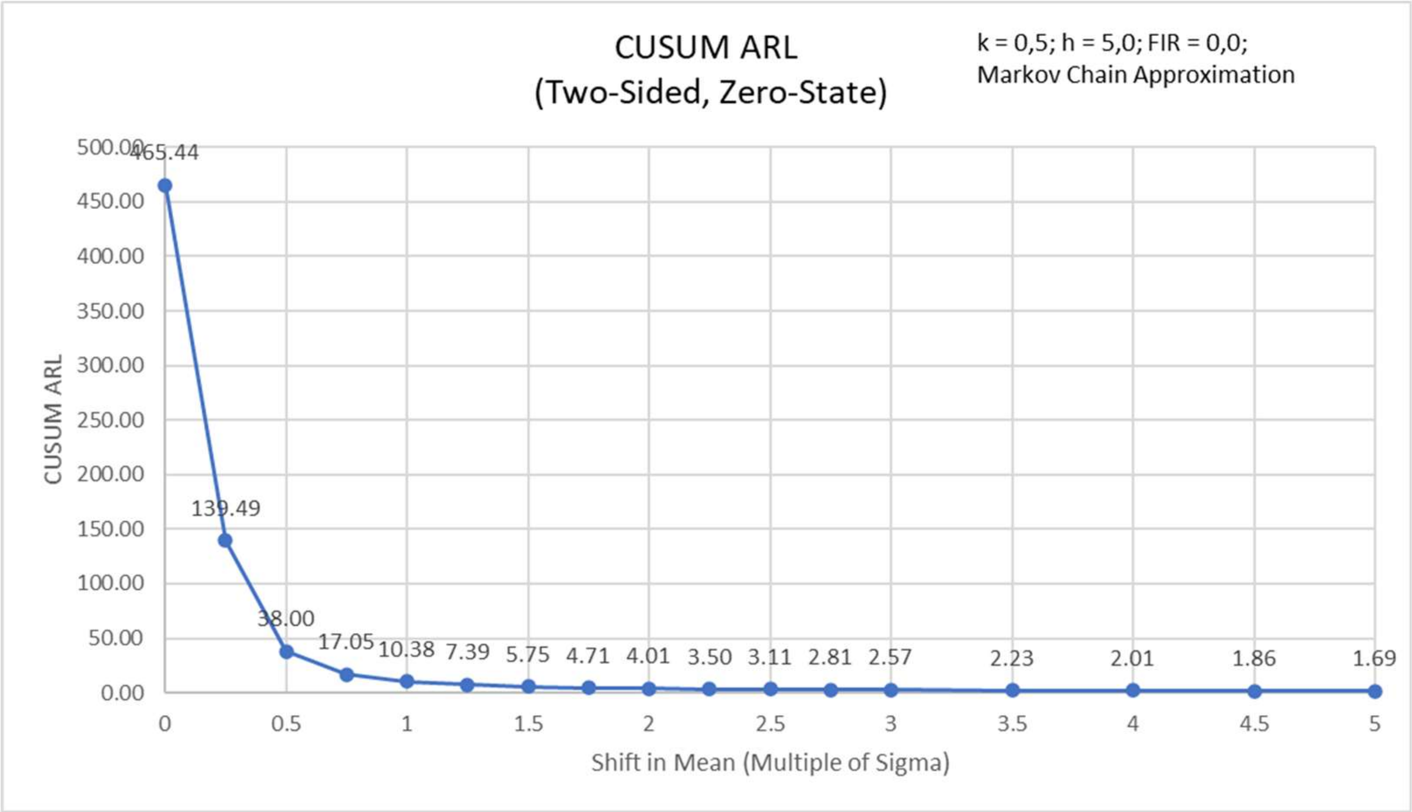
Example: EWMA ARL Calculator

SigmaXL > Control Charts > Control Chart Templates > Average Run Length (ARL) Calculators > EWMA ARL



Example: CUSUM ARL Calculator

SigmaXL > Control Charts > Control Chart Templates > Average Run Length (ARL) Calculators > CUSUM ARL



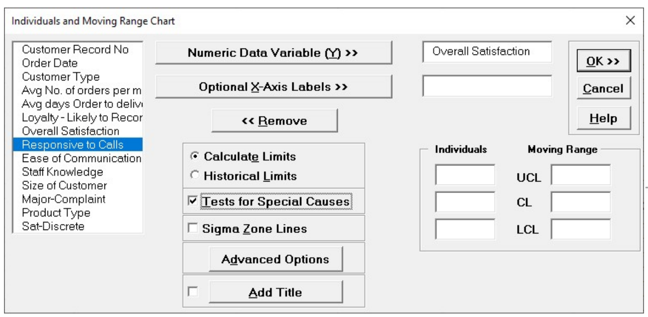
Tests for Special Causes

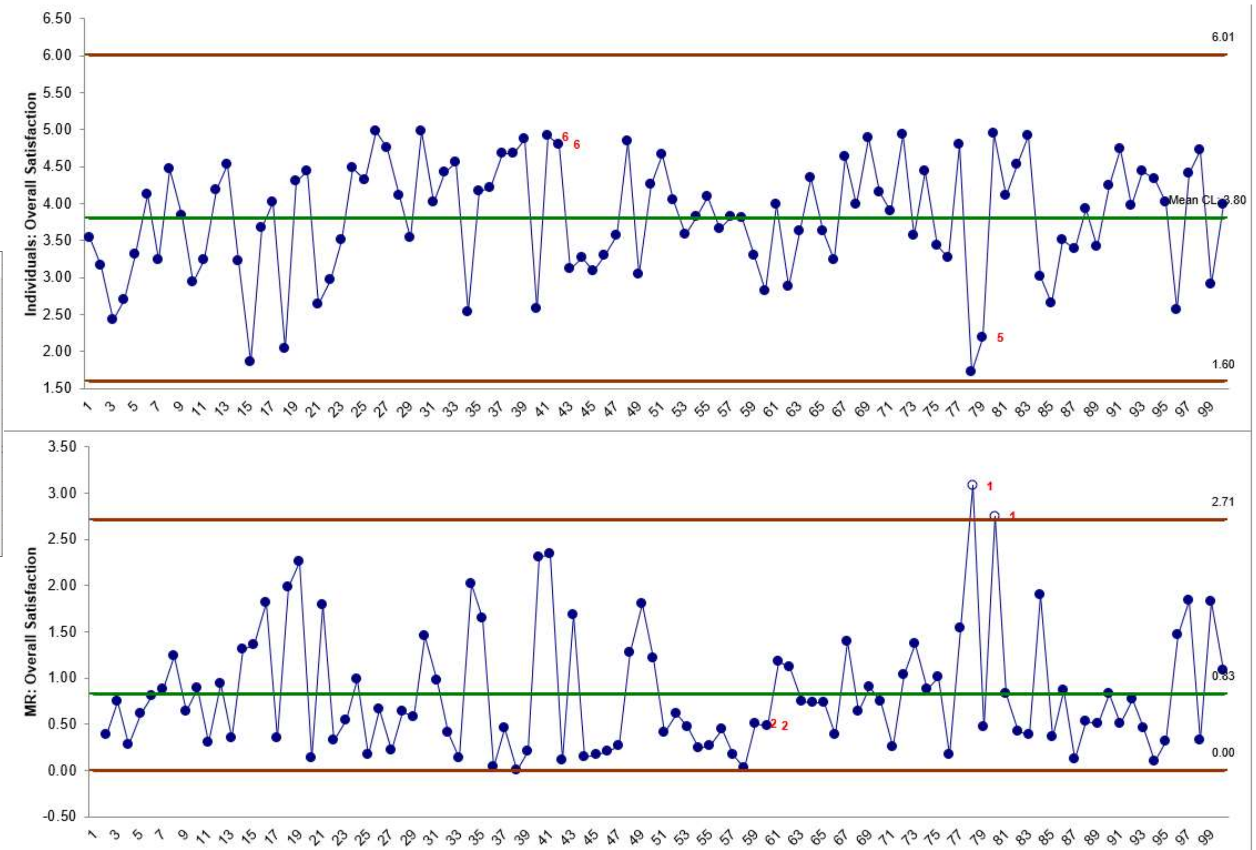
- Tests for Special Causes now supported for menu-based control charts:
  - MR/Range/StDev Charts (Tests 1-4)
  - Varying Subgroup Sizes (Moving Limits)
  - Historical Groups
- Tests for Special Causes Defaults dialog update
- I-MR-R/S menu update

Example: Individuals & Moving Range with Tests for Special Causes

SigmaXL > Control Charts > Individuals & Moving Range

Data: Overall Satisfaction in Customer Data.xlsx



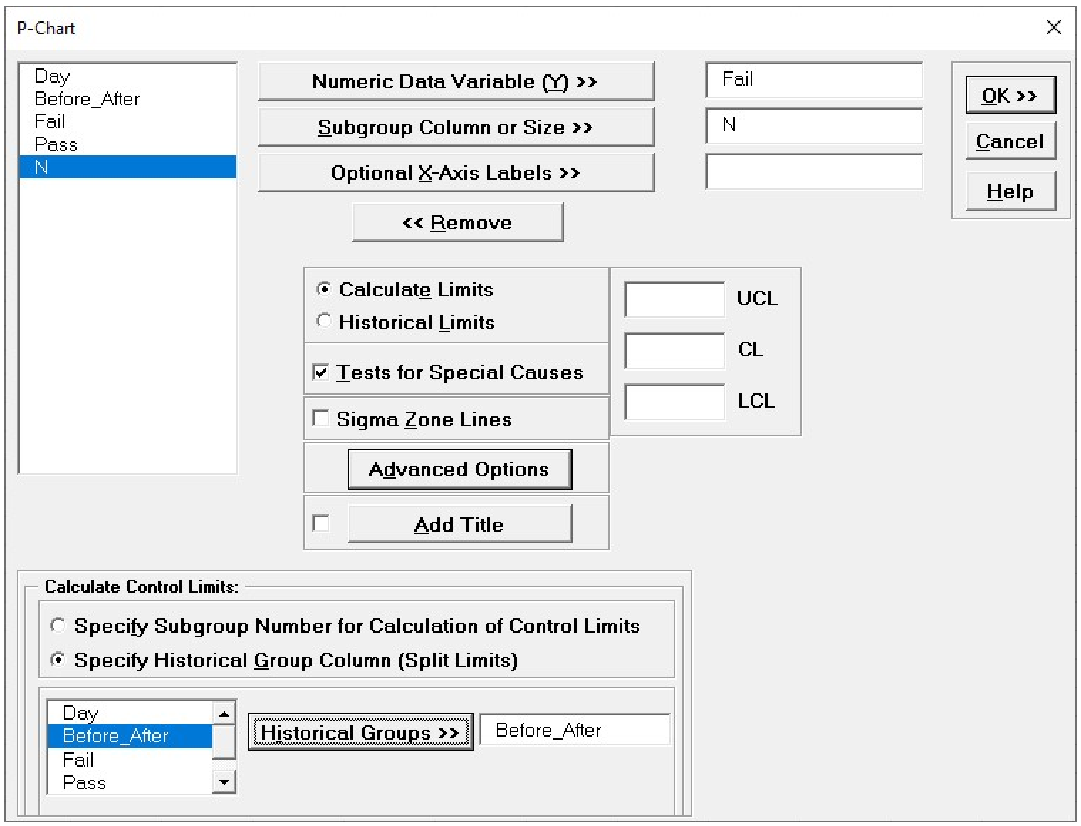


Example: Attribute P Chart with Tests for Special Causes, Varying Subgroup Size and Historical Groups

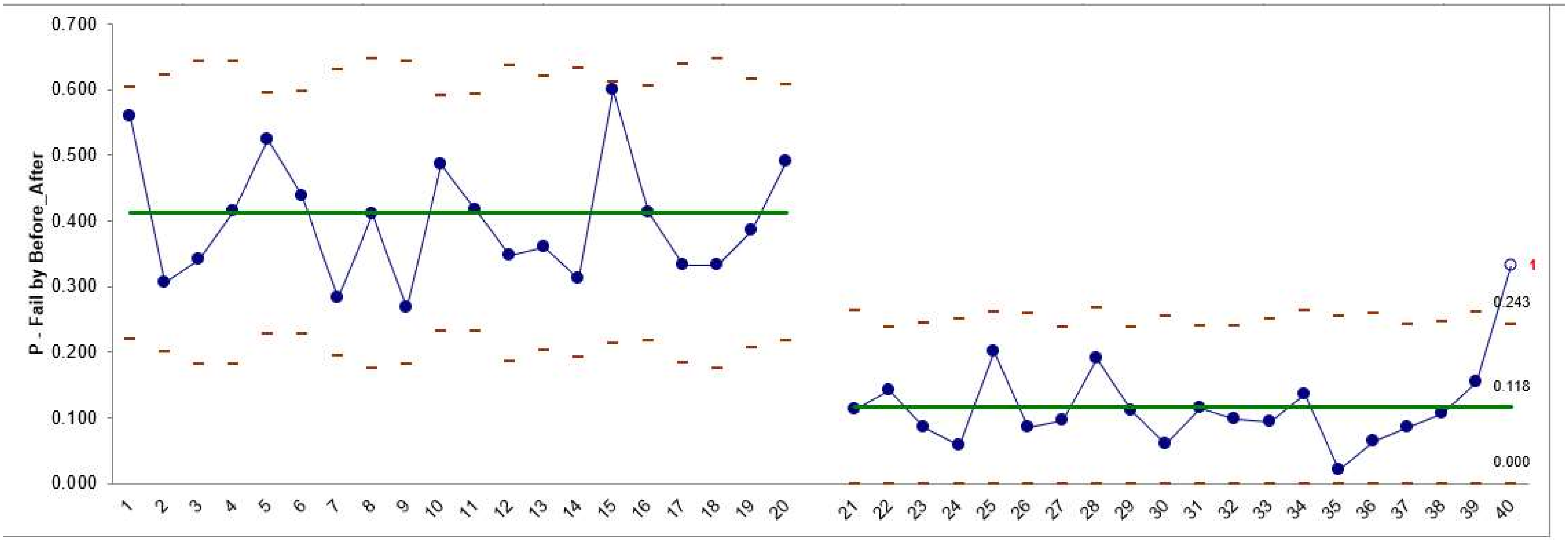
SigmaXL > Control Charts > Attribute Charts > P

Data: New York Daily Cycle Time Discrete - Before After Improvement.xlsx with Day 40 modified: Fail = 20, Pass = 40, N = 60:

40	After	20	40	60



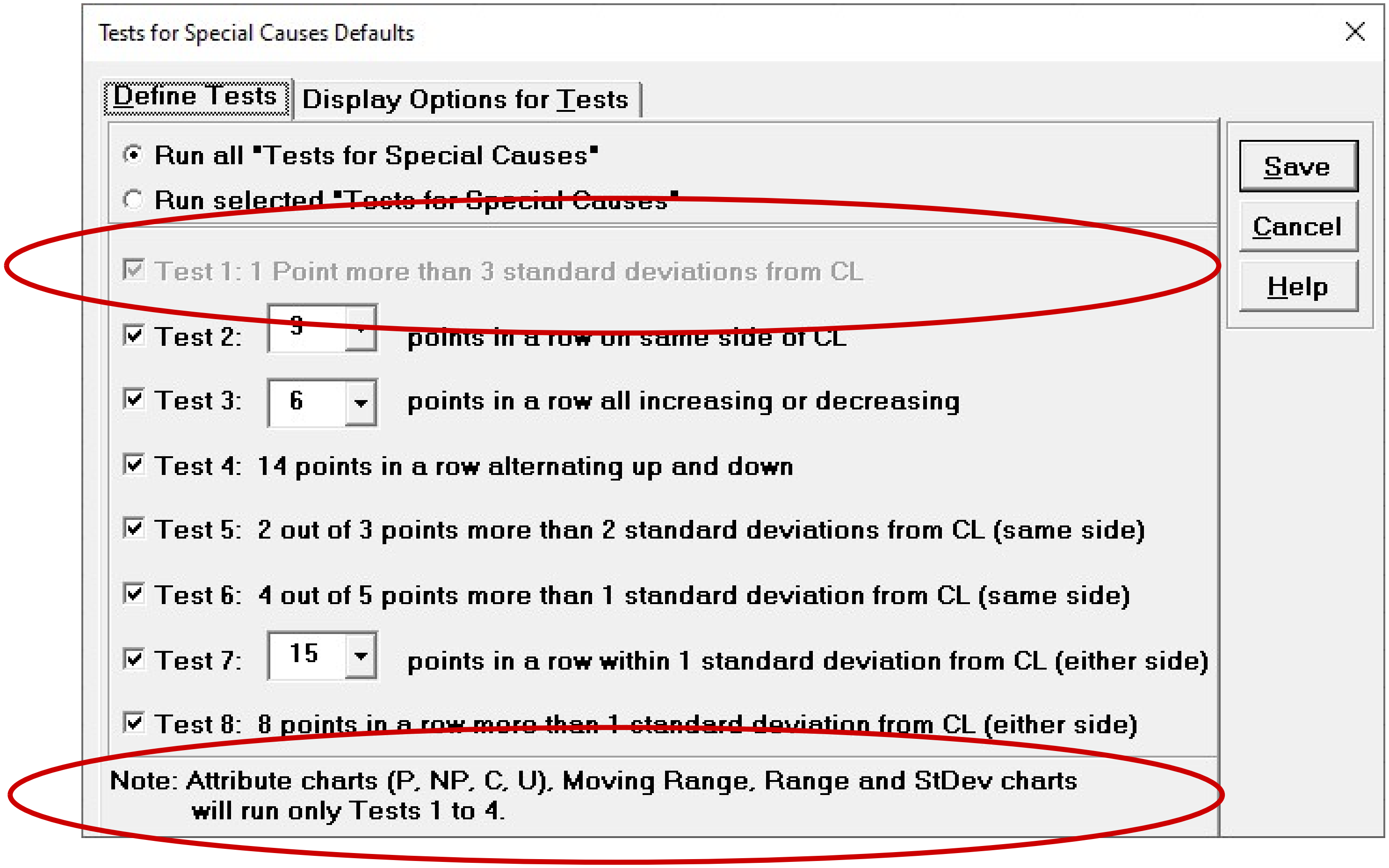
Example: Attribute P Chart with Tests for Special Causes, Varying Subgroup Size and Historical Groups



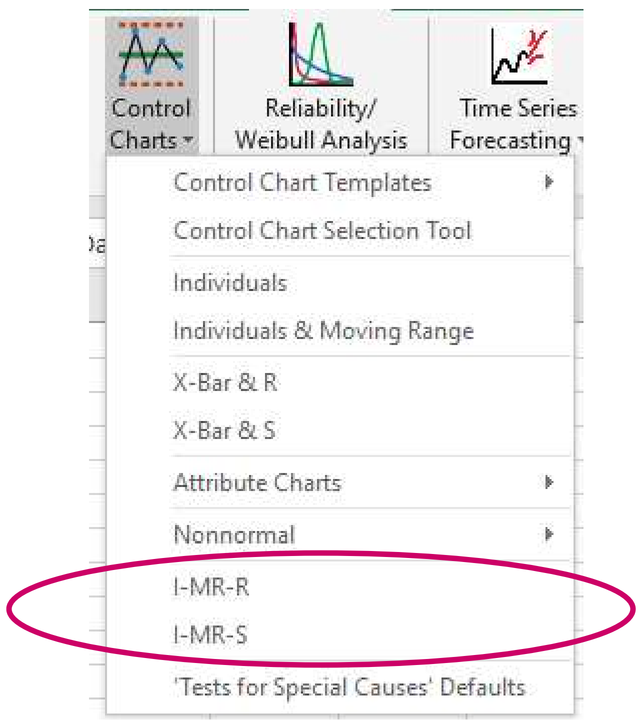
Tests for Special Causes - P - Fail
Number of Data Points Failing Tests = 1

Observation No.	Test 1: 1 point more than 3 StDev from CL	Test 2: 9 points in a row on same side of CL	Test 3: 6 points in a row all increasing or all decreasing	Test 4: 14 points in a row alternating up and down
40	x			

Tests for Specials Causes Defaults Dialog Update



I-MR-R/S Menu Update



What's New in SigmaXL® Version 9

Part 1 of 3: New and Improved Control Charts

Questions?

References for Control Chart Templates

1. Benneyan, J.C. (2001), "Performance of Number-Between g-Type Statistical Control Charts for Monitoring Adverse Events", *Health Care Management Science*, 4, pp. 319–336.
2. Montgomery, D.C. (2013), *Introduction to Statistical Quality Control*, Seventh Ed., Wiley.
3. Nelson, L.S. (1994), "A Control Chart for Parts-Per-Million Nonconforming Items", *Journal of Quality Technology*, 26:3, pp. 239-240.
4. Provost L, Murray S. (2011), *The Health Care Data Guide: Learning from Data for Improvement*. San Francisco: Jossey-Bass, pp. 230-231.

References for ARL Calculators

1. Borror, C. M., Montgomery, D.C. and Runger G. C. (1999). “Robustness of the EWMA Control Chart to Nonnormality,” *Journal of Quality Technology*, 31(3), 309–316.

2. Chakraborti, S. (2007), “Run Length Distribution and Percentiles: The Shewhart Chart with Unknown Parameters”, *Quality Engineering* 19, 119–127.

3. Champ, C.W. and Woodall, W.H. (1987), "Exact results for Shewhart control charts with supplementary runs rules", *Technometrics* 29, 393-399.

4. Hawkins, D. M. and Olwell, D. H. (1998), *Cumulative Sum Charts and Charting for Quality Improvement (Information Science and Statistics)*, Springer, New York.

5. Lucas, J.M. and Crosier R.B. (1982), “Fast Initial Response for CUSUM Quality-Control Schemes: Give Your CUSUM A Headstart”, *Technometrics* 24, 199-205.

6. Lucas J.M. and Saccucci M.S. (1990), “Exponentially weighted moving average control schemes: Properties and enhancements”, *Technometrics* 32, 1-12.

References for ARL Calculators

7. Montgomery, D.C. (2013), *Introduction to Statistical Quality Control*, Seventh Ed., Wiley.

8. Schilling, E. G., and P. R. Nelson (1976), “The Effect of Nonnormality on the Control Limits of X-bar Charts,” *Journal of Quality Technology*, Vol. 8(4), pp. 183–188.

9. Steiner, S. H. (1999), "EWMA control charts with time-varying control limits and fast initial response", *Journal of Quality Technology* 31(1), 75-86.

10. Stoumbos, Z. G. and Reynolds, M.R. Jr. (2004), “The Robustness and Performance of CUSUM Control Charts Based on the Double-Exponential and Normal Distributions”, In: Lenz, H. J., Wilrich, P. T. (eds) *Frontiers in Statistical Quality Control* 7, Physica, Heidelberg, 79-100.

11. Woodall, W. H. and Faltin, F.W. (2019), "Rethinking control chart design and evaluation", *Quality Engineering* 31, 596-605.

Sigma XL
Simply Smart Software

Lean Six Sigma Statistical Tools, Templates & Monte Carlo Simulation in Excel

What's New in SigmaXL® Version 9

Upcoming Webinars:

Part 2 of 3: Time Series Forecasting
Thursday, November 12, 2020 at 3 pm ET.

Part 3 of 3: Control Charts for Autocorrelated Data
Thursday, December 10, 2020 at 3 pm ET.

Or visit www.SigmaXL.com for recordings of webinars (typically available 2 weeks after scheduled webinar).