Use of Control Chart Interpretation Rules

Description and Strategy

By Polona K. Carson and John H. Carson Jr.

Q What are control chart rules, and which rules should be applied for control chart interpretation?

A Control charts used in testing are statistical tools to detect testing process upsets in terms of change in process location/mean value and process variation. The testing process is stable when measurement outputs are predictable given a known value of the measurand, the variable being measured. This implies that the reliability of measurement of unknowns is also characterized. A process demonstrated to be stable via control charting test results of control samples across a range of measurements can be expected to produce measurements of unknown test items with bias and precision comparable to the performance indicated by control samples. This assumes that the control materials are representative for the test items.

Control charts described in the practice for use of control charts in statistical process control (E2587) are based on the assumption that process data are independent and normally distributed. Charts for process location (I-chart or $\overline{X}$-chart) and for process variation (mR-chart, R-chart, or S-chart) are needed for the sound characterization of a testing process. Both types of charts are defined by similar parameters:

— Center line, averaged value of a statistic estimating mean or variation, and
— Lower and upper limits around the center line.

Western Electric (WE) rules described in E2587 are designed to differentiate between natural process variation and special cause variation. The most well-known are the first four WE rules:

1. The Shewhart rule—one value falls outside either three-sigma control limit,
2. Two out of three consecutive values fall outside the two-sigma warning limits on the same side,
3. Four out of five consecutive values fall outside the one-sigma limits on the same side, and
4. Eight consecutive values either fall above or fall below the center line.

A control chart signals when one of its evaluation rules is violated. Rule 1 is fairly sensitive to a large shift in the process (>2.5 sigma) but has very little power to detect shifts of < 1.5 sigma within a reasonable time. Rules 2-4 make the control chart more sensitive to smaller shifts.

All rules (decision procedures) must be evaluated in terms of how they balance risk between Type I and Type II error. A Type I error occurs when a chart falsely signals that process is out of control. A Type II error occurs when a chart fails to signal when the testing process is in fact out of control. The Type II error rate depends on the properties of the special cause, the process departure from stable state.

Assuming that the process is stable, that the mean and standard deviation are perfectly known and that the testing process truly follows a normal distribution, the Type I error rate under Rules 1, 2, 3, and 4 are 0.27, 0.20, 0.13, and 0.78 percent, respectively. The WE rules are often used in combination, which can substantially increase the aggregate Type I error rate.

Another important aspect of chart behavior is the distribution of run lengths (RL). A run is the sequence of points between events when the chart is signaling. In the scenario above, we expect on average one false positive signal per 370 points under Rule 1, an in-control average run length (ARL) of 370.

In practice, process parameters are not known but estimated. The Type I error rate of all the WE rules is substantially increased in this case unless the amount of data used to estimate process parameters is quite large. We examined the false positive rates of control chart signaling for an I/MR chart with various WE rules applied and using different initial sample sizes to estimate chart parameters. To do this, we simulated using different combinations of the first four WE rules in combination with Rule 1 applied to the MR chart (MR1). The simulation applied the rules alone and in various combination to 32,000 sequences of independent normally distributed values, each sequence of length 5,000. The results for in-control

Table 1— In-Control ARL for I/MR Chart

Initial Sample Size	12	15	20	25	30	40	50	60	80	100
Rules Applied										
R1	56	71	93	113	130	157	181	199	228	245
R1.R2	40	49	62	74	84	101	114	125	142	153
R1.MR1	34	40	47	54	58	65	71	75	81	85
R1.R2.MR1	30	34	40	46	49	56	61	64	70	73
R1.R3.MR1	26	30	35	40	43	48	52	55	60	63
R1.R4.MR1	27	31	36	41	44	49	53	55	59	62
R1.R2.R3.MR1	24	28	32	36	39	44	48	50	54	57
R1.R2.R4.MR1	24	27	32	36	39	43	47	49	53	55
R1.R3.R4.MR1	22	25	30	33	36	40	43	45	48	50
R1.R2.R3.R4.MR1	21	24	28	31	33	37	39	41	44	46

Table 2 — In-Control Median RL for I/MR Chart

Initial Sample Size	12	15	20	25	30	40	50	60	80	100
Rules Applied										
R1	12	16	24	31	40	54	69	82	104	120
R1.R2	10	12	17	22	27	36	45	52	65	75
R1.MR1	9	12	16	19	22	27	32	36	42	46
R1.R2.MR1	8	10	13	16	19	23	27	31	36	40
R1.R3.MR1	8	10	13	15	18	22	25	28	32	35
R1.R4.MR1	10	13	16	19	22	26	29	32	36	38
R1.R2.R3.MR1	8	9	12	14	16	20	23	25	29	32
R1.R2.R4.MR1	9	11	14	16	19	22	25	28	31	34
R1.R3.R4.MR1	9	10	13	15	17	21	23	25	28	30
R1.R2.R3.R4.MR1	8	10	12	14	16	19	21	23	26	28

ARL and median RL are shown in Tables 1 and 2. Run lengths would be expected to be above the median RL one half of the time. The "Rules Applied" entries are interpreted as, for example, R1.R2.MR1 means apply WE Rules 1 and 2 with MR1.

The distribution of run lengths is extremely skewed to the right. Therefore, the values of median RL in Table 2 are substantially smaller than the ARL values in Table 1, which are inflated because of some very long run lengths, especially in those cases in which the process standard deviation is significantly overestimated. The median RL is much closer to the intuitive "feel" that a control chart user would have when thinking about run length.

RECOMMENDATION

As demonstrated above, we need to be cautious using interpretation rules, since we want to avoid too many false positive signals on control charts. We recommend the following strategy.

Start by applying WE Rule 1 to the control chart for monitoring mean value and MR1 to interpret the chart for variation. This will quickly signal large perturbations (>2.5 sigma) in the testing process that can be catastrophic for laboratory data quality. Once the control chart shows no evidence of process instability, more WE rules can be added to increase the chart sensitivity to smaller perturbations in the testing process.

However, a better strategy is to use a companion EWMA (exponentially weighted moving average) or CUSUM (cumulative sum) chart as described in the previous Data Points article, "Control Charts for Monitoring Test Method Performance" (*SN*, March/April 2016) to monitor process for medium and small changes (< 1.5 sigma).

Updating the chart parameters as more points are added will be helpful in either case.

In a future article, we will present some results on the power of control charts to detect process shifts and compare WE rules to companion charts in terms of their ARL and median RL.

Polona K. Carson, Ph.D., MBA, P&J Carson Consulting LLC, Ohio, is vice chair of the committee on quality and statistics (E11).

John Carson, Ph.D., of P&J Carson Consulting LLC, Findlay, Ohio, is the Data Points column coordinator. He is chairman of the subcommittee on statistical quality control (E11.30), part of the committee on quality and statistics (E11), and a member of the committees on petroleum products, liquid fuels, and lubricants (D02); air quality (D22); and environmental assessment, risk management, and corrective action (E50).