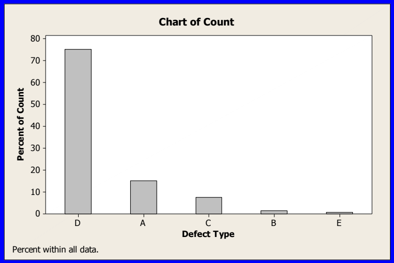
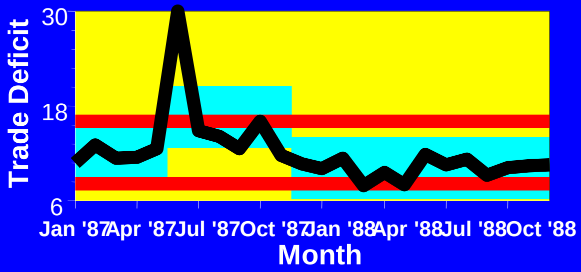
Normalizing the I Control Chart

Dr. Wayne Taylor





Version: September 30, 2017

The Problem

- Average-S, U, P, Laney U’ and Laney P’ control charts all allow the charts to be normalized based on the sample size or number of opportunities.
- The only commonly used control chart that cannot be normalized is the Individuals (I) chart.

The Solution

- A normalized version of the Individuals (I) chart referred to as the the Normalized (I_N)

The Data

- **Values: $X_1, X_2, \ldots, X_n$**
- **Opportunities: $O_1, O_2, \ldots, O_n$**
- **Examples:**
  - **Complaints where the volumes vary**
  - **Process with lot-to-lot variation where the sample size varies (I –chart of lot averages)**
  - **Linear trends with unequally spaced samples checking for OOT points**

Complaint Data

- Donald Wheeler recommends an I chart for count data
- Richard Laney (2002) points out that the I chart cannot be normalized to account for differences in sample size or opportunities, resulting in constant control limits.
  - He offers the Laney U' chart as a solution
  - This is specific to counts
- For count data the I_N chart is nearly identical to the Laney U' chart. However, the I_N chart has other applications as well.

Assumptions

- **Assumption:**

$$X_i \sim N\left(\mu_i O_i, \sigma\sqrt{O_i}\right)$$

- **Based on the property of addition:**

$$X = Y_1 + Y_2 + Y_3 + \cdots + Y_O$$

$$\mu_X = \mu_Y O \qquad \sigma_X = \sigma_Y \sqrt{O}$$

Control Limits

- **The normalized values are then:**

$$N_i = \frac{X_i}{O_i} \sim N\left(\mu_i, \frac{\sigma}{\sqrt{O_i}}\right)$$

- **Control Limits:**

$$CL_i = \bar{N} \pm 3\,\frac{\bar{S}}{\sqrt{O_i}}$$

$$\bar{N} = \frac{X_1 + X_2 + \ldots + X_n}{O_1 + O_2 + \ldots + O_n}$$

Estimating S

- Start with:

$$N_i - N_{i-1} \sim N\left(\mu_i - \mu_{i-1}, \sigma\sqrt{\frac{1}{O_i} + \frac{1}{O_{i-1}}}\right)$$

- The following has the standard half-normal distribution:

$$\frac{\left|N_i - N_{l-1}\right|}{\sigma\sqrt{\frac{1}{O_i} + \frac{1}{O_{i-1}}}}$$

Standard half-normal distribution

Mean:	$\sqrt{\frac{2}{\pi}}$
Median:	$\Phi^{-1}(0.75)$
Standard Deviation:	$\sqrt{1-\frac{2}{\pi}}$

Standard Deviation

- **Unbiased estimate of the standard deviation:**

$$S_i = \sqrt{\frac{\pi}{2}} \frac{\left|N_i - N_{I-1}\right|}{\sqrt{\frac{1}{O_i} + \frac{1}{O_{i-1}}}}$$

Standard Deviation

- **Average estimator:**

$$\bar{S} = \text{Average}\left(S_2, S_3, \cdots, S_n\right)$$

- **Median estimator:**

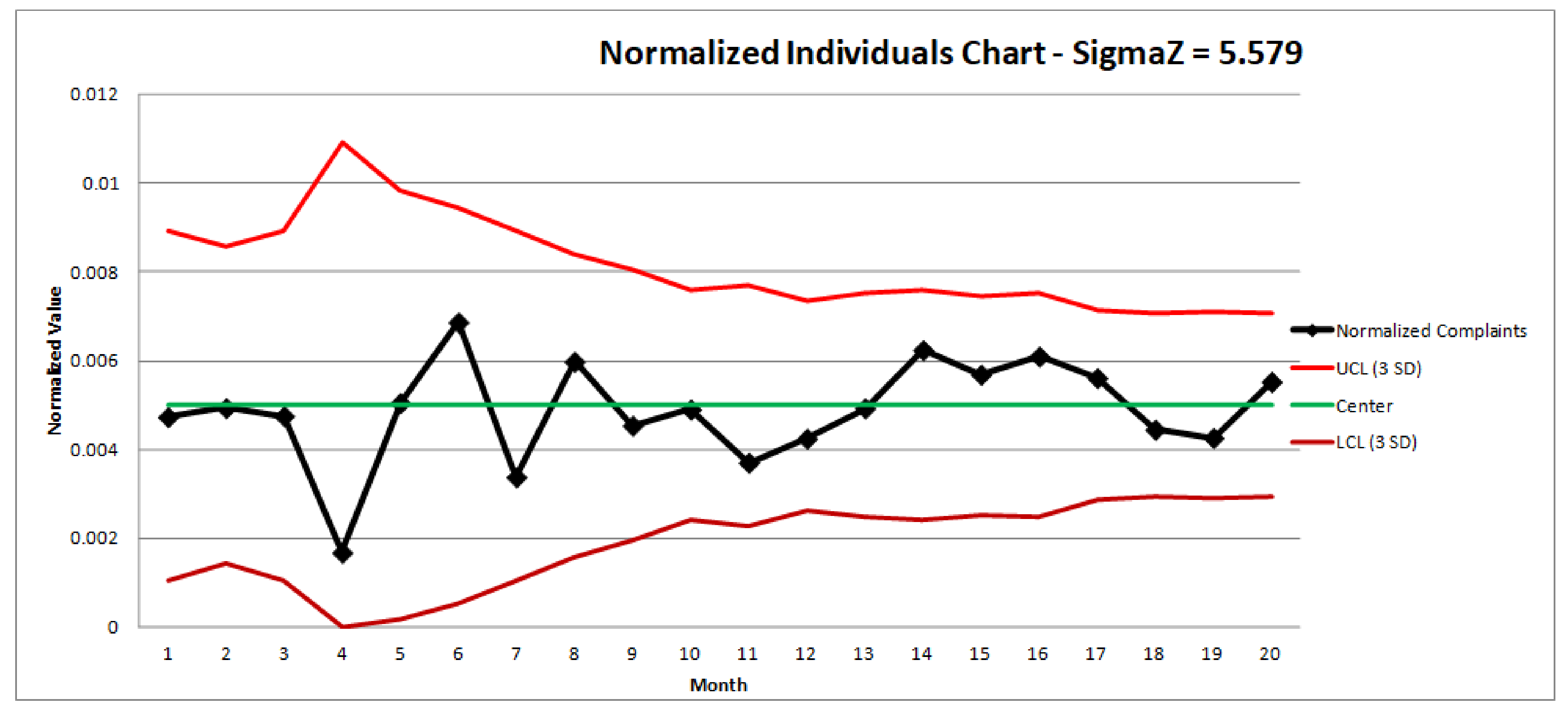
$$\tilde{S} = \sqrt{\frac{2}{\pi}}\,\frac{\text{Median}\left(S_2, S_3, \cdots, S_n\right)}{\Phi^{-1}\left(0.75\right)}$$

TE

Table 4: Example Complaint Data

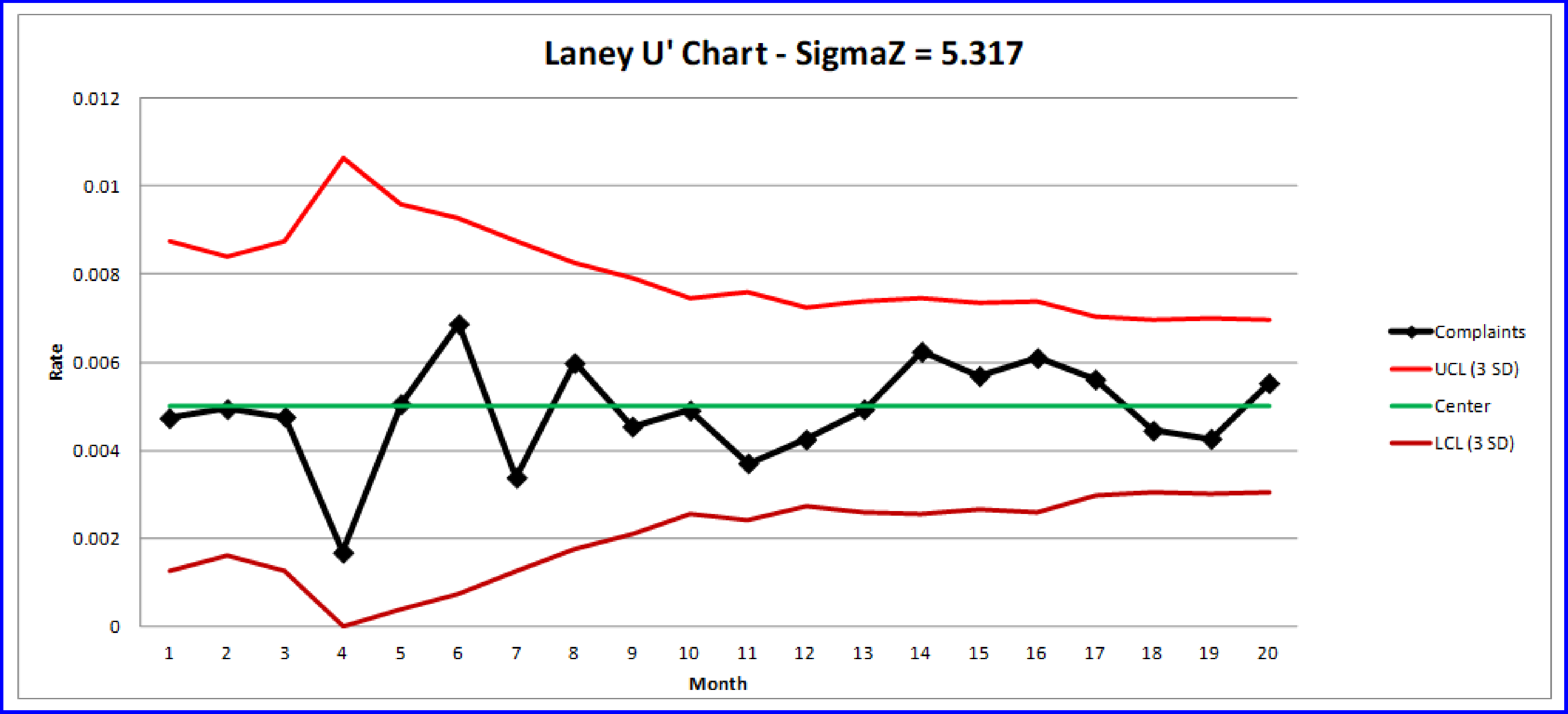
Month	Complaints (X_i)	Sales Volume (O_i)	Normalized Complaints Rates (N_i)
1	426	90000	0.004733333
2	543	110000	0.004936364
3	428	90000	0.004755556
4	67	40000	0.001675
5	303	60000	0.00505
6	481	70000	0.006871429
7	304	90000	0.003377778
8	718	120000	0.005983333
9	681	150000	0.00454
10	1030	210000	0.004904762
11	704	190000	0.003705263
12	1062	250000	0.004248
13	1085	220000	0.004931818
14	1311	210000	0.006242857
15	1309	230000	0.005691304
16	1342	220000	0.0061
17	1740	310000	0.005612903
18	1468	330000	0.004448485
19	1364	320000	0.0042625
20	1824	330000	0.005527273

I_N - Complaints



TE

Laney U' - Complaints



TE

Difference

- **Taylor:**

$$\bar{S}_2 = \sqrt{\frac{\pi}{2}}\,\frac{\left|N_2 - N_1\right|}{\sqrt{\frac{1}{O_2} + \frac{1}{O_1}}}$$

- **Laney:**

$$\hat{S}_2 = \frac{\sqrt{\pi}}{2}\left|\frac{N_2 - \bar{N}}{\sqrt{\frac{1}{O_2}}} - \frac{N_1 - \bar{N}}{\sqrt{\frac{1}{O_1}}}\right|$$

Comparison

Table 3: Comparison of Laney and Taylor Estimators of the Moving Standard Deviation

			$\hat{S}_2$		$\bar{S}_2$	
O_1	O_2	O_3	Average	SD	Average	SD
1	1	1	1.000	0.756	1.000	0.756
1	10	1	0.897	0.658	1.000	0.756
1	1	10	1.000	0.756	1.000	0.756
1	10	10	0.943	0.712	1.000	0.756
1	1	100	1.000	0.756	1.000	0.756
1	10	100	0.989	0.748	1.000	0.756
1	1	1000	1.000	0.756	1.000	0.756
1	10	1000	0.999	0.755	1.000	0.756

O_3 is number of opportunities for all the other data points combined

TE

Comparison

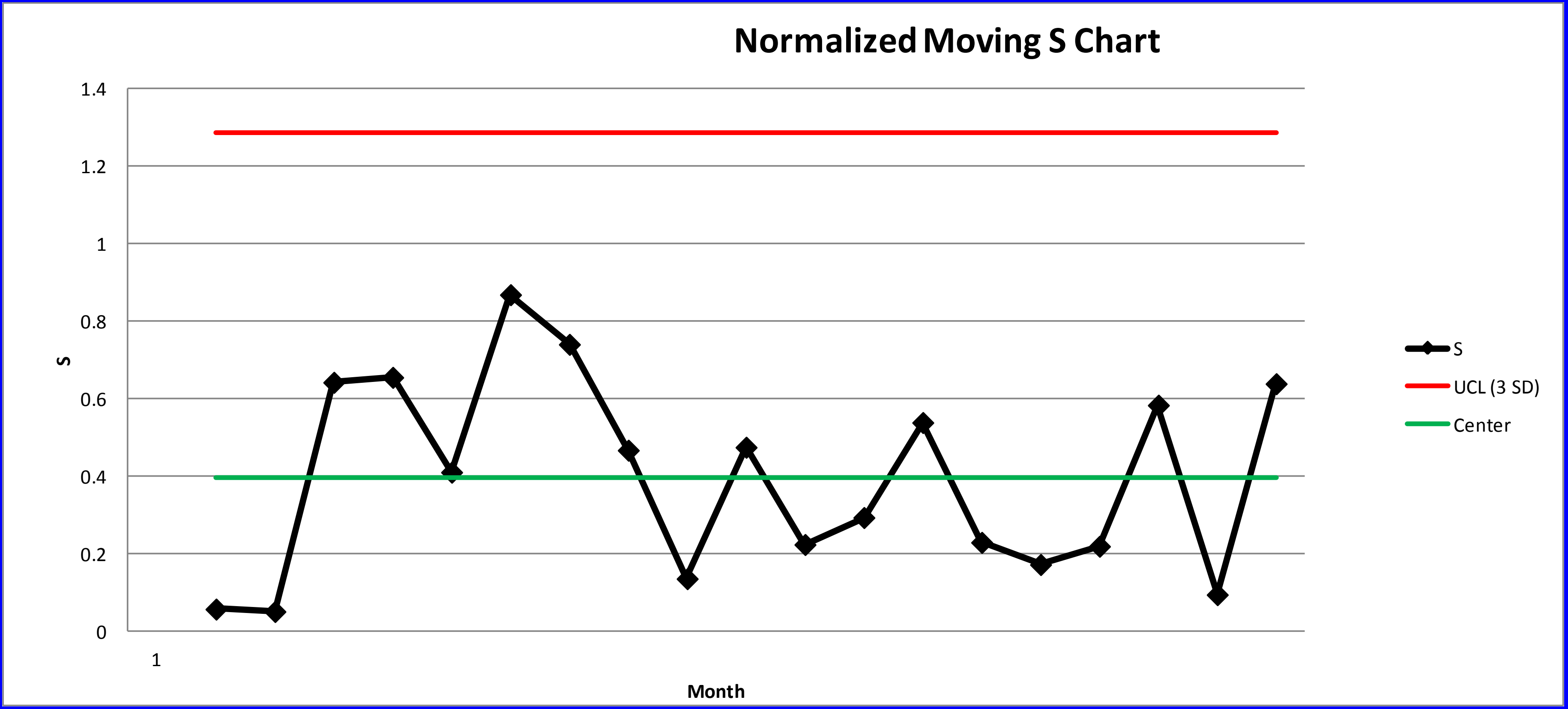
- Taylor estimator
  - Unbiased
- Laney estimator
  - May be biased but bias is small when there is lots of data
- From a practical point of view it makes little difference which is used.

SigmaZ

- **Ratio of the standard deviation to that of the Poisson distribution**

$$\sigma_Z = \frac{\bar{S}}{\sqrt{\bar{N}}}$$

Moving S Chart



Moving SigmaZ

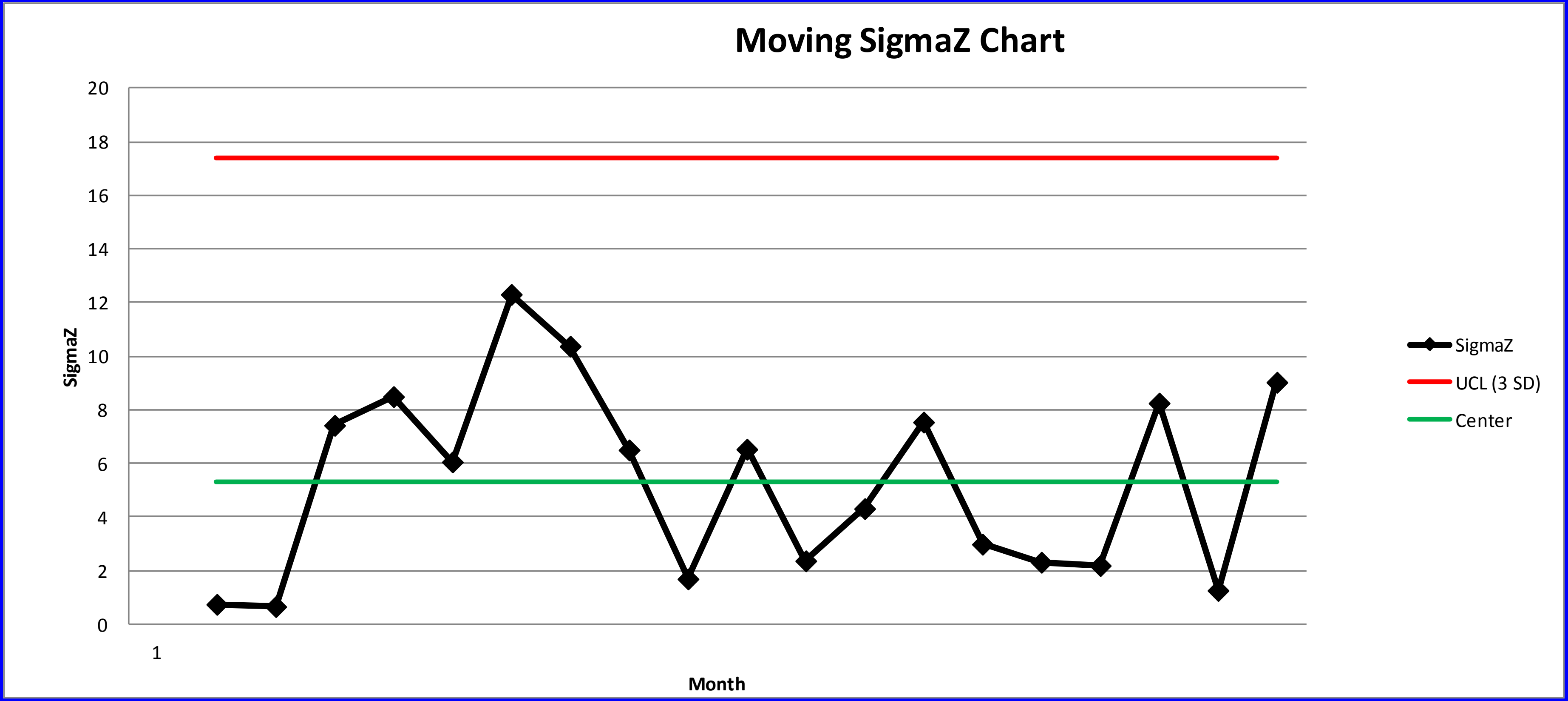


Table 5: Example Between Lot Variation Data

Lot	Value (Sum)	N	Normalized Value (Average)
1	1265.983327	13	97.38333286
2	1265.494062	13	97.34569709
3	1234.095432	13	94.93041783
4	467.0938368	5	93.41876736
5	480.5207487	5	96.10414975
6	1263.362522	13	97.18173242
7	471.6089798	5	94.32179597
8	483.3272231	5	96.66544462
9	1247.482101	13	95.96016158
10	1252.155511	13	96.31965468
11	475.8327811	5	95.16655622
12	1244.424527	13	95.72496359
13	479.2992493	5	95.85984986
14	1225.883257	13	94.29871211
15	1246.284863	13	95.86806642
16	1247.863834	13	95.98952569
17	476.4285921	5	95.28571841
18	480.9507213	5	96.19014427
19	1246.33594	13	95.8719954
20	1244.075692	13	95.69813015

I_N – Between Lots

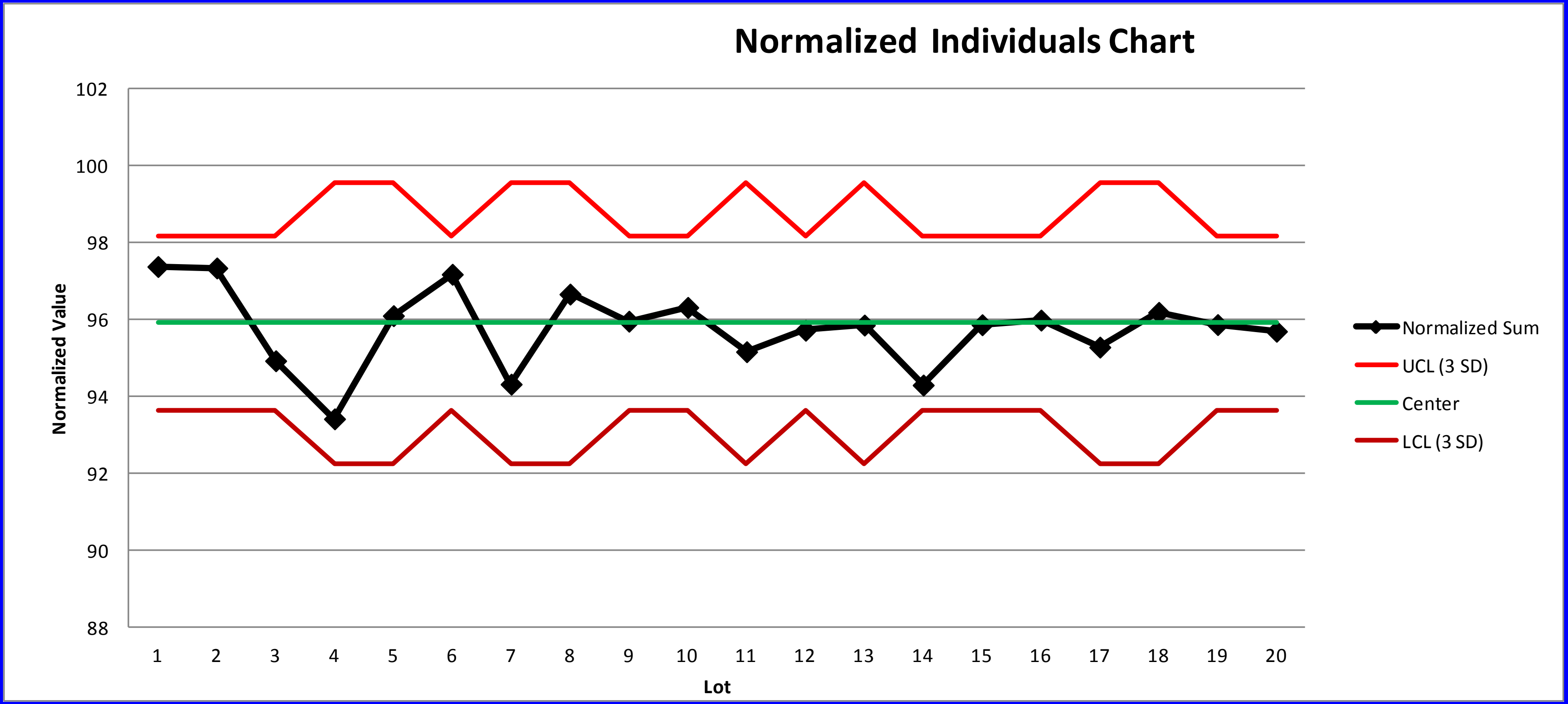
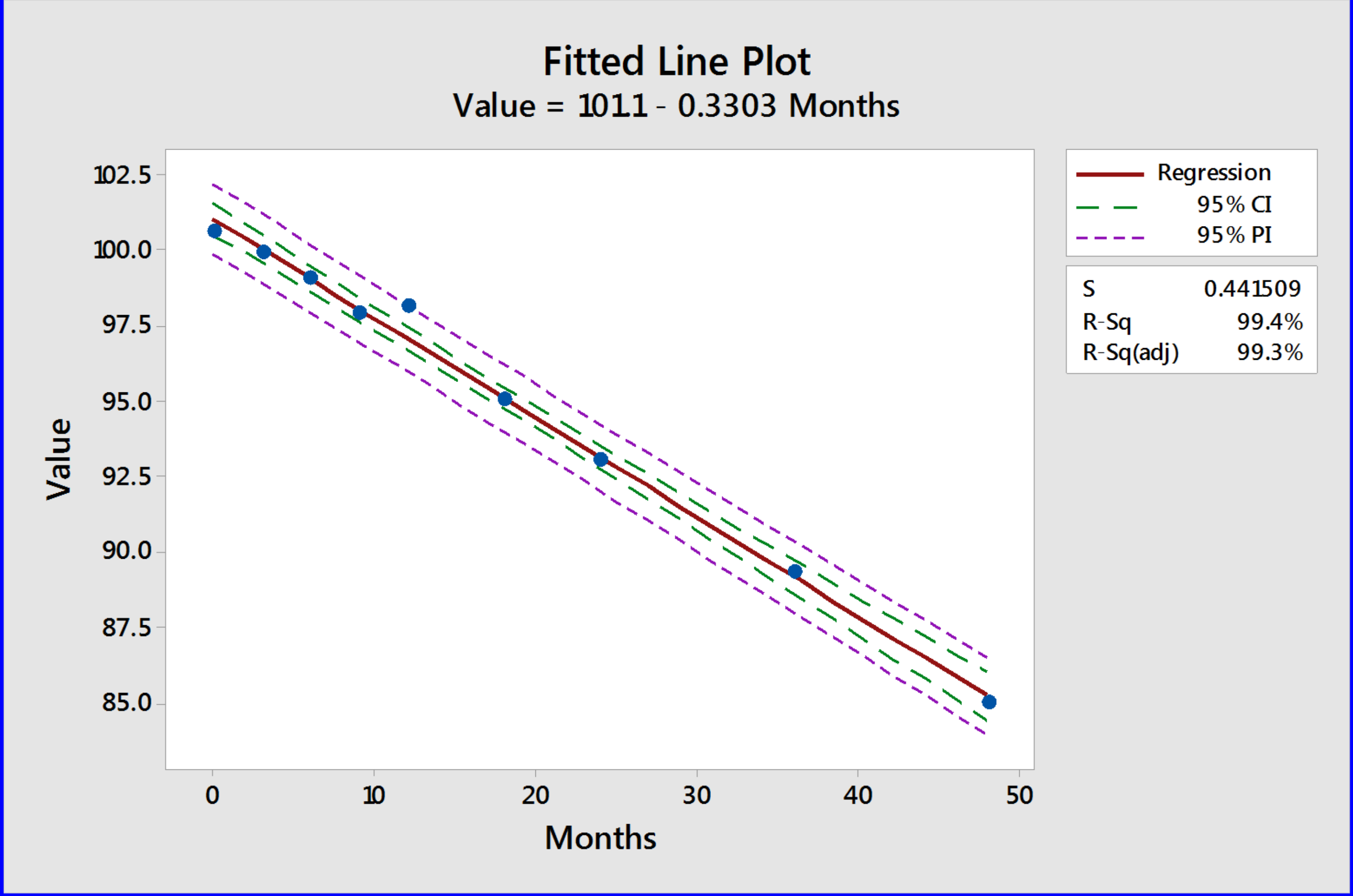


Table 6: Example Stability Data

Month	Value
0	100.8281
3	99.8719
6	99.16905
9	98.08248
12	98.12253
18	94.95572
24	93.00887
36	88.96801
48	85.21958

Regression

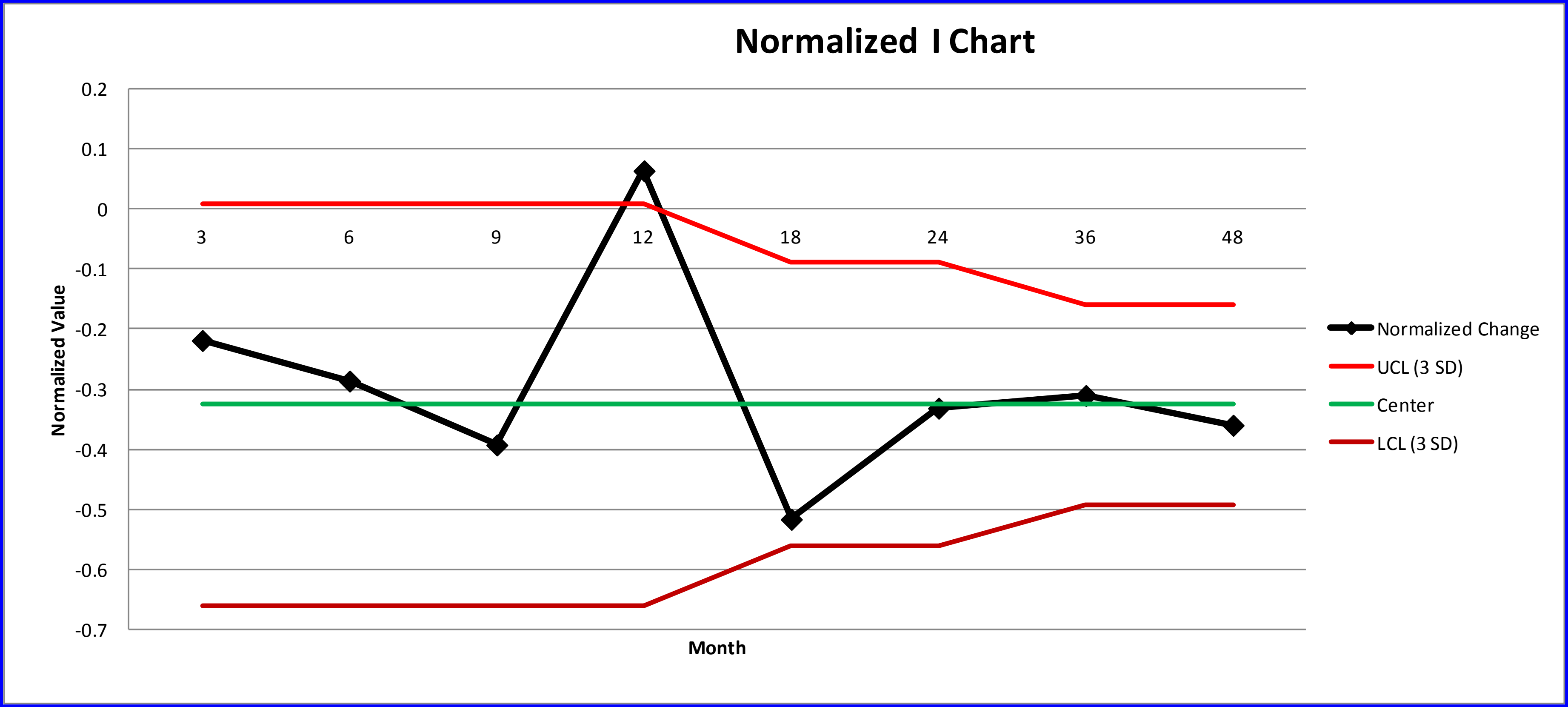


Fitted Line Plot
Value = 101.1 - 0.3303 Months
Regression
95% CI
95% PI
S 0.441509
R-Sq 99.4%
R-Sq(adj) 99.3%
Value
102.5
100.0
97.5
95.0
92.5
90.0
87.5
85.0
0
10
20
30
40
50
Months

Table 7: Changes and Slopes of Example Stability Data

Month	Change	Length Interval	Normalized Change (Slope)
3	-0.65563532	3	-0.218545106
6	-0.85848342	3	-0.286161138
9	-1.17796488	3	-0.392654961
12	0.193200231	3	0.064400077
18	-3.09996878	6	-0.516661463
24	-1.98809354	6	-0.331348924
36	-3.72314215	12	-0.310261846
48	-4.31874831	12	-0.359895693

I_N – Slope



Normalized I Chart
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
-0.7
3
6
9
12
18
24
36
48
Normalized Value
Month
Normalized Change
UCL (3 SD)
Center
LCL (3 SD)

TE

Conclusions

- The I_N chart handles count and pass/fail data where a Laney U’ or P’ chart might be used. It also handles many other situations involving non-count data where a Laney U’ or P’ chart do not apply.
- The X-bar and I_N charts handle most needs, simplifying the selection of a chart. These are the only 2 charts needed in most cases. The decision between them is based on whether there are multiple values per time period or not.

Conclusions

- The exception to this rule is nonnormal data. One such case is when counts are low and follow the binomial or Poisson distributions. In this case U and P charts with adjusted control limits should be used as described in Taylor (2017a, b).
- If the Laney U' or P' chart is used, consider accompanying it with a moving SigmaZ chart. It is inconsistent to show Xbar-S, and I-Moving S charts, but not to do the same for a Laney U' or P' chart, as they are all based on a time ordered series of estimates of the standard deviation.

References

- Laney, David (2002), Improved Control Charts for Attributes, Quality Engineering, 14(4), 531–537.
- Wheeler, Donald (2011), What About p-Charts?, Quality Digest, http://www.qualitydigest.com/inside/quality-insider-article/what-about-p-charts.html.
- Taylor, Wayne (2017d), Normalized Individuals (I_N) Chart, Taylor Enterprises, Inc., www.variation.com/techlib/brief1.html.
- Taylor, Wayne (2017a), Adjusted Control Limits for U Charts, Taylor Enterprises, Inc., www.variation.com/techlib/brief2.html.
- Taylor, Wayne (2017b), Adjusted Control Limits for P Charts, Taylor Enterprises, Inc., www.variation.com/techlib/brief3.html.
- Taylor, Wayne (2017c), Statistical Procedures for the Medical Device, Taylor Enterprises, Inc., www.variation.com/procedures.