



Advanced Topics in Statistical Process Control

The Power of Shewhart's Charts

Second Edition

Donald J. Wheeler



SPC Press
Knoxville, Tennessee



Contents

	Preface to the Second Edition	viii
	Preface	ix
	The Shewhart Catechism	x
Chapter One	Shewhart's Control Charts	1
	1.1 Shewhart's Definition of Statistical Control	2
	1.2 Shewhart's "Operation of Statistical Control"	4
	1.3 The Four Possibilities for Any Process	6
	1.4 Process Behavior Charts and the Probability Approach	14
	1.5 Enumerative Studies and Analytic Studies	18
	1.6 The Uses of Shewhart's Charts	19
	1.7 The Four Foundations of Shewhart's Charts	21
	1.8 Process Behavior Charts and Chaos Theory	25
	1.9 The Transformation of Data	28
Chapter Two	Statistics Versus Parameters	31
	2.1 Measures of Location and Dispersion	31
	2.2 The Notion of a Probability Distribution	33
	2.3 Statistics, Parameters, and Process Behavior Chart Limits	37
	2.4 Skewness and Kurtosis	46
Chapter Three	Estimating Dispersion Parameters	55
	3.1 Estimators of $SD(X)$ and $V(X)$ for <i>One</i> Subgroup of Size n	56
	3.2 Three Ways to Estimate $SD(X)$ for k Subgroups of Size n	60
	3.3 The Second Foundation of Shewhart's Charts	71
	3.4 Within-Subgroup Estimates of Dispersion	74
	3.5 Time Series Data (Subgroups of Size One)	79
	3.6 Degrees of Freedom	80
	3.7 Summary	83
Chapter Four	Process Behavior Charts for Measurements	85
	4.1 Average and Range Charts	88
	4.2 Average and Root Mean Square Deviation Charts	92
	4.3 Average and Standard Deviation Charts	96
	4.4 Average and Variance Charts	103
	4.5 So What's the Difference?	106
	4.6 Charts for Individual Values	108
	4.7 Median and Range Charts	113

Contents

Chapter Five	Three Sigma Limits	115
	5.1 Why Three-Sigma Limits?	116
	5.2 But What About the Central Limit Theorem?	120
	5.3 The Recalculation of Limits	122
	5.4 The Role of the Normal Distribution	123
	5.5 Calculating Limits With Small Amounts of Data	126
	5.6 Inadequate Measurement Units	128
	5.7 Detecting Assignable Causes	135
	5.8 Summary	140
Chapter Six	Rational Sampling and Rational Subgrouping	141
	6.1 Rational Sampling	142
	6.2 Rational Subgrouping	143
	6.3 Principles for Subgrouping	155
	6.4 Three-Way Charts	157
	6.5 Summary	165
Chapter Seven	The Quality of the Limits	167
	7.1 The Distribution of the Average Range Statistic	168
	7.2 The Distribution of the Average Root Mean Square Statistic	170
	7.3 The Distribution of the Average Standard Deviation Statistic	172
	7.4 The Distribution of the Pooled Variance Statistic	174
	7.5 The Distribution of the Median Range Statistic	176
	7.6 The Distribution of the Average Moving Range	178
	7.7 Degrees of Freedom	180
	7.8 Effective Degrees of Freedom	182
	7.9 Using Degrees of Freedom	183
	7.10 The Quality of Process Behavior Chart Limits	185
	7.11 Summary	186
Chapter Eight	Capability Confusion	187
	8.1 Capability, Capability Ratios, and Performance Ratios	188
	8.2 The Capability of a Predictable Process	188
	8.3 Capability Measures	191
	8.4 Converting Capabilities Into Fraction Nonconforming	194
	8.5 Performance Ratios	197
	8.6 World-Class Quality	197
	8.7 Summary	204
Chapter Nine	Power Functions for Process Behavior Charts	205
	9.1 The Power of Charts for Location	206
	9.2 The Formulas for the Power Functions	212
	9.3 The Average Run Length for a Procedure	221
	9.4 Summary	224

Contents

Chapter Ten	Comparing Different Types of Process Behavior Charts	225
	10.1 Process Behavior Charts for Subgroup Averages	227
	10.2 Charts for Moving Averages	237
	10.3 Moving Averages or Individual Values?	239
	10.4 <i>Contra</i> Two-Sigma Limits	245
	10.5 Summary	250
Chapter Eleven	Charts for Count Data	251
	11.1 Charts for Count Data	252
	11.2 <i>np</i> -Charts	255
	11.3 <i>p</i> -Charts	256
	11.4 A Problem With Binomial Counts	257
	11.5 <i>c</i> -Charts	259
	11.6 <i>u</i> -Charts	261
	11.7 A Problem With Poisson Counts	262
	11.8 Degrees of Freedom	267
	11.9 Process Behavior Charts for Rare Events	267
	11.10 Summary	271
Chapter Twelve	Charts for Autocorrelated Data	273
	12.1 Correlation	273
	12.2 Significant Correlations	275
	12.3 Autocorrelation	276
	12.4 What Does Autocorrelation Say About the Process?	278
	12.5 The Effect of Autocorrelation	279
	12.6 Autocorrelation's Effect Upon Chart Limits	282
	12.7 So What Shall I Do With My Autocorrelated Data?	287
Chapter Thirteen	The Cumulative Sum Technique	289
	13.1 Cumulative Sums for Individual Values	290
	13.2 Cumulative Sums for Subgroup Averages	299
	13.3 Average Run Lengths for Selected Cusum Procedures	307
	13.4 Summary	311
Chapter Fourteen	Exponentially Weighted Moving Averages	313
	14.1 The EWMA Algorithm	313
	14.2 EWMA for Individual Values	315
	14.3 Average Run Lengths for the EWMA	320
	14.4 Then What Does an EWMA Do?	323
	14.5 Summary	327

Contents

Chapter Fifteen	Multivariate Charts	329
	15.1 Simple Multivariate Charts	329
	15.2 Charts for Hotelling's T^2 Statistic	334
Chapter Sixteen	Miscellaneous Techniques	341
	16.1 Zone Charts	341
	16.2 Modified Control Charts	346
	16.3 PreControl	347
	16.4 Summary	354
Chapter Seventeen	The Characterization of Product	355
	17.1 Characterizing the Measured Item	356
	17.2 Characterizing Product Which Has Not Been Measured	366
	17.3 The Fallacy of Acceptance Sampling	368
	17.4 Estimating the Fraction Nonconforming	369
	17.5 Summary	371
Chapter Eighteen	The Analysis of Means	373
	18.1 Comparing k Treatments Using Average Charts	374
	18.2 Experimental Data vs. Production Data	375
	18.3 The Decision Limits for ANOM	379
	18.4 The Role of the Range Chart with ANOM	382
	18.5 ANOM With Multifactor Studies	386
	18.6 Interaction Effects and ANOM	388
	18.7 ANOM With Unequal Subgroup Sizes	390
	18.8 The Analysis of Mean Ranges	392
	18.9 Range-Based Analysis of Means	395
	18.10 Summary	397
Chapter Nineteen	Some Differences Between Theory and Practice	399
	19.1 The Scientific Process	400
	19.2 Statistical Tests of Hypotheses	402
	19.3 Enumerative Studies	404
	19.4 Analytical Studies	406
	19.5 Shewhart on Probability Models and Processes	407
	19.6 Finding Decision Rules for Analytic Studies	408
	19.7 Conclusions	409

Contents

Appendix		411
Glossary of Symbols		412
Bibliography		414
Tables		416-464
Table 1:	Bias Correction Factors	416
Table 2:	Bias Correction Factors for Estimating $V(X)$	417
Table 3:	Charts for Individual Values and Moving Ranges	418
Table 4:	Average and Range Charts Using the Average Range	419
Table 5:	Average and Range Charts Using the Median Range	420
Table 6:	Average and Range Charts Based on $\sigma(X)$	421
Table 7:	Average and RMS Dev. Charts Using Average RMS Dev.	422
Table 8:	Average and RMS Dev. Charts Using Median RMS Dev.	423
Table 9:	Average and RMS Dev. Charts Based on $\sigma(X)$	424
Table 10:	Average and Std. Dev. Charts Using Average Std. Dev.	425
Table 11:	Average and Std. Dev. Charts Using Median Std. Dev.	426
Table 12:	Average and Std. Dev. Charts Based on $\sigma(X)$	427
Table 13:	Average and Std. Dev. Charts Using Pooled Variance	428
Table 14:	Median and Range Charts Using the Average Range	430
Table 15:	Median and Range Charts Using the Median Range	431
Table 16:	Median and Range Charts Based on $\sigma(X)$	432
Table 17:	Average and Variance Charts	433
Table 18:	Moments for Range Distributions	434
Table 19:	Moments for Distributions of the RMS Deviation	436
Table 20:	Moments for Distributions of the Standard Deviation	438
Table 21:	Degrees of Freedom for Average Ranges	442
Table 22:	Degrees of Freedom for Median Ranges	444
Table 23:	Degrees of Freedom for Average Moving Ranges	446
Table 24:	Degrees of Freedom for Median Moving Ranges	447
Table 25:	Degrees of Freedom for Average Std. Dev. and RMS Dev.	448
Table 26:	Degrees of Freedom for Median Std. Dev. and RMS Dev.	450
Table 27:	The Standard Normal Distribution	452
Table 28:	ANOM Critical Values, H	454
Table 29:	Range-Based ANOM Critical Values	458
Table 30:	Range-Based ANOME Critical Values	459
Table 31:	Range-Based ANOMR Critical Values	460
Table 32:	Central Lines and Upper Limits for Hotelling's T^2	462
Table 33:	37 Reference Tables and Figures Located Within Text	463
Table 34:	Alpha Levels for Average Charts	464
Index		465



The Shewhart Catechism

1. *What if a state of predictability is beyond the capability of our process?*
This cannot happen. A process is operating predictably when it is being operated as consistently as it can—no more, and no less. See Sections 1.1, 1.2
2. *Is quality defined by conformance to specifications?*
No. Today world-class quality is defined as “on-target with minimum variance.” Conformance to specifications is merely the starting point, not the finish line. See Section 8.6
3. *What do capability ratios do?*
Capability ratios may be used to compare the Voice of the Process with the Voice of the Customer. However, it is easy to go overboard with this. Some other measures are also useful, and no numerical summary is an adequate replacement for the basic graphs of the data. See Section 8.3
4. *How do Dr. Deming’s 14 Points relate to Statistical Process Control?*
They may be thought of as corollaries to the observation that organizations do not behave in a rational manner and therefore cannot effectively use the knowledge gained from the charts. See Section 1.6
5. *How do Shewhart’s process behavior charts differ from tests of hypotheses?*
Shewhart’s charts characterize a time series as being either predictable or unpredictable. When a test of hypotheses is applied to a time series it *presumes* the time-series to be predictable, and looks for changes in parameters of the predictable time-series. See Sections 1.3 through 1.5 and Chapter 19
6. *Does it matter how we calculate the limits of a process behavior chart?*
Yes. There are right and wrong ways of computing limits. While *all* of the right ways will yield essentially the same limits, the wrong ways will result in incorrect limits and faulty decisions. See Section 3.2
7. *Do we have to remove the “outliers” from the data before we calculate the limits?*
No. The correct computational procedures for process behavior chart limits are not severely affected by outliers. See Sections 3.2 and 4.1
8. *Must the data “be in control” prior to placing them on a process behavior chart?*
No. One purpose of a process behavior chart is to detect the presence or absence of a state of predictability. Therefore, you do not have to wait for “pristine” data, nor do you have to clean up the data prior to computing the limits. See Section 3.2 and Chapter 4





-
9. *How can we calculate good limits from data which are “out of control?”*
The correct approaches to computing limits will minimize the impact of exceptional variation upon the limits. See Chapter 4
10. *Should we calculate limits for the Average Chart when the Range Chart is unpredictable?*
Yes. You will not discover anything by *not* calculating limits!
See Sections 4.1 and 5.4
11. *Can we use two-sigma limits?*
No. Two-sigma limits are inappropriate for a process behavior chart. See Section 10.4
12. *Do three-sigma limits depend upon having normally distributed data?*
No. Three-sigma limits are sufficiently general to work with virtually any distribution.
See Sections 5.1 and 5.2
13. *Are the process behavior chart constants dependent upon having normally distributed data?*
No. These constants are remarkably similar for a wide range of different distributions.
See Section 5.4
14. *Do we have to wait until we have 25 or 30 subgroups before calculating limits?*
No. Limits may be calculated using small amounts of data. The degrees of freedom will characterize just how “reliable” the calculated limits have become.
See Sections 3.6, 5.4, and 7.9
15. *What about data which display autocorrelation?*
Shewhart placed autocorrelated data on charts. You can too. See Chapter 12
16. *Do we have to use subgroups of size 5?*
No. The subgroup size is not the most important characteristic of the chart. Rational subgrouping is more important than the subgroup size, even if that requires subgroups of size one.
See Sections 1.6, 10.1, and Chapter 6
17. *What is a rational subgrouping?*
Among other things, a rational subgrouping is one in which the measurements inside each subgroup are judged to have been obtained under essentially the same conditions. Moreover, the subgrouping must respect the structure of the data.
See Section 1.2 and Chapter 6
18. *Does the Central Limit Theorem require charts to use subgroups of size 4 or 5?*
While the Central Limit Theorem is true, it is not the basis on which the process behavior chart is built. Therefore, this theorem cannot *require* anything of a process behavior chart. The charts will work with subgroups of size 1, and they will do so even when the data are not normally distributed. See Sections 5.2, 10.1, and 10.3





-
19. *Does a subgroup always consist of consecutive measurements?*
No. While many processes may be tracked using subgroups of consecutive measurements, this is not a requirement of the chart. Once again, it is more important that the subgrouping be rational. See Section 1.6 and Chapter 6
20. *What happens to a process behavior chart when the measurement units are too large to detect the variation in the data?*
False alarms will proliferate. This problem is easy to spot, but it must be considered if you are to avoid looking for nonexistent assignable causes. See Section 5.5
21. *Why do we use so many symbols for dispersion?*
There is a multiplicity of dispersion statistics as well as a plurality of dispersion parameters. A lack of a standard nomenclature here is a real source of confusion. See Sections 3.1 through 3.4
22. *Is the standard deviation statistic better than the range?*
Not for small subgroup sizes. See Sections 4.1, 4.3, and Appendix Tables 18, 19, and 20
23. *Is not the Pooled Variance a better measure of dispersion than the Average Standard Deviation or the Average Range?*
Yes and No. The Pooled Variance will have more degrees of freedom than the other measures, but it will be more severely affected by a lack of homogeneity than will the other measures. See Sections 3.4 and 7.7
24. *What are Degrees of Freedom?*
Degrees of Freedom is the name given to a characterization of the variation of a measure of dispersion. The greater the Degrees of Freedom, the smaller the variation of the measure of dispersion. See Sections 3.6, 7.7, 7.8, and 7.9
25. *How do we measure dispersion when our subgroup size is one?*
We must use a two-point moving range to measure dispersion when placing a time series of individual values on a process behavior chart. See Sections 3.5 and 4.6
26. *What role does the Moving Range Chart play with an Individual Value Chart?*
It signals that the user knows the correct way of computing limits, it identifies breaks in the original time series, and it provides those who look at the chart with the ability to easily check that the computation of the limits is correct. See Section 4.6
27. *Are “unbiased” estimators better than “biased” estimators?*
No. The property of being unbiased does not imply closeness to the parameter. See Section 3.1
28. *Should I be using the skewness and kurtosis values from my printout?*
No. You will never have enough data to make these “shape statistics” useful. See Section 2.4





-
29. *Will Chaos Theory replace process behavior charts?*
No. Since Dr. Shewhart's approach is holistic and empirical, it can accommodate the paradigm shift represented by the new theories of quasi-chaotic variation.
See Section 1.7
30. *What about transforming the data prior to placing them on a chart?*
Transformations to achieve a quasi-normal histogram or transformations to achieve other statistical properties are not recommended. They introduce more complexity than they are worth. Transformations to achieve *clarity of interpretation* are always appropriate.
See Sections 1.8, 6.1, 11.9, and Chapter 19
31. *Is a Moving Average Chart better than a chart for Individual Values?*
Only occasionally. Most times the Individual Value Chart will be the better chart.
See Section 10.3
32. *What is the difference between Attribute Charts and XmR Charts?*
XmR Charts use empirical limits while Attribute Charts use theory to construct limits. Therefore, with Attribute Charts you must check to see if the theory is appropriate.
See Chapter 11
33. *How can we chart rare events?*
Counts of rare events will result in weak charts. It is better to measure the area of opportunity between the rare events and use these measurements on the charts.
See Section 11.9
34. *Is it true that the Cumulative Sum (Cusum) is better than a process behavior chart?*
No. The Cusum technique may be easier to program than a process behavior chart, but contrary to much of what has been written, it does not really work any better than process behavior charts. Moreover, by transforming the original data, the Cusum technique can actually obscure that which is easily seen in the running records of the process behavior charts.
See Chapter 13
35. *Can we use an Exponentially Weighted Moving Average instead of a process behavior chart?*
An EWMA attempts to model a time series. Such models may be used for process monitoring. However, EWMA's are no more sensitive than a process behavior chart, they will always lag behind a chart, and they are more complex than a chart.
See Chapter 14
36. *Can we use PreControl instead of a process behavior chart?*
No. PreControl is totally different from a process behavior chart, and inferior as well.
See Section 16.3
37. *Are Zone Charts a valid alternative to a process behavior chart?*
No. Zone charts result in too many false alarms. In addition, they are actually more complex than a process behavior chart, while providing nothing more than a parody of the process behavior chart.
See Section 16.1

