

P' Chart

Summary	1
Data Input.....	3
P' Chart.....	4
P' Chart Report.....	7
Analysis Summary	8
Analysis Options.....	9
Runs Tests.....	13
Save Results	15
Calculations.....	15
References.....	15

Summary

The **P' Chart** procedure creates a control chart for data that describe the proportion of times an event occurs in m samples taken from a product or process. The data might represent the proportion of defective items in a manufacturing process, the proportion of customers that return a product, or any other event that has exactly 2 possible outcomes. Unlike the standard P chart, however, the P' chart does not assume that the samples all come from a binomial distribution with the same mean. It instead allows the mean proportion to vary over the sampling period, a phenomenon known as *overdispersion*. The methodology used in constructing the chart is due to Laney (2002).

Out-of-control signals are highlighted, including both points beyond the control limits and any unusual runs in the data. The chart may be constructed in either *Initial Study* (Phase 1) mode, where the current data determine the control limits, or in *Control to Standard* (Phase 2) mode, where the limits come from either a known standard or from prior data.

The data for this analysis consist of m samples from a population detailing:

n_j = number of items in sample j

d_j = number of defective or otherwise unacceptable items in sample j

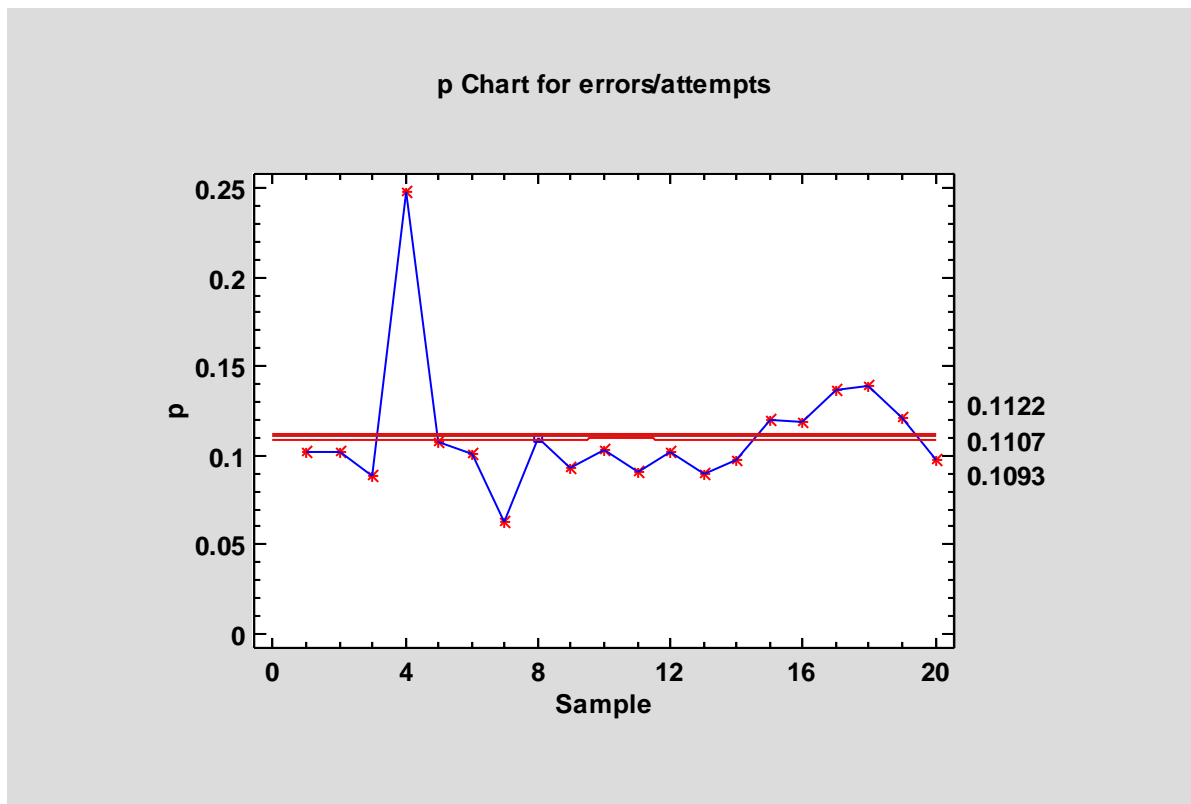
Sample StatFolio: *pprimechart.sgp*

Sample Data

The file *internet access.sgd* contains data presented by Hung describing the number of Internet access attempts on $m = 20$ consecutive days, together with the number of attempts that returned an error message. The first several rows of that file are shown below:

Day	Attempts	Errors
1	412,670	42,104
2	395,736	40,286
3	401,765	35,399
4	395,422	97,981
5	422,223	45,346
6	433,234	43,699

As is typical with very large samples, the variability between days greatly exceeds that expected from a binomial distribution. The limits on a standard P chart would therefore be much too tight and of little use in separating special variation from common cause variation:



The P' Chart widens the limits to allow for changes in the binomial mean.

Data Input

The data required by this procedure is a numeric column containing the proportions of unacceptable items in the m samples.

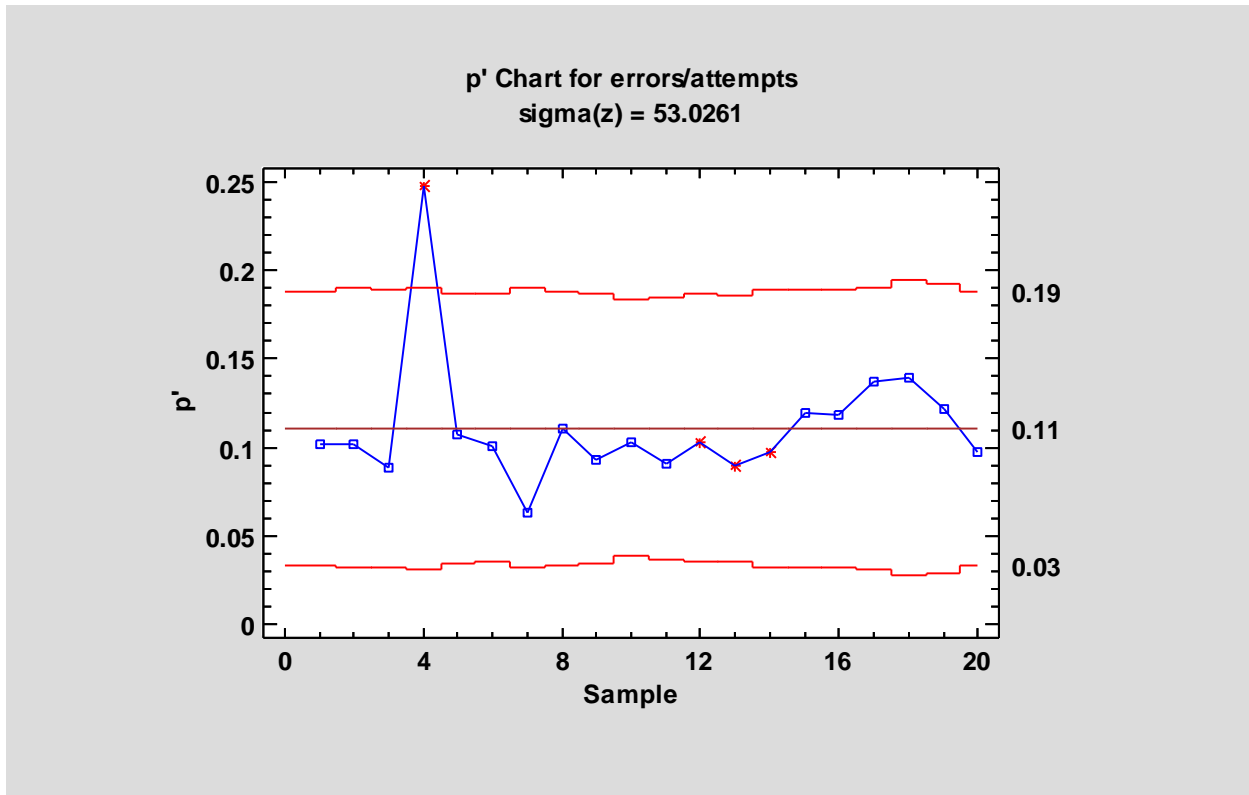
- **Proportions:** a column containing the proportion of unacceptable items in each sample.
- **Sample Sizes:** the number of items in each sample. If all sample sizes are equal, you can enter a single value such as “50”. Otherwise, enter a column with the sizes corresponding to each proportion.
- **Sample Labels:** optional labels for each sample.
- **LSL, Nominal, USL:** optional lower specification limit, nominal (target) value, and upper specification limit.
- **Select:** subset selection.

In this example, the proportions are calculated “on-the-fly” by dividing the numbers of errors by the number of attempts:

$$p_j = \frac{d_j}{n_j} \quad (1)$$

P' Chart

The *P' Chart* plots the sample proportions p_j :



In Phase 1 (*Initial Studies*) mode, the centerline and control limits are determined from the data. The centerline is located at the weighted average of the sample proportions:

$$\bar{p} = \frac{\sum_{j=1}^m n_j p_j}{\sum_{j=1}^m n_j} \tag{2}$$

The control limits are placed above and below the centerline at:

$$\bar{p} \pm k \sqrt{\frac{\bar{p}(1-\bar{p})}{n}} \hat{\sigma}_z \tag{3}$$

where k is the sigma multiple specified on the *Control Charts* tab of the *Preferences* dialog box ($k = 3$ except in rare cases), n is the sample size, and $\hat{\sigma}_z$ is an estimate of the relative amount of process variation not explained by the binomial distribution alone.

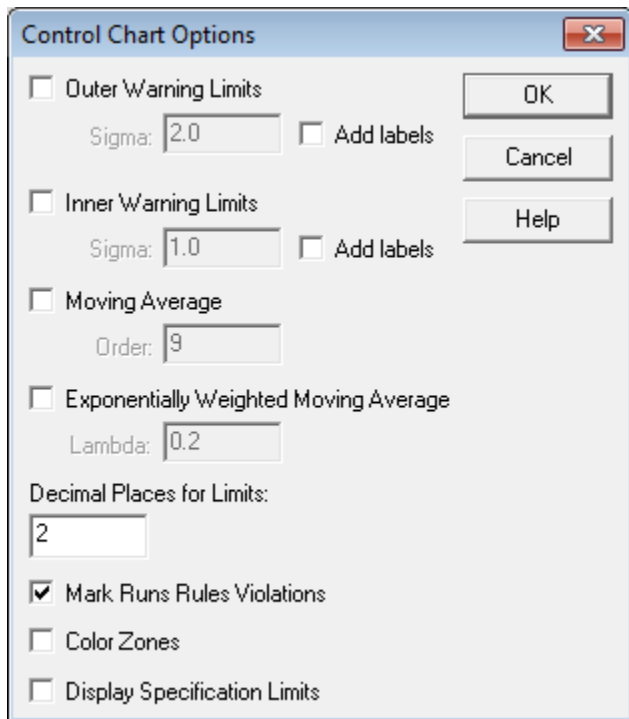
If the sample sizes are not equal, then depending on the *Analysis Options* settings, n is replaced by either:

(1) \bar{n} , the average sample size. In this case, the control limits are the same for all samples.

(2) n_j , the individual sample sizes. In this case, the control limits are step functions.

Any points beyond the control limits will be flagged using a special point symbol. Any point excluded from the analysis, usually by clicking on a point on the chart and pressing the *Exclude/Include* button, will be indicated by an X. If so indicated on the *Pane Options* dialog box, unusual sequences of points may also be flagged. In the current chart, there is one value beyond the control limits and 3 values flagged because of a run of 8 or more points below the centerline.

Pane Options



- **Outer Warning Limits:** check this box to add warning limits at the specified multiple of sigma, usually at 2 sigma.
- **Inner Warning Limits:** check this box to add warning limits at the specified multiple of sigma, usually at 1 sigma.
- **Moving Average:** check this box to add a moving average smoother to the chart. In addition to the sample proportions, the average of the most recent q points will also be displayed, where q is the *order* of the moving average. The default value $q = 9$ since the 1-sigma inner warning limits for the original proportions are equivalent to the 3-sigma control limits for that order moving average.

- **Exponentially Weighted Moving Average:** check this box to add an EWMA smoother to the chart. In addition to the sample proportions, an exponentially weighted moving average of the sample proportions will also be displayed, where λ is the smoothing parameter of the EWMA. The default value $\lambda = 0.2$ since the 1-sigma inner warning limits for the original proportions are equivalent to the 3-sigma control limits for that EWMA.
- **Decimal Places for Limits:** the number of decimal places used to display the control limits.
- **Mark Runs Rules Violations:** flags with a special point symbol any unusual sequences or runs. The runs rules applied by default are specified on the *Runs Tests* tab of the *Preferences* dialog box.
- **Color Zones:** check this box to display green, yellow and red zones.
- **Display Specification Limits:** whether to add horizontal lines to the chart displaying the location of the specification limits (if any).

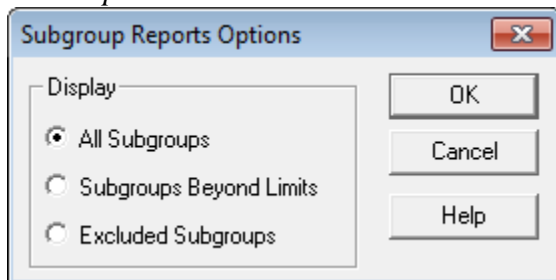
P' Chart Report

This pane tabulates the values plotted on the control chart:

p' Chart Report		
All Samples		
X = Excluded * = Beyond Limits		
Sample	Size	p
1	412670	0.102028
2	395736	0.1018
3	401765	0.0881087
4	395422	* 0.247788
5	422223	0.107398
6	433234	0.100867
7	396788	0.0623809
8	411383	0.110338
9	423348	0.0925456
10	474053	0.102689
11	446823	0.0904273
12	431661	0.102391
13	434353	0.0898969
14	406232	0.0971243
15	402454	0.119994
16	403312	0.11832
17	387782	0.137121
18	355500	0.139167
19	372441	0.121421
20	415813	0.0975992

Out-of-control points are indicated by an asterisk. Points excluded from the calculations are indicated by an X.

Pane Options



- **Display:** specify the samples (subgroups) to display in the report.

Analysis Summary

The *Analysis Summary* summarizes the data and the control chart:

p' Chart - errors/attempts
 Number of subgroups = 20
 Average subgroup size = 411150.
 0 subgroups excluded

p' Chart

<i>Period</i>	<i>#1-20</i>
UCL: +3.0 sigma	0.188119
Centerline	0.110712
LCL: -3.0 sigma	0.0333051

1 beyond limits

Estimates

<i>Period</i>	<i>#1-20</i>
Process mean	0.110712
Sigma	0.0259483
Sigma(z)	53.0261

Included in the table are:

- **Sample Information:** the number of samples *m* and the average sample size

$$\bar{n} = \frac{\sum_{j=1}^m n_j}{m} \tag{4}$$

If any samples have been excluded from the calculations, that number is also displayed.

- **p' Chart:** a summary of the centerline and control limits for the p' chart. Using *Analysis Options*, separate limits may be calculated for different periods (sets of samples).
- **Estimates:** estimates of the process mean proportion of defective items:

$$\hat{\theta} = \bar{p} \tag{5}$$

and the process sigma

$$\hat{\sigma} = \sqrt{\frac{\bar{p}(1-\bar{p})}{\bar{n}}} \hat{\sigma}_z \tag{6}$$

The estimated relative amount of additional process variation $\hat{\sigma}_z$ is also displayed.

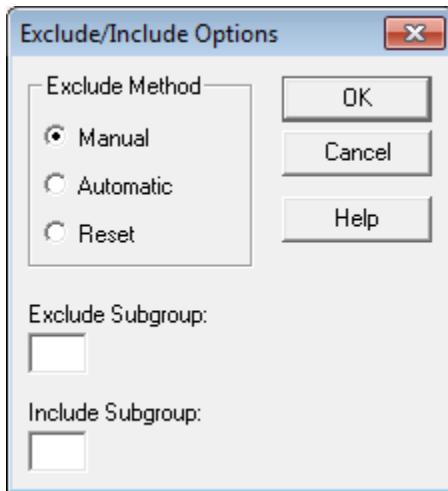
Analysis Options

- **Type of Study:** determines how the control limits are set. For an *Initial Study* (Phase 1) chart, the limits are estimated from the current data. For a *Control to Standard* (Phase 2) chart, the control limits are determined from the information in the *Control to Standard* section of the dialog box.
- **Normalize:** if selected, all statistics will be normalized by calculating Z-scores and the Z-scores plotted on the charts. The centerline on such a chart is always located at 0, and the control limits are always located at $\pm k$.
- **Avg. Subgroup Size:** if checked, the control limits will be horizontal lines based on the average sample size. If not checked, the individual sample sizes will be used, resulting in step function control limits if the sample sizes are not all equal.
- **Use Zone Format:** if checked, the statistics will be plotted using a zone chart rather than the usual format. See the example below.
- **Recalculate at:** the control limits of the chart may be recalculated at up to 4 locations along the X axis by specifying the sample numbers at which new estimates are to be initiated. Separate estimates of the mean proportion will be obtained using the data in the different sections. In such cases, the control limits will be adjusted at the start of each new period.
- **P' Chart Control Limits:** specifies the multiple k to use in determining the upper and lower control limits on the p' chart. To suppress a limit completely, enter 0.
- **Control to Standard:** to perform a Phase 2 analysis, select *Control to Standard* for the *Type of Study* and then enter either of the following:

Specify Parameters: specify the established standard mean proportion of defective items and the assumed value for σ_z .

Specify Control Limits: specify the location of the centerline and control limits exactly where you wish them to be placed.

- **Exclude button:** Use this button to exclude specific samples from the calculations. It will display the following dialog box:



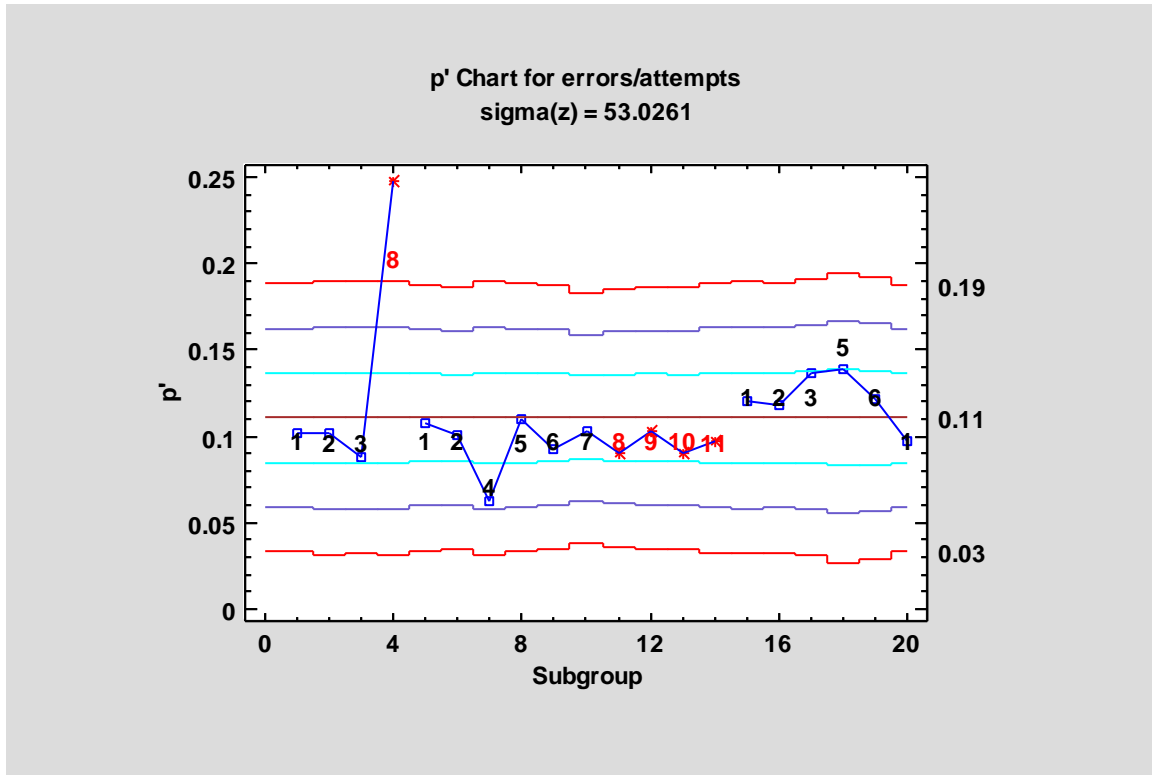
To exclude one or more values from the calculations:

1. Select *Manual*, enter the sample (subgroup) number to exclude, and press OK. The control limits will be recalculated without that subgroup and the chart replotted.
2. Select *Automatic* and press OK. The program will then remove points one at a time from the control chart, recalculating the limits after each point is removed. It will stop removing points once all remaining points are within the current control limits. This option should be used with care and is designed primarily for “what if” types of analyses.

Points may also be excluded by clicking on them while viewing a control chart and then pressing the *Exclude/Include* button on the analysis toolbar.

Example: Zone Chart

If the *Zone Format* option is selected, then the p' chart will be displayed in the following manner:



Above each observation is a running count based on the latest run of observations either above the centerline or below the centerline. Beginning at 0, a cumulative score is attached to each point on the chart in the following manner:

- If the point is within the inner warning limits, the score is incremented by 1.
- If the point is beyond the inner warning limits but within the outer warning limits, the score is incremented by 2.
- If the point is beyond the outer warning limits but within the control limits, the score is incremented by 4.
- If the point is beyond the control limits, the score is incremented by 8.

Whenever the process crosses the centerline, the score is reset to 0. Values exactly equal to the centerline do not change the score. When the score reaches a critical value (8 is the default), an out-of-control signal is generated.

The above rule is the default but may be overridden on the *Runs Tests* tab of the *Preferences* dialog box, accessible through the *Edit* menu.

Zone Chart Weights

Beyond 3 sigma:

Beyond 2 sigma:

Beyond 1 sigma:

Within 1 sigma:

Signal at:

Reset after signal

In addition, the score may be reset to 0 immediately after each out-of-control signal.

In the chart above, the maximum score reached 8 or more twice.

Runs Tests

The *Runs Tests* pane displays the results of standard tests applied to the p' chart to look for unusual sequences of points.

Runs Tests	
Rules	
(A) runs above or below centerline of length 8 or greater.	
(B) runs up or down of length 8 or greater.	
(C) sets of 5 subgroups with at least 4 beyond 1.0 sigma.	
(D) sets of 3 subgroups with at least 2 beyond 2.0 sigma.	
Violations	
<i>Observation</i>	<i>p' Chart</i>
12	A
13	A
14	A

Depending on the default settings on the *Runs Tests* tab of the *Preferences* dialog box, STATGRAPHICS will look for up to 7 different types of patterns:

- A. A group of 8 or more points, all above or all below the centerline.
- B. A group of 8 or more points, all increasing or all decreasing.
- C. A group of 5 points in which at least 4 are more than 1-sigma away from the centerline, on the same side of the centerline.
- D. A group of 3 points in which at least 2 are more than 2-sigma away from the centerline, on the same side of the centerline.
- E. A group of 15 or more points, all within 1-sigma.
- F. A group of 8 or more points, all beyond 2-sigma, but not necessarily on the same side of the centerline.
- G. A group of 8 or more points, all following an alternating up and down pattern.

Any such runs will be indicated in the above table and also on the control charts (unless suppressed).

Runs tests are designed to make standard Shewhart charts more sensitive to small shifts in the process. In the current example, rule A detects a run of 8 or more points below the centerline beginning at time period $j = 12$.

Pane Options

Runs Tests Options

- Runs of or more on same side of centerline
- Runs of or more increasing or decreasing
- Sets of with at least beyond sigma
- Sets of with at least beyond sigma
- Runs of or more at or within sigma
- Runs of or more beyond sigma (either side)
- Runs of or more alternating up and down

OK Cancel Help

Select the runs tests to be applied and the parameters that define those tests. For example, some practitioners prefer to test for runs of length 7 rather than 8.

Save Results

The following results can be saved to the datasheet:

1. *Proportions* – the sample proportions.
2. *Sizes* – the sample sizes.
3. *Labels* – the sample labels.
4. *Mean p* – the estimated mean proportion of unacceptable items.
5. *Sigma* – the estimated standard error of the proportions.

Calculations

$$\text{Estimate of process variability for sample } j: \quad \hat{\sigma}_{p_j} = \sqrt{\frac{\bar{p}(1-\bar{p})}{n_j}} \quad (7)$$

$$\text{Normalized value for sample } j: \quad z_j = \frac{p_j - \bar{p}}{\hat{\sigma}_{p_j}} \quad (8)$$

$$\text{Average moving range:} \quad \bar{R} = \frac{1}{m-1} \sum_{j=2}^m |z_j - z_{j-1}| \quad (9)$$

$$\text{Estimated relative amount of additional process variation:} \quad \hat{\sigma}_z = \frac{\bar{R}}{1.128} \quad (10)$$

References

Laney, David B. (2002) *Improved Control Charts for Attributes*. ASQ and Marcel Dekker Inc.

Hung, Chin-liang *Control Charts for Attributes: Some Variations*, www.iastate.edu.