Cusum Charts (Tabular)

Summary

Cumulative sum charts monitor variable data by plotting the cumulative sum of deviations from target. By so doing, they incorporate the past history of the process into the plotted points, achieving greater sensitivity and shorter run lengths than X-bar or X charts against small shifts in the process mean. STATGRAPHICS offers two procedures for creating cusum charts:

1. The *Cusum Charts (Tabular)* procedure described in this document, which plots two one-sided cumulative sums with horizontal control limits.
2. The *Cusum Charts (V-Mask)* procedure, which plots a single two-sided cumulative sum with a leading "V-Mask" to determine when the process goes out-of-control.

The *tabular* or *algorithmic* cusum described here is often preferred since it more closely resembles other control charts. In addition, a "headstart" feature may be added to further reduce average run lengths.

This procedure also create a "scale cusum" chart for individuals data to monitor process variability.

Sample StatFolio: *cusumtabular.sgp*

Sample Data:

The file *process shift.sgd* contains a sample of random numbers described by Montgomery (2005). The data consist of $m = 30$ observations, to be treated as individuals. A partial list of the data in that file is shown below:

Sample	*X*
1	9.45
2	7.99
3	9.29
4	11.66
5	12.16
6	10.18
7	8.04
8	11.46
9	9.20
10	10.34

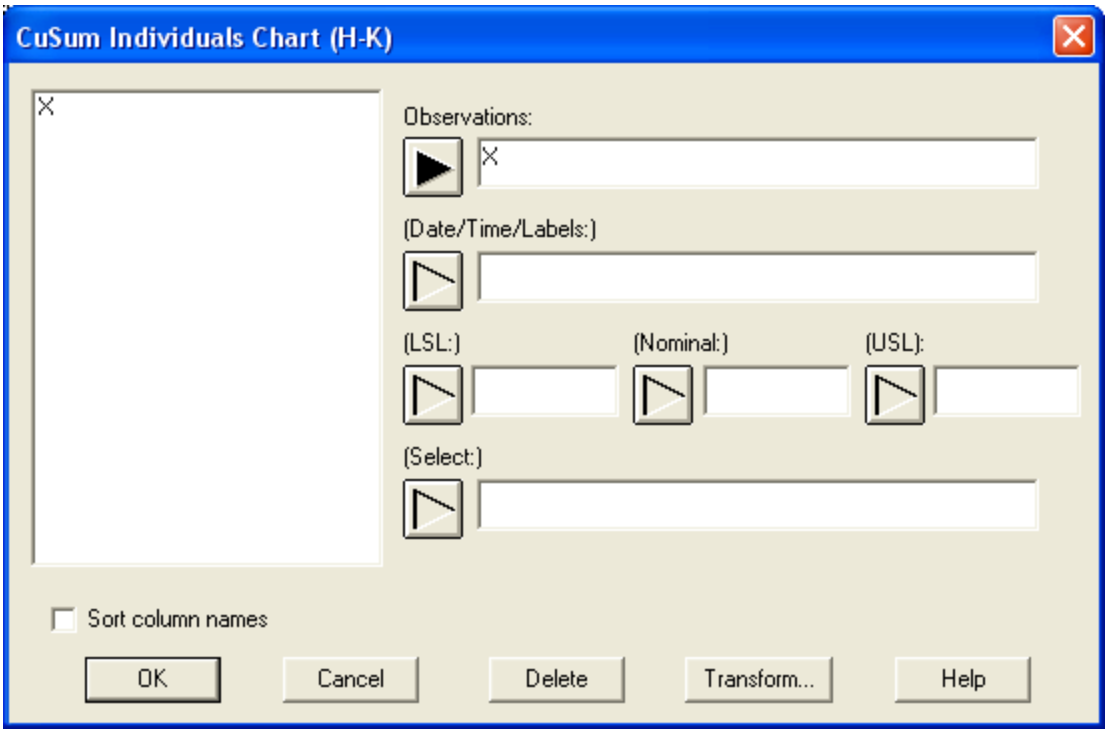
The first 20 observations were randomly generated from a normal distribution with $\mu = 10$ and $\sigma = 1$. The last 10 observations were randomly generated from a normal distribution with $\mu = 11$ and $\sigma = 1$, representing a 1-sigma shift in the process mean. Montgomery uses this example to illustrate the properties of various types of time-weighted charts.

Data Input

There are two menu selections that create tabular cusum charts, one for individuals data and one for grouped data. In the case of grouped data, the original observations may be entered, or subgroup statistics may be entered instead.

Case #1: Individuals

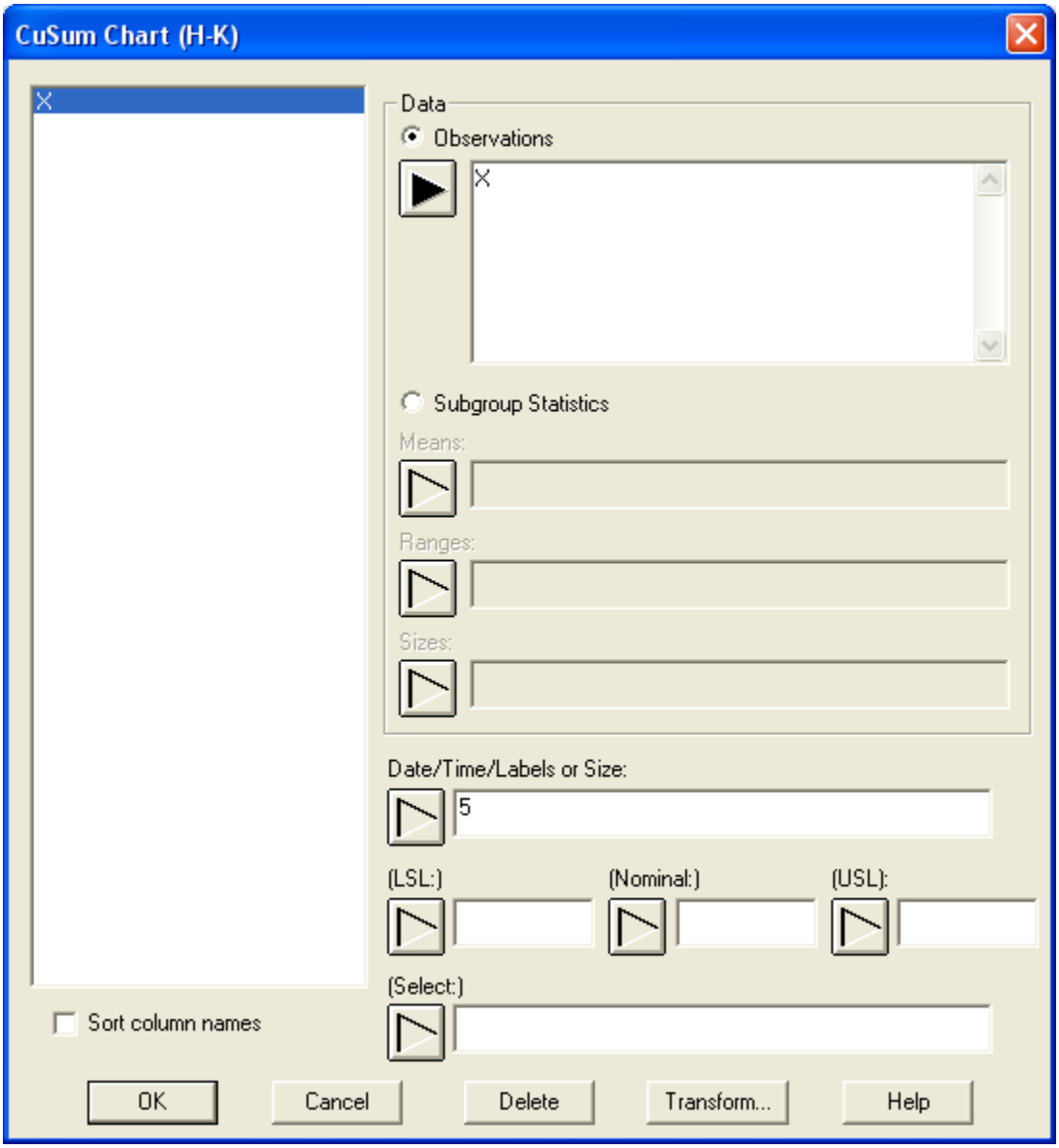
The data to be analyzed consist of a single numeric column containing *n* observations. The data are assumed to have been taken one at a time.



- **Observations:** numeric column containing the data to be analyzed.
- **Date/Time/Labels:** optional labels for each observation.
- **LSL, Nominal, USL**: optional lower specification limit, nominal (target) value, and upper specification limit.
- **Select:** subset selection.

Case #2: Grouped Data – Original Observations

The data to be analyzed consist of one or more numeric columns. The data are assumed to have been taken in groups, in sequential order by rows.



- **Observations:** one or more numeric columns. If more than one column is entered, each row of the file is assumed to represent a subgroup with subgroup size *m* equal to the number of columns entered. If only one column is entered, then the *Date/Time/Labels or Size* field is used to form the groups.

- **Date/Time/Labels or Size**: If each set of *m* rows represents a group, enter the single value *m*. For example, entering a 5 as in the example above implies that the data in rows 1-5 form the first group, rows 6-10 form the second group, and so on. If the subgroup sizes are not equal, enter the name of an additional numeric or non-numeric column containing group identifiers. The program will scan this column and place sequential rows with identical codes into the same group.

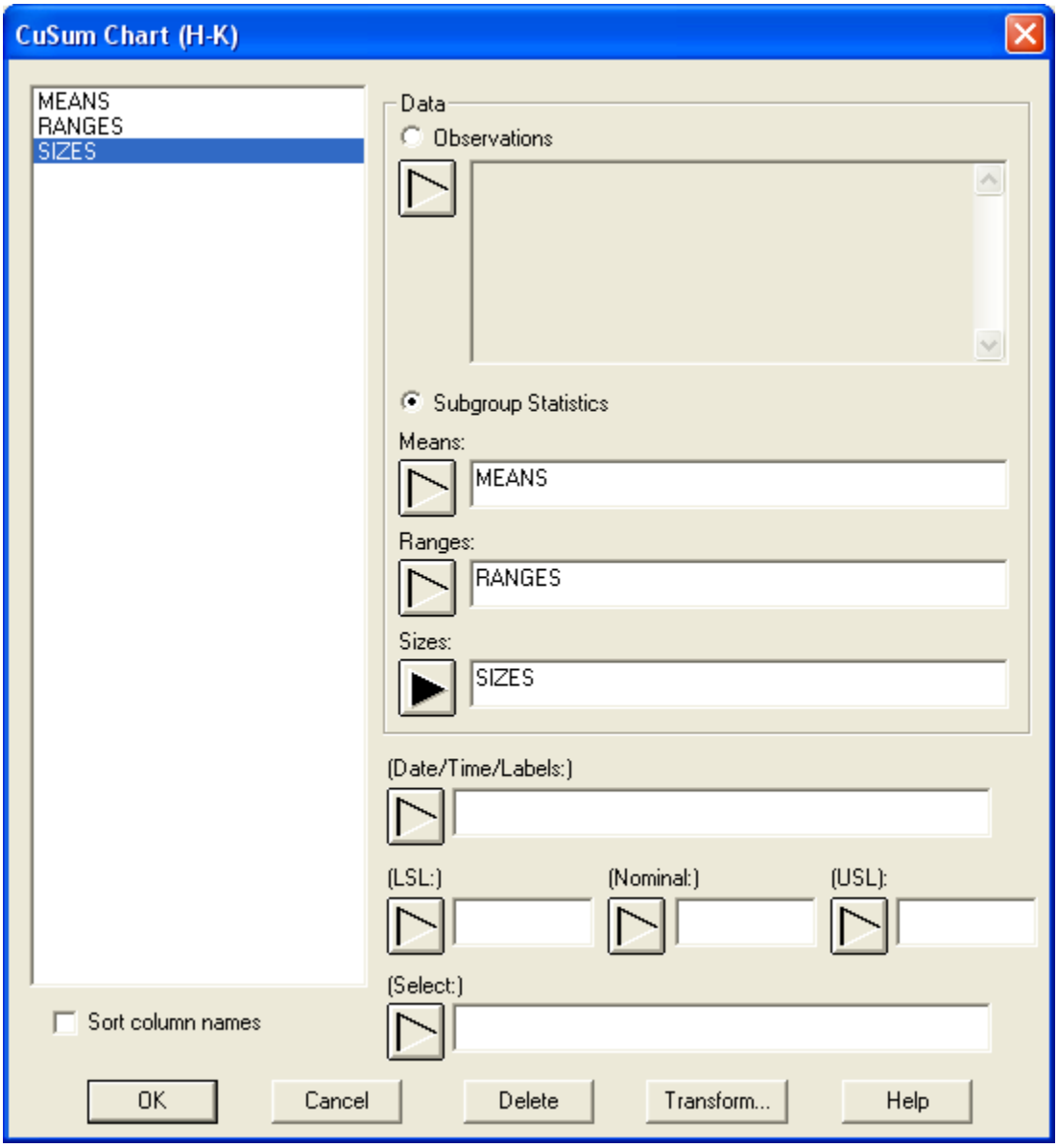
- **LSL, Nominal, USL**: optional lower specification limit, nominal (target) value, and upper specification limit.

- **Select:** subset selection.

Case #3: Grouped Data – Subgroup Statistics

In this case, the statistics for each subgroup have been computed elsewhere and entered into the datasheet, as in the table below:

Sample	*Means*	*Ranges*	*Sizes*
1	10.11	4.17	5
2	9.844	3.42	5
3	10.098	2.44	5
4	9.932	2.32	5
5	10.924	2.96	5



- **Subgroup Statistics**: the names of the column containing the subgroup means, subgroup ranges, and subgroup sizes.

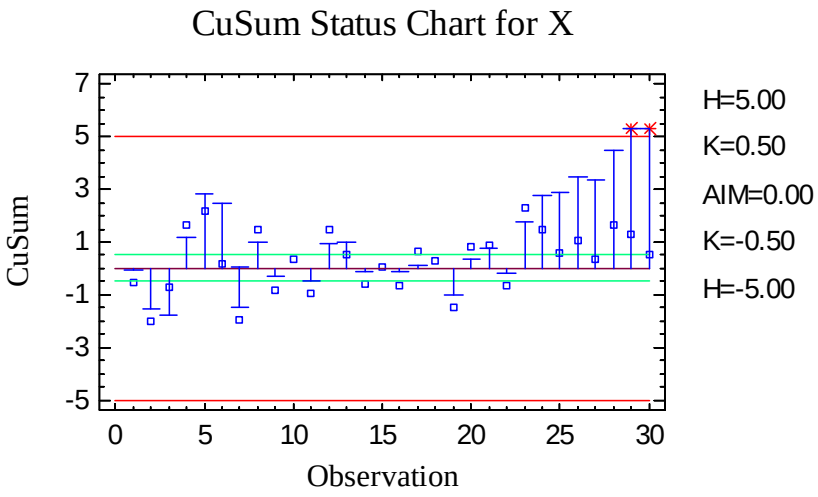
- **Date/Time/Labels:** optional labels for each subgroup.

- **LSL, Nominal, USL**: optional lower specification limit, nominal (target) value, and upper specification limit.

- **Select:** subset selection.

CuSum Chart

The tabular cusum chart plots the original data or subgroup means together with upper and lower cumulative sums.



The chart includes:

1. **Point symbols** displaying the observations or subgroup means.

2. A **centerline** at 0.

3. Horizontal lines at a short distance *K* above and below the centerline. K is called the **slack** or **reference value** and is specified on the *Analysis Options* dialog box as a multiple of the process sigma according to

$$K = k\sigma \qquad (1)$$

k affects the performance of the chart and is often set to a value close to 0.5. Points falling within this slack value do not add to the cumulative sums, as discussed below.

4. Horizontal lines at a larger distance *H* above and below the centerline. These lines define the **decision interval** and are also specified on the *Analysis Options* dialog box as a multiple of sigma:

$$H = h\sigma \qquad (2)$$

Cumulative sums falling outside the decision interval are treated as out-of-control signals.

5. Vertical bars extending above and below the centerline, representing the value of two one-sided cumulative sums, defined for individuals by:

$$C_j^+ = \max[0, x_j - (\mu + k\sigma) + C_{j-1}^+] \quad (3)$$

$$C_j^- = \max[0, (\mu - k\sigma) - x_j + C_{j-1}^-] \quad (4)$$

where $C_0^+ = C_0^- = 0$. These sums accumulate the deviations of the observations or subgroup means from the overall mean μ, plus and minus the slack. The two cusums are designed to detect shifts above and below the mean, respectively.

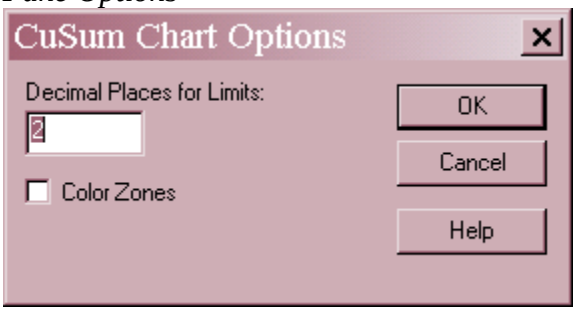
While the process is in control, as it appears to be for the first 22 periods, the cusums will bounce up and down around the centerline. However, if the process mean shifts away from the assumed mean μ, one or the other of the cusums will begin to grow and eventually move outside the decision interval, as is the case with the positive cusum in the plot above.

The cusum charts are most commonly used during Phase 2 to monitor a process against an established standard. In the current example, the parameters were set at μ = 10 and σ = 1, corresponding to the distribution from which the first 20 rows were generated. If used during Phase 1, the parameters are replaced by estimates from the current data. In that case,

$$\hat{\mu} = \bar{\bar{x}} = \frac{\sum_{j=1}^{m} n_j \bar{x}_j}{\sum_{j=1}^{m} n_j} \quad (5)$$

while σ is estimated using the method specified on the *Control Charts* tab of the *Preferences* dialog box.

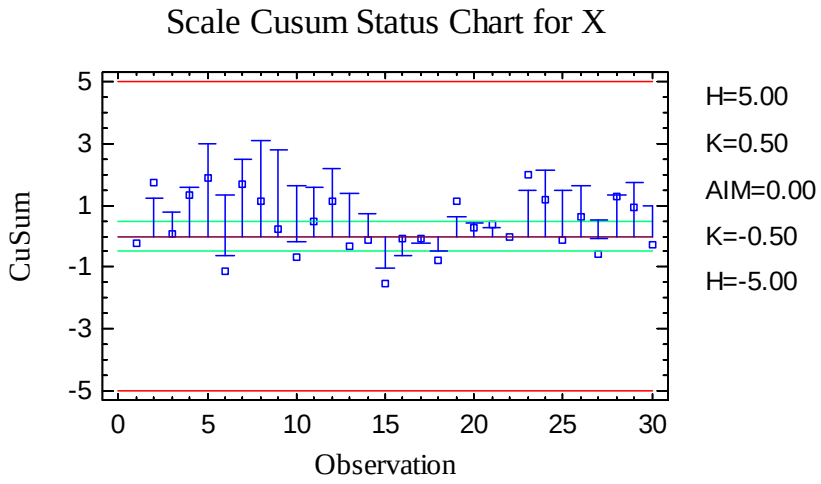
Pane Options



- **Decimal Places for Limits:** the number of decimal places used to display the control limits.
- **Color Zones:** check this box to display green, yellow and red zones.

Scale Chart

For individuals data, a second cusum chart can also be plotted to monitor the variability of the process.



This chart is designed to detect changes in the variability of the process by plotting a cumulative sum of the standardized quantities

$$v_j = \frac{\sqrt{\left|\frac{x_j - \mu}{\sigma}\right|} - 0.822}{0.349} \qquad (6)$$

It plots two one-sided cusums defined by

$$S_j^+ = \max[0, v_j - k + S_{j-1}^+] \qquad (7)$$

$$S_j^- = \max[0, -k - v_j + S_{j-1}^-] \qquad (8)$$

where $S_j^+ = S_j^- = 0$, with decision limits at $\pm h$. Any point outside the decision limits would indicate a change in the variability of the process. The scale cusum uses the same values of *h* and *k* as the cusum for deviations from the mean.

CuSum Chart Report

This pane tabulates the values plotted on the control charts:

CuSum Individuals Chart Report
All Observations
X = Excluded * = Beyond Limits

C = Mean Cusum S = Scale Cusum

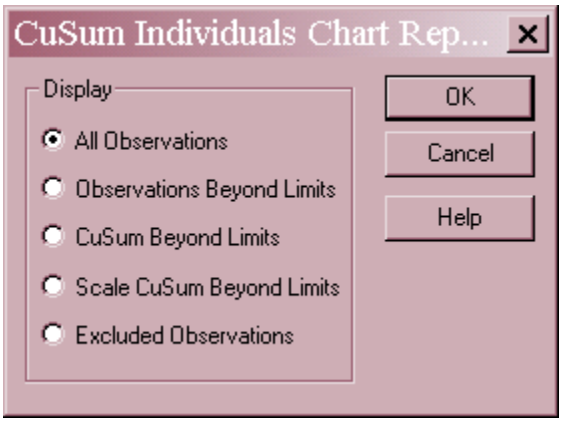
Observation	*C+*	*N+*	*C-*	*N-*	*S+*	*N+*	*S-*	*N-*
1	0.0	0	0.05	1	0.0	0	0.0	0
2	0.0	0	1.56	2	1.207	1	0.0	0
3	0.0	0	1.77	3	0.766074	2	0.0	0
4	1.16	1	0.0	0	1.60249	3	0.0	0
5	2.82	2	0.0	0	2.95835	4	0.0	0
6	2.5	3	0.0	0	1.3187	5	0.639645	1
7	0.04	4	1.46	1	2.47486	6	0.0	0
8	1.0	5	0.0	0	3.08175	7	0.0	0
9	0.0	0	0.3	1	2.78928	8	0.0	0
10	0.0	0	0.0	0	1.60474	9	0.184541	1

The columns in the table show the calculations for each cusum chart:

C+: the value of the positive cusum.
N+: the number of periods that the positive cusum has been above 0.
C-: the value of the negative cusum.
N-: the number of periods that the negative cusum has been above 0.

When the process is deemed to be out-of-control, the values of C and N can be used to help judge when the process shifted and how much to adjust it.

Pane Options



- **Display:** specify the observations or subgroups to display in the report.

Analysis Summary

The *Analysis Summary* summarizes the data and the control charts.

Cumulative Sum Individuals Chart (H-K) - X
Number of subgroups = 30
Subgroup size = 1.0
0 subgroups excluded

Decision interval h: 5.0
Slack value k: 0.5
Headstart: 0.0
Target: 10.0

Distribution: Normal
Transformation: none

Mean CuSum Chart

Period	*#1-30*
UCL	5.0
Centerline	0.0
LCL	-5.0

2 beyond limits

Scale CuSum Chart

Period	*#1-30*
UCL	5.0
Centerline	0.0
LCL	-5.0

0 beyond limits

Estimates

Period	*#1-30*	*Standard*
Process mean	10.315	10.0
Process sigma	1.19987	1.0
Mean MR(2)	1.35345	

Sigma estimated from average moving range

Included in the table are:

- **Subgroup Information**: the number of observations or subgroups *m* and the average subgroup size $\overline{n}$. If any observations or subgroups have been excluded from the calculations, that number is also displayed.

- **Decision interval, slack value, headstart, and target value:** the parameters used to calculate the cusums and position the control limits.

- **Distribution:** the assumed distribution for the data. By default, the data are assumed to follow a normal distribution. However, one of 26 other distributions may be selected using *Analysis Options*.

- **Transformation:** any transformation that has been applied to the data. Using *Analysis Options*, you may elect to transform the data using either a common transformation such as a square root or optimize the transformation using the Box-Cox method.

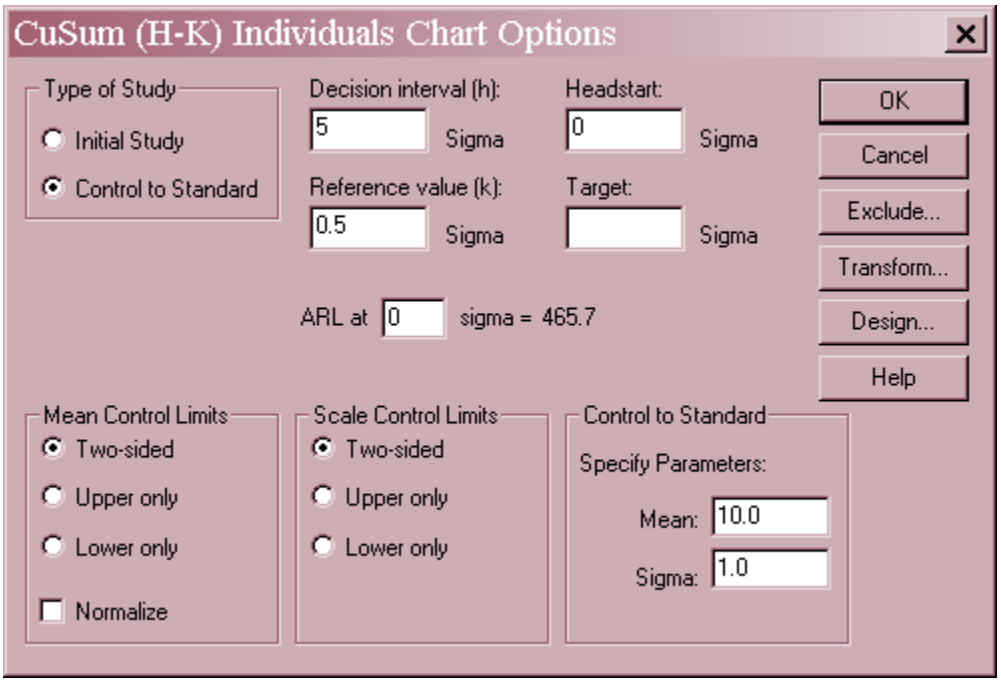
- **Mean CuSum Chart:** a summary of the centerline and control limits for the mean cusum chart.

- **Scale Cusum Chart:** a summary of the centerline and control limits for the scale cusum chart.

- **Estimates:** estimates of the process mean μ and the process standard deviation σ. The methods for estimating the process sigma depending upon the settings on the *Control Charts* tab of the *Preferences* dialog box, accessible through the *Edit* menu. The options are described in the *Individuals Control Charts* and *X-bar and R Charts* documentation.

- **Mean MR(2), Mean Range, or Mean S:** the value of the statistic used to calculate the process sigma.

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.

- **Decision interval (h):** sigma multiple used to position the control limits. A common value for *h* is 5.0.

- **Reference value (k):** sigma multiple used to position the reference or slack value. Variation within the slack limits does not count toward the cumulative sum. A common value for *k* is 0.5.

- **Headstart:** an alternative value for initializing the cusums C_0^+ and C_0^-. If specified, this gives the chart a *fast initial response,* reducing the ARL for processes that are not in control. A commonly used headstart value is *h*/2, which starts off the cusums at half the distance to the decision limits.

- **Target:** an alternative value for the target mean μ to be used when calculating the cusums. If left blank, the sample mean $\bar{x}$ will be used as the target in *Initial Studies* mode, or the specified process mean will be used if in *Control to Standard* mode.

- **Mean Control Limits**: either or both of the two one-sided cusums may be plotted.

- **Normalize**: creates a standardized cusum chart by first normalizing each data value according to

$$z_j = \frac{x_j - \mu}{\sigma} \tag{9}$$

  All of the calculations are then performed using these normalized values.

- **Scale Control Limits**: controls the scale cusum chart.

- **ARL:** the average run length at a selected number of standard deviations δ away from the mean. The dialog box displays the ARL assuming the process shifts from the original mean μ to a new mean ($\mu + \delta\sigma$).

- **Control to Standard:** To perform a Phase 2 analysis, select *Control to Standard* for the *Type of Study* and then enter the established standard process mean and sigma (or other parameters if not assuming a normal distribution).

- **Exclude button:** Use this button to exclude specific subgroups from the calculations.

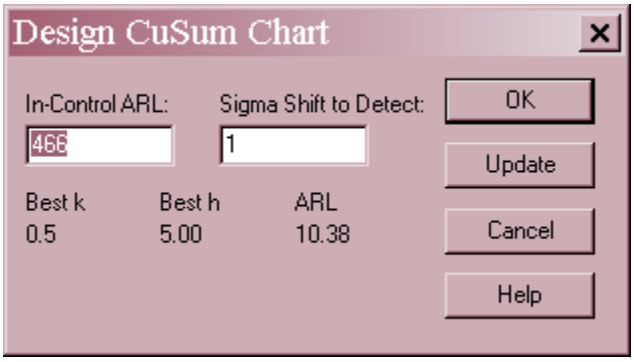
- **Transform Button**: Use this button to specify a transformation or non-normal distribution.

- **Design Button:** Use this button to select values for *h* and *k* based on the ARL at a target shift in the process mean.

For a discussion of the *Exclude* and *Transform* features, see the documentation for *Individuals Control Charts*.

Designing Cusum Charts

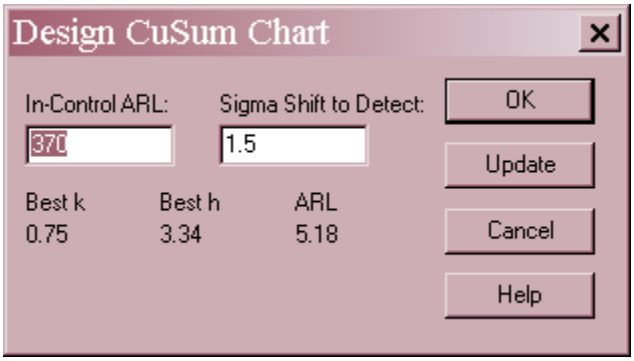
The *Analysis Options* dialog box contains a button labeled *Design*. This button allows you to design a cusum chart by helping select optimal values of *h* and *k*. When pressed, the button displays the following dialog box:



There are two input fields:

- **In-Control ARL** - the desired average run length when the process is in control. By default, the ARL for the current chart is displayed.
- **Sigma Shift to Detect** - the magnitude shift in the mean δ that the chart should be designed to detect.

The dialog box calculates the optimal *h* and *k* in order to minimize the ARL at the desired sigma shift. For example, setting the *In-Control ARL* to 370 (the value of a 3-sigma Shewhart chart), the *Sigma Shift to Detect* to 1.5, and pressing *Update* yields:



The optimal values are $k = 0.75$ and $h = 3.34$, which gives an ARL = 5.18 against a 1.5 sigma shift in the mean. When the OK button is pressed, the new *h* and *k* values will be entered into the *Analysis Options* dialog box.

Capability Indices

The *Capability Indices* pane displays the values of selected indices that measure how well the data conform to the specification limits.

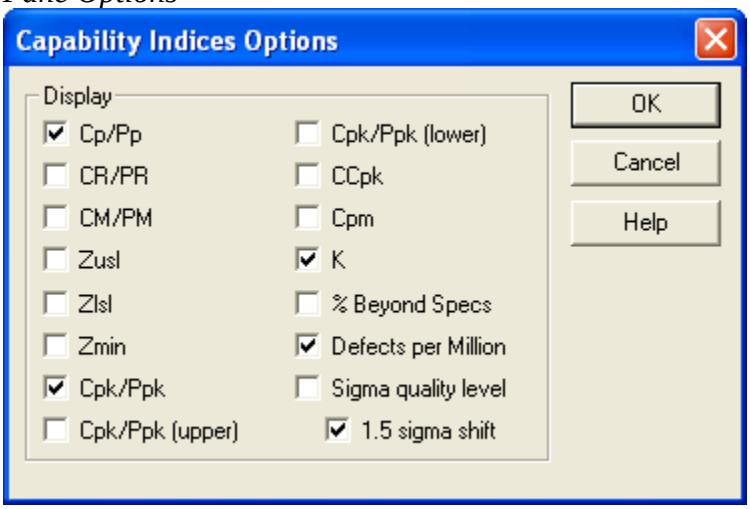
Capability Indices for X
Specifications
USL = 15.0
Nom = 10.0
LSL = 5.0

	Short-Term Capability	*Long-Term Performance*
Sigma	1.19987	1.15354
Cp/Pp	1.38904	1.44483
CR/PR	71.9919	69.2122
Cpk/Ppk	1.30153	1.35381
K		0.063

Based on 6 sigma limits. Short-term sigma estimated from average moving range.

The indices displayed by default depend on the settings of the *Capability* tab on the *Preferences* dialog box. A detailed discussion of these indices may be found in the documentation for *Process Capability (Variables).*

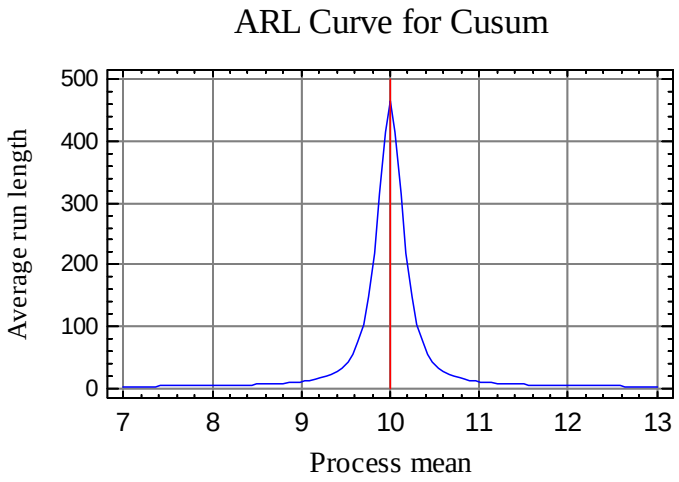
Pane Options



- **Display**: select the indices to be displayed.

ARL Curve

The *ARL Curve* is another way to view the performance of a Phase 2 chart.



The ARL curve plots the average run length (average number of values plotted up to and including the first point beyond the control limits) as a function of the true process mean. Assuming that the process mean suddenly shifts to a new value, the plot shows how long it takes on average until an out-of-control signal is generated. For very small shifts, it can take in excess of 400 points on average to detect the shift. At a shift to $\mu = 11$, the ARL is approximately 10.

Save Results

The following results can be saved to the datasheet, depending on whether the data are individuals or grouped:

1. *Cumulative sum* – the larger of the two cumulative sums at each time period.
2. *Ranges, sigmas, or moving ranges* – the values used to estimate the process sigma.
3. *Sizes* – the subgroup sizes.
4. *Labels* – the subgroup labels.
5. *Process Mean* – the estimated process mean.
6. *Process Sigma* – the estimated process standard deviation.
7. *Included Observations* – a column of 0's and 1's for excluded and included observations, respectively. This column can then be used in the *Select* field on other data input dialog boxes.

Calculations

ARL

The ARL of a two-sided cusum chart is calculated from the ARLs of the two one-sided cusums according to

$$\frac{1}{ARL} = \frac{1}{ARL_+} + \frac{1}{ARL_-} \tag{10}$$

where for $\Delta \neq 0$

$$ARL_+ = \frac{\exp(-2\Delta b) + 2\Delta b - 1}{2\Delta^2}, \quad \Delta = \delta - k \tag{11}$$

$$ARL_- = \frac{\exp(-2\Delta b) + 2\Delta b - 1}{2\Delta^2}, \quad \Delta = -\delta - k \tag{12}$$

and $b = h + 1.166$. If $\Delta = 0$, then ARL_+ *or* $ARL_- = b^2$.

Grouped Data

The formulas in this document assume that the data to be analyzed are individual measurements. If analyzing grouped data, substitute $\bar{x}_j$ in place of x_j and $\sigma / \sqrt{n}$ in place of σ.

