# Chapter 12

## XCHART Statement

## Chapter Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>OVERVIEW</td>
<td>395</td>
</tr>
<tr>
<td>GETTING STARTED</td>
<td>396</td>
</tr>
<tr>
<td>Creating a V-Mask Cusum Chart from Raw Data</td>
<td>396</td>
</tr>
<tr>
<td>Creating a V-Mask Cusum Chart from Subgroup Summary Data</td>
<td>399</td>
</tr>
<tr>
<td>Saving Summary Statistics</td>
<td>401</td>
</tr>
<tr>
<td>Creating a One-Sided Cusum Chart with a Decision Interval</td>
<td>402</td>
</tr>
<tr>
<td>Saving Cusum Scheme Parameters</td>
<td>405</td>
</tr>
<tr>
<td>Reading Cusum Scheme Parameters</td>
<td>408</td>
</tr>
<tr>
<td>SYNTAX</td>
<td>410</td>
</tr>
<tr>
<td>Summary of Options</td>
<td>411</td>
</tr>
<tr>
<td>Dictionary of Special Options</td>
<td>419</td>
</tr>
<tr>
<td>DETAILS</td>
<td>425</td>
</tr>
<tr>
<td>Basic Notation for Cusum Charts</td>
<td>425</td>
</tr>
<tr>
<td>Formulas for Cumulative Sums</td>
<td>426</td>
</tr>
<tr>
<td>Defining the Decision Interval for a One-Sided Cusum Scheme</td>
<td>428</td>
</tr>
<tr>
<td>Defining the V-Mask for a Two-Sided Cusum Scheme</td>
<td>429</td>
</tr>
<tr>
<td>Designing a Cusum Scheme</td>
<td>431</td>
</tr>
<tr>
<td>Cusum Charts Compared with Shewhart Charts</td>
<td>435</td>
</tr>
<tr>
<td>Methods for Estimating the Standard Deviation</td>
<td>435</td>
</tr>
<tr>
<td>Output Data Sets</td>
<td>438</td>
</tr>
<tr>
<td>ODS Tables</td>
<td>440</td>
</tr>
<tr>
<td>Input Data Sets</td>
<td>440</td>
</tr>
<tr>
<td>Missing Values</td>
<td>443</td>
</tr>
<tr>
<td>EXAMPLES</td>
<td>444</td>
</tr>
<tr>
<td>Example 12.1 Cusum and Standard Deviation Charts</td>
<td>444</td>
</tr>
<tr>
<td>Example 12.2 Upper and Lower One-Sided Cusum Charts</td>
<td>447</td>
</tr>
<tr>
<td>Example 12.3 Combined Shewhart–Cusum Scheme</td>
<td>448</td>
</tr>
</tbody>
</table>
Chapter 12
XCHART Statement

Overview

The XCHART statement creates cumulative sum control charts from subgroup means or individual measurements. You can create these charts for one-sided cusum (decision interval) schemes or for two-sided (V-mask) schemes. A one-sided scheme is designed to detect either a positive or a negative shift from the target mean, and a two-sided scheme is designed to detect positive and negative shifts from the target mean.

You can use options in the XCHART statement to

- specify parameters for a decision interval or V-mask
- specify the shift $\delta$ to be detected
- specify the target mean $\mu_0$
- specify a known (standard) value $\sigma_0$ for the process standard deviation or estimate the standard deviation from the data using various methods
- tabulate the information displayed on the chart
- save the information displayed on the chart in an output data set
- read parameters for the cusum scheme from a data set
- display a secondary chart that plots a time trend that has been removed from the data
- add block legends and special symbol markers to reveal stratification in process data
- superimpose stars at each point to represent related multivariate factors
- display vertical and horizontal reference lines
- modify the axis values and labels
- modify the chart layout and appearance
Getting Started

This section introduces the XCHART statement with simple examples that illustrate the most commonly used options. Complete syntax for the XCHART statement is presented in the “Syntax” section on page 410, and advanced examples are given in the “Examples” section on page 444.

Creating a V-Mask Cusum Chart from Raw Data

A machine fills eight-ounce cans of two-cycle engine oil additive. The filling process is believed to be in statistical control, and the process is set so that the average weight of a filled can is $\mu_0 = 8.100$ ounces. Previous analysis shows that the standard deviation of fill weights is $\sigma_0 = 0.050$ ounces. A two-sided cusum chart is used to detect shifts of at least one standard deviation in either the positive or negative direction from the target mean of 8.100 ounces.

Subgroup samples of four cans are selected every hour for twelve hours. The cans are weighed, and their weights are saved in a SAS data set named OIL.

```sas
data oil;
  label hour = 'Hour';
  input hour @;
  do i=1 to 4;
    input weight @;
    output;
  end;
  drop i;
cards;
1  8.024  8.135  8.151  8.065
2  7.971  8.165  8.077  8.157
3  8.125  8.031  8.198  8.050
4  8.123  8.107  8.154  8.095
5  8.068  8.093  8.116  8.128
6  8.177  8.011  8.102  8.030
7  8.129  8.060  8.125  8.144
8  8.072  8.010  8.097  8.153
9  8.066  8.067  8.055  8.059
10 8.089  8.064  8.170  8.086
11 8.058  8.098  8.114  8.156
12 8.147  8.116  8.116  8.018
;
proc print noobs;
  run;
```

The data set OIL is partially listed in Figure 12.1.
Each observation contains one value of WEIGHT along with its associated value of HOUR, and the values of HOUR are in increasing order. The CUSUM procedure assumes that DATA= input data sets are sorted in this “strung-out” form.

The following statements request a two-sided cusum chart with a V-mask for the average weights:

```sas
symbol v=dot;
title 'Cusum Chart for Average Weights of Cans';
proc cusum data=oil;
   xchart weight*hour /
      mu0 = 8.100 /* target mean */
      sigma0 = 0.050 /* known standard deviation */
      delta = 1 /* shift to be detected */
      alpha = 0.10 /* Type 1 error probability */
      vaxis = -5 to 3 ;
   label weight='Cumulative Sum';
run;
```

The CUSUM procedure is invoked with the PROC CUSUM statement. The DATA= option in the PROC CUSUM statement specifies that the SAS data set OIL is to be read. The variables to be analyzed are specified in the XCHART statement. The process measurement variable (WEIGHT) is specified before the asterisk (this variable is referred to more generally as a process). The time variable (HOUR) is specified after the asterisk (this variable is referred to more generally as a subgroup-variable because it determines how the measurements are classified into rational subgroups).

The option ALPHA=0.10 specifies the probability of a Type 1 error for the cusum scheme (the probability of detecting a shift when none occurs).

The cusum chart is shown in Figure 12.2.
Part 2. The CAPABILITY Procedure

Figure 12.2. Two-Sided Cusum Chart with V-Mask

The cusum $S_1$ plotted at HOUR=1 is simply the standardized deviation of the first subgroup mean from the target mean.

$$S_1 = \frac{8.09375 - 8.100}{0.050/\sqrt{4}} = -0.250$$

The cusum $S_2$ plotted at HOUR=2 is $S_1$ plus the standardized deviation of the second subgroup mean from the target mean.

$$S_2 = S_1 + \frac{8.0925 - 8.100}{0.050/\sqrt{4}} = -0.550$$

In general, the cusum plotted at HOUR=$t$ is $S_{t-1}$ plus the standardized deviation of the $t$th subgroup mean from the target mean.

$$S_t = S_{t-1} + \frac{\bar{X}_t - \mu_0}{\sigma_0/\sqrt{n}}$$

For further details, see “Two-Sided Cusum Schemes” on page 427.

You can interpret the chart by comparing the points with the V-mask whose right edge is centered at the most recent point (HOUR=12). Since none of the points cross the arms of the V-mask, there is no evidence that a shift has occurred, and the fluctuations in the cusums can be attributed to chance variation. In general, crossing the lower arm is evidence of an increase in the process mean, whereas crossing the upper arm is evidence of a decrease in the mean.
Chapter 12. Getting Started

Creating a V-Mask Cusum Chart from Subgroup Summary Data

The previous example illustrates how you can create a cusum chart using raw process measurements read from a DATA= data set. In many applications, however, the data are provided in summarized form as subgroup means. This example illustrates the use of the XCHART statement when the input data set is a HISTORY= data set.

The following data set provides the subgroup means, standard deviations, and sample sizes corresponding to the variable WEIGHT in the data set OIL (see page 396):

```sas
data oilstat;
  input hour weightx weights weightn;
cards;
1  8.0938  0.0596  4
2  8.0925  0.0902  4
3  8.1010  0.0763  4
4  8.1198  0.0256  4
5  8.1013  0.0265  4
6  8.0800  0.0756  4
7  8.1145  0.0372  4
8  8.0830  0.0593  4
9  8.0618  0.0057  4
10 8.1023  0.0465  4
11 8.1065  0.0405  4
12 8.0993  0.0561  4
;
proc print;
run;
```

The data set OILSTAT is listed in Figure 12.3.

Since the data set contains a subgroup variable, a mean variable, a standard deviation variable, and a sample size variable, it can be read as a HISTORY= data set. Note that the names WEIGHTX, WEIGHTS, and WEIGHTN satisfy the naming conventions for summary variables since they begin with a common prefix (WEIGHT) and end with the suffix letters \(X\), \(S\), and \(N\).

The following statements create the cusum chart:

```sas
```
title 'Cusum Chart for Average Weights of Cans';

proc cusum history=oilstat;
  xchart weight*hour /
    mu0 = 8.100 /* target mean */
    sigma0 = 0.050 /* known standard deviation */
    delta = 1 /* shift to be detected */
    alpha = 0.10 /* Type 1 error probability */
    vaxis = -5 to 3;
  label weightx = 'Cumulative Sum';
run;

Note that the process WEIGHT specified in the XCHART statement is the prefix of the summary variable names in PARTS. Also note that the vertical axis label is specified by associating a variable label with the subgroup mean variable (PART-GAPX). The horizontal axis label is specified by associating a variable label with the subgroup-variable (HOUR). The chart (not shown here) is identical to the one in Figure 12.2.

In general, a HISTORY= input data set used with the XRCHART statement must contain the following four variables:

- subgroup variable
- subgroup mean variable
- subgroup range variable
- subgroup sample size variable

Furthermore, the names of subgroup mean, standard deviation, and sample size variables must begin with the prefix process specified in the XRCHART statement and end with the special suffix characters X, S, and N, respectively.

Note that the interpretation of process depends on the input data set specified in the PROC CUSUM statement.

- If raw data are read using the DATA= option (as in the previous example), process is the name of the SAS variable containing the process measurements.
- If summary data are read using the HISTORY= option (as in this example), process is the common prefix for the names containing the summary statistics.

For more information, see “DATA= Data Set” on page 440 and “HISTORY= Data Set” on page 442.
Saving Summary Statistics

In this example, the CUSUM procedure is used to save summary statistics and cusums in an output data set. The summary statistics can subsequently be analyzed by the CUSUM procedure (as in the preceding example). The following statements read the raw measurements from the data set OIL (see page 396) and create a summary data set named OILHIST:

```
proc cusum data=oil;
  xchart weight*hour / nochart
    outhistory = oilhist            /* target mean */
    mu0 = 8.100 /* known standard deviation */
    sigma0 = 0.050 /* shift to be detected */
    delta = 1 /* Type 1 error probability */
    alpha = 0.10 /* Type 1 error probability */
    vaxis = -5 to 3 ;
  label weight = 'Cumulative Sum';
run;
```

```
proc print data=oilhist;
  format weightx weights weightc 6.4 ;
run;
```

The OUTHISTORY= option names the SAS data set containing the summary information, and the NOCHART option suppresses the display of the charts (since the purpose here is simply to create an output data set). Figure 12.4 lists the data set OILHIST.

<table>
<thead>
<tr>
<th>Obs</th>
<th>hour</th>
<th>weight</th>
<th>weight</th>
<th>weight</th>
<th>weight</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>X</td>
<td>S</td>
<td>C</td>
<td>N</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>8.0938</td>
<td>0.0596</td>
<td>-.2500</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>8.0925</td>
<td>0.0902</td>
<td>-.5500</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>8.1010</td>
<td>0.0763</td>
<td>-.5100</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>8.1198</td>
<td>0.0256</td>
<td>0.2800</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>8.1013</td>
<td>0.0265</td>
<td>0.3300</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>8.0800</td>
<td>0.0756</td>
<td>-.4700</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>8.1145</td>
<td>0.0372</td>
<td>0.1100</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>8.0830</td>
<td>0.0593</td>
<td>-.5700</td>
<td>4</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>8.0618</td>
<td>0.0057</td>
<td>-2.100</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>8.1023</td>
<td>0.0465</td>
<td>-2.010</td>
<td>4</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
<td>8.1065</td>
<td>0.0405</td>
<td>-1.750</td>
<td>4</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td>8.0993</td>
<td>0.0561</td>
<td>-1.780</td>
<td>4</td>
</tr>
</tbody>
</table>

Figure 12.4.  Listing of the Data Set OILHIST

There are five variables in the data set.

- HOUR contains the subgroup index
- WEIGHTX contains the subgroup means
- WEIGHTS contains the subgroup standard deviations
- WEIGHTC contains the cumulative sums
- WEIGHTN contains the subgroup sample sizes
Note that the variables in the OUTHISTORY= data set are named by adding the suffix characters \( X, S, N, \) and \( C \) to the process \texttt{WEIGHT} specified in the \texttt{XCHART} statement. In other words, the variable naming convention for OUTHISTORY= data sets is the same as for HISTORY= data sets.

For more information, see “OUTHISTORY= Data Set” on page 439.

### Creating a One-Sided Cusum Chart with a Decision Interval

An alternative to the V-mask cusum chart is the one-sided cusum chart with a decision interval, which is sometimes referred to as the “computational form of the cusum chart.” This example illustrates how you can create a one-sided cusum chart for individual measurements.

A can of oil is selected every hour for fifteen hours. The cans are weighed, and their weights are saved in a SAS data set named CANS:

```sas
proc print;
  var label: hour comment:
```

<table>
<thead>
<tr>
<th>hour</th>
<th>weight</th>
<th>comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8.024</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>7.971</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>8.125</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>8.123</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>8.068</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>8.177</td>
<td>Pump Adjusted</td>
</tr>
<tr>
<td>7</td>
<td>8.229</td>
<td>Pump Adjusted</td>
</tr>
<tr>
<td>8</td>
<td>8.072</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>8.066</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>8.089</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>8.058</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>8.147</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>8.141</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>8.047</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>8.125</td>
<td></td>
</tr>
</tbody>
</table>
```

Suppose the problem is to detect a positive shift in the process mean of one standard deviation \( \delta = 1 \) from the target of 8.100 ounces. Furthermore, suppose that

- a known value \( \sigma_0 = 0.050 \) is available for the process standard deviation
- an in-control average run length (ARL) of approximately 100 is required
- an ARL of approximately five is appropriate for detecting the shift

Table 12.17 on page 433 indicates that these ARLs can be achieved with the decision interval \( h = 3 \) and the reference value \( k = 0.5 \). The following statements use these parameters to create the chart and tabulate the cusum scheme:

```sas
* This data set is used by later examples in this chapter.
```
title 'One-Sided Cusum Analysis';
proc cusum data=cans;
   xchart weight*hour /
      mu0 = 8.100 /* target mean for process */
      sigma0 = 0.050 /* known standard deviation */
      delta = 1 /* shift to be detected */
      h = 3 /* cusum parameter h */
      k = 0.5 /* cusum parameter k */
      scheme = onesided /* one-sided decision interval */
      tableall; /* table */
   label weight = 'Cusum of Weight';
run;

The chart is shown in Figure 12.5.

Figure 12.5. One-Sided Cusum Chart with Decision Interval

The cusum plotted at HOUR=t is

$$S_t = \max(0, S_{t-1} + (z_t - k))$$

where \(S_0 = 0\), and \(z_t\) is the standardized deviation of the \(t\)\(^{th}\) measurement from the target.

$$z_t = \frac{x_t - \mu_0}{\sigma_0}$$

The cusum \(S_t\) is referred to as an upper cumulative sum. A shift is signaled at the seventh hour since \(S_7\) exceeds \(h\). For further details, see “One-Sided Cusum Schemes” on page 426.
The option TABLEALL requests the tables shown in Figure 12.6, Figure 12.7, and Figure 12.8. The table in Figure 12.6 summarizes the cusum scheme, and it confirms that an in-control ARL of 117.6 and an ARL of 6.4 at $\delta = 1$ are achieved with the specified $h$ and $k$.

### One-Sided Cusum Analysis

#### The CUSUM Procedure

**Cusum Parameters**

- **Process Variable**: weight (Cusum of Weight)
- **Subgroup Variable**: hour (Hour)
- **Scheme**: One-Sided
- **Target Mean (Mu0)**: 8.1
- **Sigma0**: 0.05
- **Delta**: 1
- **Nominal Sample Size**: 1
- **h**: 3
- **k**: 0.5
- **Average Run Length (Delta)**: 6.40390895
- **Average Run Length (0)**: 117.595692

### Figure 12.6.  Summary Table

<table>
<thead>
<tr>
<th>hour</th>
<th>Process Variable</th>
<th>Individual Value</th>
<th>Cusum Interval Exceeded</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8.0240000</td>
<td>0.0000000</td>
<td>3.0000</td>
</tr>
<tr>
<td>2</td>
<td>7.9710000</td>
<td>0.0000000</td>
<td>3.0000</td>
</tr>
<tr>
<td>3</td>
<td>8.1250000</td>
<td>0.0000000</td>
<td>3.0000</td>
</tr>
<tr>
<td>4</td>
<td>8.1230000</td>
<td>0.0000000</td>
<td>3.0000</td>
</tr>
<tr>
<td>5</td>
<td>8.0680000</td>
<td>0.0000000</td>
<td>3.0000</td>
</tr>
<tr>
<td>6</td>
<td>8.1770000</td>
<td>1.0400000</td>
<td>3.0000</td>
</tr>
<tr>
<td>7</td>
<td>8.2290000</td>
<td>3.1200000</td>
<td>3.0000</td>
</tr>
<tr>
<td>8</td>
<td>8.0720000</td>
<td>2.0600000</td>
<td>3.0000</td>
</tr>
<tr>
<td>9</td>
<td>8.0660000</td>
<td>0.8800000</td>
<td>3.0000</td>
</tr>
<tr>
<td>10</td>
<td>8.0890000</td>
<td>0.1600000</td>
<td>3.0000</td>
</tr>
<tr>
<td>11</td>
<td>8.0580000</td>
<td>0.0000000</td>
<td>3.0000</td>
</tr>
<tr>
<td>12</td>
<td>8.1470000</td>
<td>0.4400000</td>
<td>3.0000</td>
</tr>
<tr>
<td>13</td>
<td>8.1410000</td>
<td>0.7600000</td>
<td>3.0000</td>
</tr>
<tr>
<td>14</td>
<td>8.0470000</td>
<td>0.0000000</td>
<td>3.0000</td>
</tr>
<tr>
<td>15</td>
<td>8.1250000</td>
<td>0.0000000</td>
<td>3.0000</td>
</tr>
</tbody>
</table>

### Figure 12.7.  Tabulation of One-Sided Chart

The table in Figure 12.7 tabulates the information displayed in Figure 12.5.

### Figure 12.8.  Computational Form of Cusum Scheme

The table in Figure 12.8 presents the computational form of the cusum scheme described by Lucas (1976).
The CUSUM Procedure

Computational Cumulative Sum for weight

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>Sample</th>
<th>Individual</th>
<th>Upper</th>
<th>Consecutive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Value</td>
<td>Cusum</td>
<td>Number of</td>
</tr>
<tr>
<td></td>
<td>hour</td>
<td>Size</td>
<td></td>
<td>Upper Sums &gt; 0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>8.0240000</td>
<td>0.0000000</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>7.9710000</td>
<td>0.0000000</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>8.1250000</td>
<td>0.0000000</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>8.1230000</td>
<td>0.0000000</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>8.0680000</td>
<td>0.0000000</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>8.1770000</td>
<td>1.0400000</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>8.2290000</td>
<td>3.1200000</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>8.0720000</td>
<td>2.0600000</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>8.0660000</td>
<td>0.8800000</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>8.0890000</td>
<td>0.1600000</td>
<td>5</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>8.0580000</td>
<td>0.0000000</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>8.1470000</td>
<td>0.4400000</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>8.1410000</td>
<td>0.7600000</td>
<td>2</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
<td>8.0470000</td>
<td>0.0000000</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>8.1250000</td>
<td>0.0000000</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 12.8. Computational Form of Cusum Scheme

Following the method of Lucas (1976), the process average at the out-of-control point (HOUR=7) can be estimated as

\[
\mu_0 + \sigma_0 (N_7 k + S_7) / (N_7 \sqrt{n}) \\
= 8.10 + 0.05(2(0.5) + 3.12) / 2 \\
= 8.203 \text{ ounces}
\]

where \( S_7 = 3.12 \) is the upper sum at HOUR=7, and \( N_7 = 2 \) is the number of successive positive upper sums at HOUR=7.

Saving Cusum Scheme Parameters

This example is a continuation of the previous example that illustrates how to save cusum scheme parameters in a data set specified with the OUTLIMITS= option. This enables you to apply the parameters to future data or to subsequently modify the parameters with a DATA step program.

```
proc cusum data=cans lineprinter;
   xchart weight*hour /
      mu0 = 8.100 /* target mean */
      sigma0 = 0.050 /* known standard deviation */
      delta = 1 /* shift to be detected */
      h = 3 /* decision interval */
      k = 0.5 /* reference value */
      scheme = onesided /* one-sided scheme */
      outlimits = cusparm;
      label weight = 'Cusum of Weight';
proc print data=cusparm;
run;
```
The chart, shown in Figure 12.9, is similar to the one in Figure 12.5 but is created for output on a line printer since the LINEPRINTER option is included in the PROC CUSUM statement.

The OUTLIMITS= data set is listed in Figure 12.10. The data set contains one observation with the parameters for process WEIGHT. The variables _TYPE_, _H_, _K_, _MU0_, _DELTA_, and _STDDEV_ save the parameters specified with the options SCHEME=, H=, K=, MU0=, DELTA=, and SIGMA0=, respectively. The variable _MEAN_ saves an estimate of the process mean, and the variable _LIMITN_ saves the nominal sample size. The variables _ARLIN_ and _ARLOUT_ save the average run lengths for δ = 0 and for δ = 1.

The variables _VAR_ and _SUBGRP_ save the process and subgroup-variable. The variable _TYPE_ is a bookkeeping variable that indicates whether the value

*In Release 6.12 and previous releases of SAS/QC software, the keyword GRAPHICS was required in the PROC CUSUM statement to specify that the chart be created with a graphics device. In Version 7, you can specify the LINEPRINTER option to request line printer plots.
of _STDDEV_ is an estimate or a standard value.
For more information, see “OUTLIMITS= Data Set” on page 438.
Reading Cusum Scheme Parameters

This example shows how the cusum parameters saved in the previous example can be applied to new measurements saved in a data set named CANS2:

```sas
data cans2;
  length pump $ 8;
  label hour = 'Hour';
  input hour weight pump $ 8. ;
  cards;
  16 8.1765 Pump 3
  17 8.0949 Pump 3
  18 8.1393 Pump 3
  19 8.1491 Pump 3
  20 8.0473 Pump 1
  21 8.1602 Pump 1
  22 8.0633 Pump 1
  23 8.0921 Pump 1
  24 8.1573 Pump 1
  25 8.1304 Pump 1
  26 8.0979 Pump 1
  27 8.2407 Pump 1
  28 8.0730 Pump 1
  29 8.0986 Pump 2
  30 8.0785 Pump 2
  31 8.2308 Pump 2
  32 8.0986 Pump 2
  33 8.0782 Pump 2
  34 8.1435 Pump 2
  35 8.0666 Pump 2
;```

The following statements create a one-sided cusum chart for the measurements in CANS2 using the parameters in CUSPARM:

```sas
title 'One-Sided Cusum Analysis for New Data';
proc cusum data=cans2 limits=cusparm;
  xchart weight*hour (pump);
  label weight = 'Cusum of Weight';
run;
```

The LIMITS= option in the PROC CUSUM statement specifies the data set containing preestablished cusum parameters. The chart, shown in Figure 12.11, indicates that the process is in control. Levels of the variable PUMP (referred to as a block-variable) do not enter into the analysis but are displayed in a block legend across the top of the chart. See “Block Variable Legend Options” in Table 12.10 on page 415.

*In Release 6.09 and in earlier releases, it is also necessary to specify the READLIMITS option or the READINDEX= option to read cusum parameters from a LIMITS= data set.
In general, the parameters for a specified process and subgroup-variable are read from the first observation in the LIMITS= data set for which

- the value of _VAR_ matches the process (in this case, WEIGHT)
- the value of _SUBGRP_ matches the subgroup-variable name (in this case, HOUR)

If you are maintaining more than one set of cusum parameters for a particular process, you will find it convenient to include a special identifier variable named _INDEX_ in the LIMITS= data set. This must be a character variable of length 16. Then, if you specify READINDEX='value' in the XCHART statement, the parameters for a specified process and subgroup-variable are read from the first observation in the LIMITS= data set for which

- the value of _VAR_ matches process
- the value of _SUBGRP_ matches the subgroup-variable name
- the value of _INDEX_ matches value

In this example, the LIMITS= data set was created in a previous run of the CUSUM procedure. You can also create a LIMITS= data set with the DATA step. See “LIMITS= Data Set” on page 441 for details concerning the variables that you must provide.
Part 2. The CAPABILITY Procedure

Syntax

Specifying One-Sided Schemes
The basic syntax for a one-sided (decision interval) scheme using the XCHART statement is as follows:

```
XCHART process*subgroup-variable / SCHEME=ONESIED MU0=target DELTA=shift H=h <options> ;
```

The general form of this syntax is as follows:

```
XCHART (processes)*subgroup-variable <(block-variables )> < =symbol-variable | ='character' >/SCHEME=ONESIED MU0=target DELTA=shift H=h <options> ;
```

Note that the options SCHEME=ONESIED, MU0=, DELTA=, and H= are required unless their values are read from a LIMITS= data set.

Specifying Two-Sided Schemes
The basic syntax for a two-sided (V-mask) scheme is as follows:

```
XCHART process*subgroup-variable / MU0=target DELTA=shift ALPHA=alpha|H=h <options> ;
```

The general form of this syntax is as follows:

```
XCHART (processes)*subgroup-variable <(block-variables )> < =symbol-variable | ='character' >/MU0=target DELTA=shift ALPHA=alpha|H=h <options> ;
```

Note that the options MU0=, DELTA=, and either ALPHA= or H= are required unless their values are read from a LIMITS= data set.

Components of the XCHART Statement
You can use any number of XCHART statements in the CUSUM procedure. The components of the XCHART statement are described as follows.

```
process
processes
```

identify one or more processes to be analyzed. The specification of process depends on the input data set specified in the PROC CUSUM statement.

- If raw data are read from a DATA= data set, process must be the name of the variable containing the raw measurements. For an example, see “Creating a V-Mask Cusum Chart from Raw Data” on page 396.
- If summary data are read from a HISTORY= data set, process must be the common prefix of the summary variables in the HISTORY= data set. For an exam-
A process is required. If more than one process is specified, enclose the list in parentheses. The parameters specified in the XCHART statement are applied to all of the processes.*

subgroup-variable
is the variable that classifies the data into subgroups. The subgroup-variable is required. In the examples on pages 396 and 399, HOUR is the subgroup variable.

block-variables
are optionally specified variables that group the data into blocks of consecutive subgroups. The blocks are labeled in a legend, and each block-variable provides one level of labels in the legend. See Figure 12.11 for an example.

classification
is an optionally specified variable whose levels (unique values) determine the plotting character or symbol marker used to plot the cusums.

- If you produce a chart on a line printer, an ‘A’ marks points corresponding to the first level of the symbol-variable, a ‘B’ marks points corresponding to the second level, and so on.
- If you produce a chart on a graphics device, distinct symbol markers are displayed for points corresponding to the various levels of the symbol-variable. You can specify the symbol markers with SYMBOLn statements.

classification
specifies a plotting character for charts produced on line printers. See Figure 12.9 for an example.

options
specify optional cusum parameters, enhance the appearance of the chart, request additional analyses, save results in data sets, and so on. The “Summary of Options” section, which follows, lists all options by function.

Summary of Options

The following tables list the XCHART statement options by function. Options unique to the CUSUM procedure are listed in Table 12.1 to Table 12.5, and they are described in detail in “Dictionary of Special Options” on page 419. Options that are common to both the CUSUM and SHEWHART procedures are listed in Table 12.7 to Table 12.16. They are described in detail beginning on page 1733 of Part 9, “The SHEWHART Procedure,”.

*For this reason, it may be preferable to read distinct cusum parameters for each process from a LIMITS= data set.
Table 12.1. Options for Specifying a One-Sided (Decision Interval) Cusum Scheme*

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DELTA=value</td>
<td>specifies shift to be detected as a multiple of standard error</td>
</tr>
<tr>
<td>H=value</td>
<td>specifies decision interval $h$ ($h &gt; 0$) as a multiple of standard error</td>
</tr>
<tr>
<td>HEADSTART=value</td>
<td>specifies headstart value $S_0$ as a multiple of standard error</td>
</tr>
<tr>
<td>K=value</td>
<td>specifies reference value $k$ ($k &gt; 0$)</td>
</tr>
<tr>
<td>LIMITN=n</td>
<td>specifies fixed nominal sample size for cusum scheme</td>
</tr>
<tr>
<td>LIMITN=VARYING</td>
<td>specifies that cusums are to be computed for all subgroups regardless of sample size</td>
</tr>
<tr>
<td>MU0=value</td>
<td>specifies target $\mu_0$ for mean</td>
</tr>
<tr>
<td>NOREADLIMITS</td>
<td>specifies that cusum parameters are not to be read from LIMITS= data set</td>
</tr>
<tr>
<td>READINDEX='value'</td>
<td>reads cusum scheme parameters from a LIMITS= data set</td>
</tr>
<tr>
<td>READLIMITS</td>
<td>specifies that cusum parameters are to be read from LIMITS= data set</td>
</tr>
<tr>
<td>SCHEME=ONESIDED</td>
<td>specifies a one-sided scheme</td>
</tr>
<tr>
<td>SHIFT=value</td>
<td>specifies shift to be detected in data units</td>
</tr>
<tr>
<td>SIGMA0=value</td>
<td>specifies standard (known) value $\sigma_0$ for process standard deviation</td>
</tr>
</tbody>
</table>

Table 12.2. Options for Specifying a Two-Sided (V-Mask) Cusum Scheme*

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALPHA=value</td>
<td>specifies probability of Type 1 error</td>
</tr>
<tr>
<td>BETA=value</td>
<td>specifies probability of Type 2 error</td>
</tr>
<tr>
<td>H=value</td>
<td>specifies vertical distance between V-mask origin and upper (or lower) arm</td>
</tr>
<tr>
<td>K=value</td>
<td>specifies slope of lower arm of V-mask</td>
</tr>
<tr>
<td>LIMITN=n</td>
<td>specifies fixed nominal sample size for cusum scheme</td>
</tr>
<tr>
<td>LIMITN=VARYING</td>
<td>specifies that cusums are to be computed for all subgroups regardless of sample size</td>
</tr>
<tr>
<td>NOREADLIMITS</td>
<td>specifies that cusum parameters are not to be read from LIMITS= data set</td>
</tr>
<tr>
<td>READINDEX='value'</td>
<td>reads cusum scheme parameters from a LIMITS= data set</td>
</tr>
<tr>
<td>READLIMITS</td>
<td>specifies that cusum parameters are to be read from LIMITS= data set</td>
</tr>
<tr>
<td>READSIGMAS</td>
<td>reads _SIGMAS_ instead of _ALPHA_ from LIMITS= data set when both variables are available</td>
</tr>
<tr>
<td>SIGMAS=value</td>
<td>specifies probability of Type 1 error as probability that standard normally distributed variable exceeds \textit{value} in absolute value</td>
</tr>
</tbody>
</table>

* The options in these tables are described in the “Dictionary of Special Options” section on page 419.
Table 12.3. Options for Estimating Process Standard Deviation †

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMETHOD=keyword</td>
<td>specifies method for estimating process standard deviation ( \sigma )</td>
</tr>
<tr>
<td>TYPE=keyword</td>
<td>identifies whether _STDDEV_ in OUTLIMITS= data set is an estimate or standard, and specifies value of _TYPE_ in OUTLIMITS= data set</td>
</tr>
</tbody>
</table>

Table 12.4. Options for Displaying Decision Interval or V-Mask †

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CINFILL=color</td>
<td>specifies color for area under decision interval line or inside V-mask</td>
</tr>
<tr>
<td>CLIMITS=color</td>
<td>specifies color of decision interval line</td>
</tr>
<tr>
<td>CMASK=color</td>
<td>specifies color of V-mask outline</td>
</tr>
<tr>
<td>LLIMITS=linetype</td>
<td>specifies line type for decision interval</td>
</tr>
<tr>
<td>LMASK=linetype</td>
<td>specifies line type for V-mask arms</td>
</tr>
<tr>
<td>NOMASK</td>
<td>suppresses display of V-mask</td>
</tr>
<tr>
<td>ORIGIN=value</td>
<td>specifies value of subgroup-variable locating origin of V-mask</td>
</tr>
<tr>
<td>WLIMITS=n</td>
<td>specifies line width for decision interval</td>
</tr>
<tr>
<td>WMASK=n</td>
<td>specifies line width for V-mask</td>
</tr>
</tbody>
</table>

Table 12.5. Tabulation Options †

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>TABLEALL</td>
<td>specifies the options TABLECHART, TABLECOMP, TABLEID, TABLEOUT, and TABLESUMMARY</td>
</tr>
<tr>
<td>TABLECHART</td>
<td>tabulates the information displayed in the cusum chart</td>
</tr>
<tr>
<td>TABLECOMP</td>
<td>tabulates the computational form of the cusum scheme as described by Lucas (1976) and Lucas and Crosier (1982)</td>
</tr>
<tr>
<td>TABLEID</td>
<td>augments TABLECHART and TABLECOMP tables with columns for ID variables</td>
</tr>
<tr>
<td>TABLEOUT</td>
<td>augments TABLECHART table with a column indicating if the decision interval or V-mask was exceeded</td>
</tr>
<tr>
<td>TABLESUMMARY</td>
<td>tabulates the parameters for the cusum scheme and the average run lengths corresponding to shifts of zero and ( \delta )</td>
</tr>
</tbody>
</table>

Note that specifying (EXCEPTIONS) after the option TABLECHART creates a table for exceptional points only.

Table 12.6. Input Data Set Options †

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>MISSBREAK</td>
<td>specifies that observations with missing values are not to be processed</td>
</tr>
</tbody>
</table>

†The options in these tables are described in the “Dictionary of Special Options” section on page 419.
Part 2. The CAPABILITY Procedure

Table 12.7. Output Data Set Options

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMAGEMAP= SAS-data-set</td>
<td>creates OUTTABLE= data set with additional graph coordinate data</td>
</tr>
<tr>
<td>OUTHISTORY= SAS-data-set</td>
<td>creates output data set containing subgroup summary statistics and cusums</td>
</tr>
<tr>
<td>OUTINDEX='string'</td>
<td>specifies value of <em>INDEX</em> in OUTLIMITS= data set</td>
</tr>
<tr>
<td>OUTLIMITS= SAS-data-set</td>
<td>creates output data set containing cusum parameters</td>
</tr>
<tr>
<td>OUTPHASE='string'</td>
<td>specifies value of <em>PHASE</em> in OUTHISTORY= data set or OUTTABLE= data set</td>
</tr>
<tr>
<td>OUTTABLE= SAS-data-set</td>
<td>creates output data set containing subgroup summary statistics, cusums, and decision interval or V-mask values</td>
</tr>
</tbody>
</table>

Table 12.8. Axis and Axis Label Options

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAXIS=color</td>
<td>specifies color for axis lines and tick marks</td>
</tr>
<tr>
<td>CFRAME=color</td>
<td>specifies fill color(s) for frame for plot area</td>
</tr>
<tr>
<td>(color-list)</td>
<td></td>
</tr>
<tr>
<td>CTEXT=color</td>
<td>specifies color for tick mark values and axis labels</td>
</tr>
<tr>
<td>HAXIS=values</td>
<td>AXISn</td>
</tr>
<tr>
<td>HEIGHT=value</td>
<td>specifies height of axis label and axis legend text</td>
</tr>
<tr>
<td>HMINOR=n</td>
<td>specifies number of minor tick marks between major tick marks on horizontal axis</td>
</tr>
<tr>
<td>HOFFSET=value</td>
<td>specifies length of offset at both ends of horizontal axis</td>
</tr>
<tr>
<td>NOHLABEL</td>
<td>suppresses label for horizontal axis</td>
</tr>
<tr>
<td>NOVANGLE</td>
<td>requests vertical axis labels that are strung out vertically</td>
</tr>
<tr>
<td>SKIPHLABELS=n</td>
<td>specifies thinning factor for tick mark labels on horizontal axis</td>
</tr>
<tr>
<td>SPLIT='character'</td>
<td>specifies splitting character for axis labels</td>
</tr>
<tr>
<td>TURNHLABELS</td>
<td>requests horizontal axis labels that are strung out vertically</td>
</tr>
<tr>
<td>VAXIS=value-list</td>
<td>name</td>
</tr>
<tr>
<td>VAXIS2=value-list</td>
<td>name</td>
</tr>
<tr>
<td>VMINOR=n</td>
<td>specifies number of minor tick marks between major tick marks on vertical axis</td>
</tr>
<tr>
<td>VOFFSET=value</td>
<td>specifies length of offset at both ends of vertical axis</td>
</tr>
<tr>
<td>WAXIS=n</td>
<td>specifies width of axis lines</td>
</tr>
</tbody>
</table>

†The options in these tables are described in Chapter 43, “Dictionary of Options,” of Part 9, “The SHEWHART Procedure,”.
### Table 12.9.  Reference Line Options

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHREF=color</td>
<td>specifies color for HREF= and HREF2= lines</td>
</tr>
<tr>
<td>CVREF=color</td>
<td>specifies color for VREF= and VREF2= lines</td>
</tr>
<tr>
<td>HREF=values</td>
<td>SAS-data-set</td>
</tr>
<tr>
<td>HREF2=values</td>
<td>SAS-data-set</td>
</tr>
<tr>
<td>HREFCHAR='character'</td>
<td>specifies line character for HREF= and HREF2= lines</td>
</tr>
<tr>
<td>HREFLABELS=</td>
<td>specifies labels for HREF= lines</td>
</tr>
<tr>
<td>HREF2LABELS=</td>
<td>specifies labels for HREF2= lines</td>
</tr>
<tr>
<td>HREFLABPOS=n</td>
<td>specifies position of HREFLABELS= and HREF2LABELS= labels</td>
</tr>
<tr>
<td>LHREF=linetype</td>
<td>specifies line type for HREF= and HREF2= lines</td>
</tr>
<tr>
<td>LVREF=linetype</td>
<td>specifies line type for VREF= and VREF2= lines</td>
</tr>
<tr>
<td>NOBYREF</td>
<td>specifies that reference line information in a data set is to be applied uniformly to charts created for all BY groups</td>
</tr>
<tr>
<td>VREF=values</td>
<td>SAS-data-set</td>
</tr>
<tr>
<td>VREF2=values</td>
<td>SAS-data-set</td>
</tr>
<tr>
<td>VREFCHAR='character'</td>
<td>specifies line character for VREF= and VREF2= lines</td>
</tr>
<tr>
<td>VREFLABELS=</td>
<td>specifies labels for VREF= lines</td>
</tr>
<tr>
<td>VREF2LABELS=</td>
<td>specifies labels for VREF2= lines</td>
</tr>
<tr>
<td>VREFLABPOS=n</td>
<td>specifies position of VREFLABELS= and VREF2LABELS= labels</td>
</tr>
</tbody>
</table>

### Table 12.10.  Block Variable Legend Options

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BLOCKLABELPOS=</td>
<td>specifies position of label for block-variable legend</td>
</tr>
<tr>
<td>BLOCKLABTYPE=</td>
<td>specifies text size of block-variable legend</td>
</tr>
<tr>
<td>BLOCKPOS=n</td>
<td>specifies vertical position of block-variable legend</td>
</tr>
<tr>
<td>BLOCKREP</td>
<td>repeats identical consecutive labels in block-variable legend</td>
</tr>
<tr>
<td>CBLOCKLAB=color</td>
<td>specifies color for filling background in block-variable legend</td>
</tr>
<tr>
<td>CBLOCKVAR=variable)</td>
<td>specifies one or more variables whose values are colors for filling background of block-variable legend</td>
</tr>
</tbody>
</table>

†The options in these tables are described in Chapter 43, “Dictionary of Options,” of Part 9, “The SHEWHART Procedure.”.
Table 12.11. Options for Plotting and Labeling Points

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALLLABEL=VALUE[variable]</td>
<td>labels every point on cusum chart</td>
</tr>
<tr>
<td>ALLLABEL2=VALUE[variable]</td>
<td>labels every point on trend chart</td>
</tr>
<tr>
<td>CCONNECT=COLOR</td>
<td>specifies color for line segments that connect points on chart</td>
</tr>
<tr>
<td>CFRAMELAB=color</td>
<td>specifies fill color for frame around labeled points</td>
</tr>
<tr>
<td>CONNECTCHAR='character'</td>
<td>specifies character used to form line segments that connect points on chart</td>
</tr>
<tr>
<td>COUT=color</td>
<td>specifies color for connecting line segments that lie above the decision interval or outside the V-mask</td>
</tr>
<tr>
<td>COUTFILL=color</td>
<td>specifies color for areas between connected points and decision interval or V-mask</td>
</tr>
<tr>
<td>NOCONNECT</td>
<td>suppresses line segments that connect points on chart</td>
</tr>
<tr>
<td>NOTRENDCONNECT</td>
<td>suppresses line segments that connect points on trend chart</td>
</tr>
<tr>
<td>OUTLABEL=VALUE[variable]</td>
<td>labels points exceeding decision interval on one-sided chart</td>
</tr>
<tr>
<td>SYMBOLCHARS='characters'</td>
<td>specifies characters indicating symbol-variable</td>
</tr>
<tr>
<td>SYMBOLLEGEND=None[name]</td>
<td>specifies LEGEND statement for levels of symbol-variable</td>
</tr>
<tr>
<td>SYMBOLORDER=keyword</td>
<td>specifies order in which symbols are assigned for levels of symbol-variable</td>
</tr>
</tbody>
</table>

Table 12.12. Phase Options

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPHASELEG=color</td>
<td>specifies text color for phase legend</td>
</tr>
<tr>
<td>OUTPHASE='string'</td>
<td>specifies value of <em>PHASE</em> in OUTHISTORY=data set</td>
</tr>
<tr>
<td>PHASEBREAK</td>
<td>disconnects last point in a phase from first point in next phase</td>
</tr>
<tr>
<td>PHASELABTYPE=value[ keyword]</td>
<td>specifies text size of phase legend</td>
</tr>
<tr>
<td>PHASELEGEND</td>
<td>displays phase labels in a legend across top of chart</td>
</tr>
<tr>
<td>PHASEREF</td>
<td>delineates phases with vertical reference lines</td>
</tr>
<tr>
<td>READPHASES=ALL</td>
<td>'label1'...'labeln'</td>
</tr>
</tbody>
</table>

Table 12.13. Grid Options

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENDGRID</td>
<td>adds grid after last plotted point</td>
</tr>
<tr>
<td>GRID</td>
<td>adds grid to chart</td>
</tr>
<tr>
<td>LENDGRID=linetype</td>
<td>specifies line type for grid requested with the ENDGRID option</td>
</tr>
<tr>
<td>LGRID=linetype</td>
<td>specifies line type for grid requested with the GRID option</td>
</tr>
<tr>
<td>WGRID=n</td>
<td>specifies width of grid lines</td>
</tr>
</tbody>
</table>

*The options in these tables are described in Chapter 43, “Dictionary of Options,” of Part 9, “The SHEWHART Procedure.”*
### Table 12.14. Graphical Enhancement Options

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANNOTATE=</td>
<td>specifies annotate data set that adds features to cusum chart</td>
</tr>
<tr>
<td><strong>SAS-data-set</strong></td>
<td></td>
</tr>
<tr>
<td>ANNOTATE2=</td>
<td>specifies annotate data set that adds features to trend chart</td>
</tr>
<tr>
<td><strong>SAS-data-set</strong></td>
<td></td>
</tr>
<tr>
<td>DESCRIPTION='string'</td>
<td>specifies string that appears in the description field of PROC GREPLAY master menu for cusum chart</td>
</tr>
<tr>
<td>FONT=font</td>
<td>specifies software font for labels and legends on charts</td>
</tr>
<tr>
<td>HTML=(variable)</td>
<td>specifies a variable whose values are URLs to be associated with subgroups</td>
</tr>
<tr>
<td>NAME='string'</td>
<td>specifies name that appears in the name field of the PROC GREPLAY master menu for cusum chart</td>
</tr>
<tr>
<td>PAGENUM='string'</td>
<td>specifies the form of the label used in pagination</td>
</tr>
<tr>
<td>NAMEPOS=</td>
<td>specifies the position of the page number requested with the PAGENUM= option</td>
</tr>
<tr>
<td>WTREND=n</td>
<td>specifies width of line segments connecting points on trend chart</td>
</tr>
</tbody>
</table>

### Table 12.15. Plot Layout Options

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALLN</td>
<td>plots summary statistics for all subgroups</td>
</tr>
<tr>
<td>BILEVEL</td>
<td>creates control charts using half-screens and half-pages</td>
</tr>
<tr>
<td>EXCHART</td>
<td>creates cusum chart only when exceptions occur</td>
</tr>
<tr>
<td>INTERVAL=keyword</td>
<td>specifies natural time interval between consecutive subgroup positions when time, date, or datetime format is associated with a numeric subgroup variable</td>
</tr>
<tr>
<td>MAXPANELS=n</td>
<td>specifies maximum number of pages or screens for chart</td>
</tr>
<tr>
<td>NMARKERS</td>
<td>requests special markers for points corresponding to sample sizes not equal to nominal sample size for fixed control limits</td>
</tr>
<tr>
<td>NOCHART</td>
<td>suppresses creation of cusum chart</td>
</tr>
<tr>
<td>NOFRAME</td>
<td>suppresses frame for plot area</td>
</tr>
<tr>
<td>NOLEGEND</td>
<td>suppresses legend for subgroup sample sizes and cusum parameters</td>
</tr>
<tr>
<td>NpanelPOS=n</td>
<td>specifies number of subgroup positions per panel on each chart</td>
</tr>
<tr>
<td>REPEAT</td>
<td>repeats last subgroup position on panel as first subgroup position of next panel</td>
</tr>
<tr>
<td>TOTPANELS=n</td>
<td>specifies number of pages or screens to be used to display chart</td>
</tr>
<tr>
<td>TRENDVAR=variable[ (variable-list)]</td>
<td>specifies list of trend variables</td>
</tr>
<tr>
<td>YPCT1=value</td>
<td>specifies length of vertical axis on cusum chart as a percentage of sum of lengths of vertical axes for cusum and trend charts</td>
</tr>
</tbody>
</table>

The options in these tables are described in Chapter 43, “Dictionary of Options,” of Part 9, “The SHEWHART Procedure.”.
Table 12.16. Star Options**

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSTARCIRCLES=color</td>
<td>specifies color for circles requested by the STARCIRCLES= option</td>
</tr>
<tr>
<td>CSTARFILL=color</td>
<td>specifies color for filling stars</td>
</tr>
<tr>
<td>CSTAROUT=color</td>
<td>specifies color for outlines of stars that exceed inner or outer circles</td>
</tr>
<tr>
<td>CSTARS=color</td>
<td>specifies color for outlines of stars</td>
</tr>
<tr>
<td>LABELFONT=font</td>
<td>specifies font for labels added with the STARLABEL= option</td>
</tr>
<tr>
<td>LSTARCIRCLES=linetypes</td>
<td>specifies line types of circles requested by the STARCIRCLES= option</td>
</tr>
<tr>
<td>LSTARS=linetype</td>
<td>specifies line types of outlines of stars requested by the STARVERTICES= option</td>
</tr>
<tr>
<td>STARBDRADIUS=value</td>
<td>specifies radius of outer bound circle for vertices of stars</td>
</tr>
<tr>
<td>STARCIRCLES=value-list</td>
<td>specifies reference circles for stars</td>
</tr>
<tr>
<td>STARINRADIUS=value</td>
<td>specifies inner radius of stars</td>
</tr>
<tr>
<td>STARLABEL=keyword</td>
<td>specifies vertices to be labeled</td>
</tr>
<tr>
<td>STARLEGEND=keyword</td>
<td>specifies style of legend for star vertices</td>
</tr>
<tr>
<td>STARLEGENDLAB='label'</td>
<td>specifies label for STARLEGEND= legend</td>
</tr>
<tr>
<td>STAROUTRADIUS=value</td>
<td>specifies outer radius of stars</td>
</tr>
<tr>
<td>STARSPEC=value</td>
<td>specifies method used to standardize vertex variables</td>
</tr>
<tr>
<td>STARSTART=value</td>
<td>specifies angle for first vertex</td>
</tr>
<tr>
<td>STARTYPE=keyword</td>
<td>specifies graphical style of star</td>
</tr>
<tr>
<td>STARVERTICES=variable</td>
<td>superimposes star at each point on chart</td>
</tr>
<tr>
<td>WSTARCIRCLES=n</td>
<td>specifies width of circles requested by the STARCIRCLES= option</td>
</tr>
<tr>
<td>WSTARS=n</td>
<td>specifies width of stars requested by the STARVERTICES= option</td>
</tr>
</tbody>
</table>

**The options in these tables are described in Chapter 43, “Dictionary of Options,” of Part 9, “The SHEWHART Procedure.”
Dictionary of Special Options

The marginal notes *Graphics* and *Line Printer* identify options that apply to graphics devices and line printers, respectively.

**ALPHA= value**

specifies the probability $\alpha$ of incorrectly deciding that a shift has occurred when the process mean is equal to the target mean. This is known as the probability of a Type 1 error. The *value* must be between zero and one, and it is typically set at 0.05 or 0.10. If you specify the ALPHA= option, the error probability approach is used to determine the V-mask. For details, see “Defining the V-Mask for a Two-Sided Cusum Scheme” on page 429.

The ALPHA= option is applicable only with two-sided cusum schemes. As an alternative to the ALPHA= value, you can specify the percentile $z_{1-\alpha/2}$ from a standard normal distribution with the SIGMAS= option. As a second alternative, you can specify the geometric parameter $h$ for the V-mask (in standard error units) with the H= option.

In addition to the ALPHA= option, you can optionally specify the probability of a Type 2 error with the BETA= option.

**BETA= value**

specifies the probability $\beta$ of failing to discover that the specified shift has occurred. This is known as the probability of a Type 2 error. The *value* must be between zero and one. The BETA= option is used in conjunction with either the ALPHA= option or the SIGMAS= option.

The interpretation of $\beta$ is based on the analogy between cusum charts and sequential probability ratio tests, and it is inexact since the cusum chart does not provide an acceptance region. Refer to Johnson (1961) and van Dobben de Bruyn (1968) for further details.

**CINFILL= color**

specifies the color for the area under the decision interval or inside the V-mask arms. By default, these areas are not filled with a color. See also the COUTFILL= option.

**CLIMITS= color**

specifies the color for the decision interval line. The default color is the first color in the device color list.

**CMASK= color**

specifies the color for the V-mask arms. The default color is the first color in the device color list.

**DATAUNITS**

computes cumulative sums without standardizing the subgroup means or individual measurements. As a result, the vertical axis of the cusum chart is scaled in the same units as the data.
Part 2. The CAPABILITY Procedure

The DATAUNITS option requires constant subgroup sample sizes. If your data do not have constant subgroup sample sizes, you need to specify a constant nominal sample size \( n \) for the V-mask or decision interval with the LIMITN= option or with the variable _LIMITN_ in the LIMITS= data set.

**DELTA=**\( value \)

specifies the absolute value of the smallest shift to be detected as a multiple \( \delta \) of the process standard deviation \( \sigma \) or the standard error \( \sigma \bar{X} \), depending on whether \( \delta \) is viewed as a shift in the population mean or a shift in the sampling distribution of the subgroup mean \( \bar{X} \), respectively.

If you specify SCHEME=ONESIDED (see the SCHEME= option later in this list) and the \( value \) is positive, a shift above the process mean is to be detected, whereas if the \( value \) is negative, a shift below the process mean is to be detected.

As an alternative to specifying the DELTA= option, you can specify the shift in the same units as the data with the SHIFT= option.

**H=**\( value \)

specifies the decision interval \( h \) for a one-sided cusum scheme. This type of scheme is completely specified by the parameters \( h \) and \( k \) (see the K= option later in this list). You can also specify the H= option as an alternative to the ALPHA= or SIGMAS= options for a two-sided cusum scheme with a V-mask. In this case, the H= option specifies the vertical distance \( h \) between the origin for the V-mask and the upper or lower arm of the V-mask. In either case, the H= \( value \) must be positive and must be expressed as a multiple of standard error.

You can use a table of average run lengths to choose \( h \) (this is typically between zero and 10). See pages 433 and 434.

**HEADSTART=**\( value \)

**HSTART=**\( value \)

specifies a headstart value \( S_0 \) for a one-sided cusum scheme. The value must be expressed as a multiple of standard error. See "Headstart Values" on page 426, and refer to Lucas and Crosier (1982), Ryan (1989), and Montgomery (1991).

**K=**\( value \)

specifies the reference value \( k \) for a one-sided (decision interval) cusum scheme. This type of scheme is completely specified by the parameters \( k \) and \( h \) (see the H= option earlier in this list). You can also specify the K= and H= options as geometric parameters for a two-sided cusum scheme with a V-mask. In this case, the K= option specifies the slope of the lower arm of the V-mask, and the K= and H= options together are alternatives to the error probability options ALPHA=, SIGMAS=, and BETA=. In either case, the K= \( value \) must be positive and must be expressed as a multiple of standard error.

You can use a table of average run lengths to choose \( k \) and \( h \) (\( k \) is typically between zero and two). See pages 433 and 434.

For a one-sided scheme, the default K= \( value \) is \( \delta / 2 \), which is referred to as the central reference value. For a two-sided scheme where the V-mask is specified geometrically with the H= option, the default K= \( value \) is \( \delta / 2 \). If, however, the V-mask is specified
by an error probability with the ALPHA= option, then the K= option should not be specified.

**CAUTION:** The interpretation of the K= value depends on the subgroup-variable and the interval between subgroups that is specified with the INTERVAL= option. For a two-sided scheme, the value is the increase in the lower V-mask arm per unit change on the subgroup axis, so the value depends on how the subgroup-variable is scaled.

- If integer values are assigned to the subgroup-variable, then a unit change is defined as one.
- If the subgroup-variable has character values, then a unit change is defined as the increment between adjacent values of the subgroup-variable.
- If the subgroup-variable is numeric and is formatted with a SAS date or time format, then a unit change is defined as the default value for the INTERVAL= option. For example, if a DATE7. format is associated with the subgroup-variable, then a unit change is defined as one day.

You can use the INTERVAL= option to modify the definition of a unit change. For example, if a DATE7. format is associated with the subgroup-variable but subgroups are collected hourly, then INTERVAL=HOUR defines a unit change as one hour rather than one day.

**LIMITN=\textit{n}**

**LIMITN=VARYING**

specifies either a fixed or varying nominal sample size for the control limits. If you specify LIMITN=\textit{n}, cusums are calculated and displayed only for those subgroups with a sample size equal to \textit{n}, although you can specify the ALLN option to force all cusums to be plotted. If you specify LIMITN=VARYING, cusums are calculated and displayed for all subgroups, regardless of sample size.

**LLIMITS=\textit{linetype}**

specifies the line type for the decision interval. The default is 4 (a dashed line).

**LMASK=\textit{linetype}**

specifies the line type for the V-mask arms. The default is 1 (a solid line).

**MU0=\textit{value}**

specifies the target mean \( \mu_0 \) for the process. The target mean must be scaled in the same units as the data.

**NOARL**

suppresses calculation of average run lengths. By default, this calculation is performed if you specify the TABLESUMMARY option or an OUTLIMITS= data set.

**NOMASK**

suppresses the display of the V-mask on charts for two-sided schemes. This option does not affect computations of cusums or V-mask parameters.

**NOREADLIMITS**

specifies that the cusum scheme parameters for each process listed in the chart statement are not to be read from the LIMITS= data set specified in the PROC CUSUM statement. The NOREADLIMITS option is available only in Release 6.10 and later.
Part 2. The CAPABILITY Procedure

releases. See the READLIMITS option later in this list.

**ORIGIN=value**
specifies the origin of the V-mask, which is defined as the horizontal coordinate of the right edge of the V-mask. If a date, time, or datetime format is associated with the subgroup-variable, you must specify the value as a date, time, or datetime constant, respectively. If the subgroup variable is character, you must specify the value as a quoted string. The default value is the last (most recent) value of the subgroup-variable.

Note that estimates for the process mean and standard deviation are calculated only from subgroups up to and including the origin subgroup.

**READINDEX=’value’**
reads cusum scheme parameters from a LIMITS= data set (specified in the PROC CUSUM statement) for each process listed in the chart statement. The \( i \)th set of control limits for a particular process is read from the first observation in the LIMITS= data set for which

- the value of \( \_VAR\_ \) matches process
- the value of \( \_SUBGRP\_ \) matches the subgroup-variable
- the value of \( \_INDEX\_ \) matches value

The value can be up to 16 characters and must be enclosed in quotes.

**READLIMITS**
specifies that cusum scheme parameters are to be read from a LIMITS= data set specified in the PROC CUSUM statement. The parameters for a particular process are read from the first observation in the LIMITS= data set for which

- the value of \( \_VAR\_ \) matches process
- the value of \( \_SUBGRP\_ \) matches the subgroup variable

The use of the READLIMITS option depends on which release of SAS/QC software you are using.

- **In Release 6.10 and later releases, the READLIMITS option is not necessary.** To read cusum scheme parameters as described previously, you simply specify a LIMITS= data set. However, even though the READLIMITS option is redundant, it continues to function as in earlier releases.

- **In Release 6.09 and earlier releases, you must specify the READLIMITS option to read cusum scheme parameters as described previously.** If you specify a LIMITS= data set without specifying the READLIMITS option (or the READINDEX= option), the cusum scheme parameters are computed from the data.

**READSIGMAS**
specifies that the variable \( \_SIGMAS\_ \) (instead of \( \_ALPHA\_ \)) is to be read from a LIMITS= data set that contains both variables. The variables \( \_SIGMAS\_ \) and
Chapter 12. Syntax

_ALPHA_ provide the same parameters as the SIGMAS= and ALPHA= options. By default, _ALPHA_ is read from the LIMITS= data set.

**SCHEME=ONESIDED**  
**SCHEME=TWOSIDED**

indicates whether the cusum scheme is a one-sided (decision interval) scheme or a two-sided scheme with a V-mask. By default, SCHEME=TWOSIDED.

**SHIFT=value**

specifies the shift to be detected in the same units as the data. The value is interpreted as the shift in the mean of the sampling distribution of the subgroup mean. The SHIFT= option is an alternative to the DELTA= option. To specify the SHIFT= option, one of the following must be true:

- The subgroup sample sizes are constant.
- A constant nominal sample size $n$ is provided for the cusum scheme with the LIMITN= option or the _LIMITN_ variable in a LIMITS= data set.

The relationship between the SHIFT= value (denoted by $\Delta$) and the DELTA= value (denoted by $\delta$) is $\delta = \Delta/(\sigma/\sqrt{n})$, where $\sigma$ is the process standard deviation.

**SIGMA0=value**

specifies a known standard deviation $\sigma_0$ for the process standard deviation $\sigma$. The value must be positive. By default, PROC CUSUM estimates $\sigma$ from the data using the formulas given in “Methods for Estimating the Standard Deviation” on page 435. You can use the variable _STDDEV_ in a LIMITS= data set as an alternative to the SIGMA0= option.

**SIGMAS=value**

specifies the probability $\alpha$ of false detection for a two-sided cusum scheme with a V-mask as the probability that the absolute value of a standard normally distributed variable is greater than the value. For example, SIGMAS=3 corresponds to the probability $\alpha = 0.0027$. The value must be positive. The SIGMAS= option is an alternative to the ALPHA= and H= options, and only one of these three options can be specified.

The SIGMAS= option is useful for defining cusum charts that correspond to Shewhart charts whose control limits are defined with the same value as the multiple of $\sigma$. Refer to Johnson and Leone (1962, 1974).

**SMETHOD=NOWEIGHT | MVLUE | RMSDF**

specifies a method for estimating the process standard deviation from subgroup observations, $\sigma$, as summarized by the following table.

<table>
<thead>
<tr>
<th>Keyword</th>
<th>Method for Estimating Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOWEIGHT</td>
<td>estimates $\sigma$ as an unweighted average of unbiased subgroup estimates of $\sigma$</td>
</tr>
<tr>
<td>MVLUE</td>
<td>calculates a minimum variance linear unbiased estimate for $\sigma$</td>
</tr>
<tr>
<td>RMSDF</td>
<td>calculates a root-mean square estimate for $\sigma$</td>
</tr>
</tbody>
</table>

For formulas, see “Methods for Estimating the Standard Deviation” on page 435.
TABLEALL
requests all the tables specified by the options TABLECHART, TABLECOMP, TABLEID, TABLEOUT, and TABLESUMMARY.

TABLECHART <(EXCEPTIONS)>
creates a table of the subgroup variable, the subgroup sample sizes, the subgroup means, the cumulative sums, and the decision interval or V-mask limits. A table is produced for each process specified in the XCHART statement. The keyword EXCEPTIONS (enclosed in parentheses) is optional and restricts the tabulation to those subgroups for which the decision interval or V-mask values are exceeded.

TABLECOMP
tabulates the computational form of the cusum scheme as described by Lucas (1976) and Lucas and Crosier (1982). Upper or lower cumulative sums (or both) are tabulated for each process given in the XCHART statement. See “Formulas for Cumulative Sums” on page 426 for more information.

TABLEID
augments the tables specified by the TABLECHART and TABLECOMP options with a column for each of the ID variables.

TABLEOUT
augments the table specified by the TABLECHART option with a column indicating whether the decision interval or V-mask values are exceeded.

TABLESUMMARY
produces a table that summarizes the cusum scheme. The table lists the parameters of the scheme and the average run lengths corresponding to shifts of zero and \( \delta \). The average run lengths are computed using the method of Goel and Wu (1971). A table is produced for each process. You can save the summary in a data set by specifying the OUTLIMITS= option. See “OUTLIMITS= Data Set” on page 438 for details.

TYPE=ESTIMATE
TYPE=STANDARD
specifies the value of _TYPE_ in an OUTLIMITS= data set. The variable _TYPE_ indicates whether the variable _STDDEV_ in the OUTLIMITS= data set represents an estimate or a standard (known) value. The default is STANDARD if the SIGMA0= option is specified; otherwise, the default is ESTIMATE.

WLIMITS=linetype
specifies the width (in pixels) of the decision interval line. The default width is 1.

WMASK=linetype
specifies the width (in pixels) of the V-mask arms. The default width is 1.
Basic Notation for Cusum Charts

The following notation is used in this chapter:

- $\mu$ denotes the mean of the population, also referred to as the process mean or the process level.

- $\mu_0$ denotes the target mean (goal) for the population. Goel and Wu (1971) refer to $\mu_0$ as the “acceptable quality level” and use the symbol $\mu_a$ instead. The symbol $\bar{X}_0$ is used for $\mu_0$ in Glossary and Tables for Statistical Quality Control. You can provide $\mu_0$ with the MU0= option or with the variable _MU0_ in a LIMITS= data set.

- $\sigma$ denotes the population standard deviation. You can provide $\sigma$ with the variable _STDDEV_ in a LIMITS= data set (where _TYPE_=STANDARD).

- $\sigma_0$ denotes a known standard deviation. You can provide $\sigma_0$ with the SIGMA0= option or the variable _STDDEV_ in a LIMITS= data set.

- $\hat{\sigma}$ denotes an estimate of $\sigma$. You can provide $\hat{\sigma}$ with the SIGMA0= option or the variable _STDDEV_ in a LIMITS= data set. To identify this value as an estimate, specify TYPE=ESTIMATE or assign the value ESTIMATE to the variable _TYPE_ in a LIMITS= data set.

- $n$ denotes the nominal sample size for the cusum scheme. You can provide $n$ with the LIMITN= option or the variable _LIMITN_ in a LIMITS= data set.

- $\delta$ denotes the shift in $\mu$ to be detected, expressed as a multiple of the standard deviation. You can provide $\delta$ with the DELTA= option or the variable _DELTA_ in a LIMITS= data set.

- $\Delta$ denotes the shift in $\mu$ to be detected, expressed in data units. If the sample size $n$ is constant across subgroups, then $\Delta = \delta \sigma \bar{X} = (\delta \sigma) / \sqrt{n}$.

Some authors use the symbol $D$ instead of $\Delta$; for example, refer to Johnson and Leone (1962, 1974) and Wadsworth and others (1986). You can provide $\Delta$ with the SHIFT= option. Although it may be more natural to specify the shift in data units, it is preferable to specify the shift as $\delta$, since this generalizes to data with unequal subgroup sample sizes.
Formulas for Cumulative Sums

One-Sided Cusum Schemes

Positive Shifts

If the shift $\delta$ to be detected is positive, the cusum computed for the $t$th subgroup is

$$S_t = \max(0, S_{t-1} + (z_t - k))$$

for $t=1, 2, \ldots, n$, where $S_0=0$, $z_t$ is defined as for two-sided schemes, and the parameter $k$, termed the reference value, is positive. The cusum $S_t$ is referred to as an upper cumulative sum. Since $S_t$ can be written as

$$\max \left( 0, S_{t-1} + \frac{\bar{X}_t - (\mu_0 + k\sigma\bar{X}_i)}{\sigma\bar{X}_i} \right)$$

the sequence $S_t$ cumulates deviations in the subgroup means greater than $k$ standard errors from $\mu_0$. If $S_t$ exceeds a positive value $h$ (referred to as the decision interval), a shift or out-of-control condition is signaled. This formulation follows that of Lucas (1976), Lucas and Crosier (1982), and Montgomery (1991).

Negative Shifts

If the shift $\delta$ to be detected is negative, the cusum computed for the $t$th subgroup is

$$S_t = \max(0, S_{t-1} - (z_t + k))$$

for $t=1, 2, \ldots, n$, where $S_0=0$, $z_t$ is defined as for two-sided cusum schemes, and the parameter $k$, termed the reference value, is positive. The cusum $S_t$ is referred to as a lower cumulative sum. Since $S_t$ can be written as

$$\max \left( 0, S_{t-1} - \frac{\bar{X}_t - (\mu_0 - k\sigma\bar{X}_i)}{\sigma\bar{X}_i} \right)$$

the sequence $S_t$ cumulates the absolute value of deviations in the subgroup means less than $k$ standard errors from $\mu_0$. If $S_t$ exceeds a positive value $h$ (referred to as the decision interval), a shift or out-of-control condition is signaled.

This formulation follows that of Lucas (1976), Lucas and Crosier (1982), and Montgomery (1991). Note that $S_t$ is always positive and $h$ is always positive, regardless of whether $\delta$ is positive or negative. For schemes designed to detect a negative shift, some authors, including van Dobben de Bruyn (1968) and Wadsworth and others (1986), define a reflected version of $S_t$ for which a shift is signaled when $S_t$ is less than a negative limit.

Headstart Values

Lucas and Crosier (1982) describe the properties of a fast initial response (FIR) feature for cusum schemes in which the initial cusum $S_0$ is set to a “headstart” value.
Average run length calculations given by Lucas and Crosier (1982) show that the FIR feature has little effect when the process is in control and that it leads to a faster response to an initial out-of-control condition than a standard cusum scheme. You can provide headstart value \( S_0 \) with the HEADSTART= option or the variable _HSTART_ in a LIMITS= data set.

### Constant Sample Sizes

When the subgroup sample sizes are constant (=\( n \)), it may be preferable to compute cusums that are scaled in the same units as the data. Refer to Montgomery (1991) and Wadsworth and others (1986). To request this, specify the DATAUNITS option. Cusums are then computed as

\[
S_t = \max(0, S_{t-1} + (\bar{X}_i - (\mu_0 + k\sigma / \sqrt{n})))
\]

for \( \delta > 0 \) and the equation

\[
S_t = \max(0, S_{t-1} - (\bar{X}_i - (\mu_0 - k\sigma / \sqrt{n})))
\]

for \( \delta < 0 \). In either case, a shift is signaled if \( S_t \) exceeds \( h' = h\sigma / \sqrt{n} \). Wadsworth and others (1986) use the symbol \( H \) for \( h' \).

If the subgroup sample sizes are not constant, you can specify a constant nominal sample size \( n \) with the LIMITN= option or the variable _LIMITN_ in a LIMITS= data set. In this case, only those subgroups with sample size \( n \) are analyzed unless you also specify the option ALLN. You can further specify the option NMARKERS to request special symbol markers for points corresponding to sample sizes not equal to \( n \).

### Two-Sided Cusum Schemes

If the cusum scheme is two-sided, the cumulative sum \( S_t \) plotted for the \( t^{th} \) subgroup is

\[
S_t = S_{t-1} + z_t
\]

for \( t=1, 2, \ldots, n \). Here \( S_0=0 \), and the term \( z_t \) is calculated as

\[
z_t = (\bar{X}_t - \mu_0) / (\sigma / \sqrt{n_t})
\]

where \( \bar{X}_t \) is the \( t^{th} \) subgroup average, and \( n_t \) is the \( t^{th} \) subgroup sample size. If the subgroup samples consist of individual measurements \( x_t \), the term \( z_t \) simplifies to

\[
z_t = (x_t - \mu_0) / \sigma
\]

Since the first equation can be rewritten as

\[
S_t = \sum_{i=1}^{t} z_i = \sum_{i=1}^{t} (\bar{X}_i - \mu_0) / \sigma \bar{X}_i
\]
the sequence $S_t$ cumulates standardized deviations of the subgroup averages from the target mean $\mu_0$.

In many applications, the subgroup sample sizes $n_i$ are constant ($n_i = n$), and the equation for $S_t$ can be simplified.

$$S_t = \frac{1}{\sigma \bar{X}} \sum_{i=1}^{t} (\bar{X}_i - \mu_0) = \frac{\sqrt{n}/\sigma}{\sigma} \sum_{i=1}^{t} (\bar{X}_i - \mu_0)$$

In some applications, it may be preferable to compute $S_t$ as

$$S_t = \sum_{i=1}^{t} (\bar{X}_i - \mu_0)$$

which is scaled in the same units as the data. Refer to Montgomery (1991), Wadsworth and others (1986), and \textit{ASQC Glossary and Tables for Statistical Quality Control}. If the subgroup sample sizes are constant ($= n$) and if you specify the \textsc{dataunits} option in the \textsc{xchart} statement, the \textsc{cusum} procedure computes cusums using the final equation above. In this case, the procedure rescales the V-mask parameters $h$ and $k$ to $h' = h\sigma / \sqrt{n}$ and $k' = k\sigma / \sqrt{n}$, respectively. Wadsworth and others (1986) use the symbols $F$ for $k'$ and $H$ for $h'$.

If the subgroup sample sizes are not constant, you can specify a constant nominal sample size $n$ with the \textsc{limitn=} option or with the variable \_\textsc{limitn} in a \textsc{limits=} data set. In this case, only those subgroups with sample size $n$ are analyzed unless you also specify the option \textsc{alln}. You can further specify the option \textsc{nmarkers} to request special symbol markers for points corresponding to sample sizes not equal to $n$.

If the process is in control and the mean $\mu$ is at or near the target $\mu_0$, the points will not exhibit a trend since positive and negative displacements from $\mu_0$ tend to cancel each other. If $\mu$ shifts in the positive direction, the points exhibit an upward trend, and if $\mu$ shifts in the negative direction, the points exhibit a downward trend.

---

**Defining the Decision Interval for a One-Sided Cusum Scheme**

The height of the decision interval is $h$, expressed as a multiple of the standard error of the subgroup mean. You can specify $h$ with the \textsc{h=} option in the \textsc{xchart} statement or with the variable \_\textsc{h} in a \textsc{limits=} data set. The decision interval is displayed as a horizontal line on the cusum chart, as illustrated in Figure 12.12.
Interpreting One-Sided Cusum Charts

A shift or out-of-control condition is signaled at time $t$ if the cusum $S_t$ plotted at time $t$ exceeds the decision interval line.

Defining the V-Mask for a Two-Sided Cusum Scheme

The dimensions of the V-mask can be specified using two distinct sets of two parameters.

- $\theta$, defined as half of the angle formed by the V-mask arms, and $d$, the distance between the origin and the vertex, as shown in Figure 12.13. This parameterization is used by many authors, including Johnson and Leone (1962, 1974) and Montgomery (1991).

- $h$, the vertical distance between the origin and the upper (or lower) V-mask arm, and $k$, the rise (drop) in the lower (upper) arm corresponding to an interval of one subgroup unit on the horizontal axis. You can specify the definition of an interval with the INTERVAL= option. This parameterization is used by Lucas (1976) and Wadsworth and others (1986). Lucas (1976) uses the symbols $h^*$ for $h$ and $k^*$ for $k$, and Wadsworth and others (1986) use the symbol $f$ in place of $k$.

The two parameterizations are related by the equations

$$\theta = \arctan(k/a)$$

$$d = h/k$$
where the aspect ratio \( a \) is the number of units on the vertical axis corresponding to one unit on the horizontal axis. The CUSUM procedure uses the \( h \) and \( k \) parameterization because it eliminates the need for working with aspect ratios, which are dependent on the graphics device. Furthermore, \( h \) and \( k \) are also useful for average run length computations and for parameterizing one-sided cusum schemes.

![Figure 12.13. V-Mask Parameters](image)

You can specify the V-mask in two ways:

- geometrically, by providing \( h \) and \( k \) (or simply \( h \)) with the H= and K= options or with the variables _H_ and _K_ in a LIMITS= data set
- in terms of error probabilities, by providing \( \alpha \) and \( \beta \) (or simply \( \alpha \)) with the ALPHA= and BETA= options or with the variables _ALPHA_ and _BETA_ in a LIMITS= data set. The SIGMAS= option is an alternative to the ALPHA= option, and the variable _SIGMAS_ is an alternative to the variable _ALPHA_ (if the READSIGMAS option is specified).

If you provide \( \alpha \) and \( \beta \), \( h \) and \( k \) are computed using the formulas

\[
h = |\delta|^{-1} \log((1 - \beta)/(\alpha/2))
\]

\[
k = |\delta|/2
\]
If you provide $\alpha$ but not $\beta$, $h$ and $k$ are computed using the formulas

\[ h = -|\delta|^{-1}\log(\alpha/2) \]
\[ k = |\delta|/2 \]

In the preceding equations, the error probability $\alpha$ is divided by two because two-sided deviations from the target mean are detected. Refer to Johnson and Leone (1962, 1974).

**Interpreting Two-Sided Cusum Charts**

The origin of the V-mask is located at the most recently plotted point, as illustrated in Figure 12.13. As additional data are collected and the cumulative sum sequence is updated, the origin is relocated at the newest point. A shift or out-of-control condition is signaled at time $t$ if one or more of the points plotted up to time $t$ cross an arm of the V-mask. An upward shift is signaled by points crossing the lower arm, and a downward shift is signaled by points crossing the upper arm. The time at which the shift occurred corresponds to the time at which a distinct change is observed in the slope of the plotted points.

**Designing a Cusum Scheme**

There are three main methods for designing a cusum scheme: the average run length (ARL) approach, the error probability approach, and the economic design approach.

**Average Run Length (ARL) Approach**

With the ARL approach, the parameters $h$ and $k$ are chosen to yield desired average run lengths when the process is operating at the target mean and when a shift of magnitude $\delta$ has occurred. The average run length is the expected number of samples taken before an out-of-control condition is signaled. Ideally, the ARL should be long when $\mu = \mu_0$ and short when $\mu$ shifts away from $\mu_0$.

The ARL method typically involves the use of a table or nomogram. Refer to Kemp (1961), van Dobben de Bruyn (1968), Goel and Wu (1971), Duncan (1974), Lucas (1976), Montgomery (1991), and Wadsworth and others (1986).

For one-sided charts, average run lengths are tabulated as a function of $h$, $k$, and $\delta$ in Table 12.17 on page 433. No headstart is assumed in this table. For two-sided charts, average run lengths are tabulated as a function of $h$, $k$, and $\delta$ in Table 12.18 on page 434, which is formatted similarly to Table 2 given by Lucas (1976).

The ARLs in Table 12.17 and Table 12.18 were calculated with the DATA step function CUSUMARL, which is described on page 1998. This function uses the method of Goel and Wu (1971). You can use this function to generate more detailed, interpolated versions of the tables or to compute ARLs with headstart values.

It can be shown that the two-sided (V-mask) cusum scheme parameterized by $h$ and $k$ is equivalent to two simultaneously operating one-sided cusum schemes, one that computes an upper cusum and one that computes a lower cusum. Both one-sided schemes use the same parameters $h$ and $k$. 

431
Part 2. The CAPABILITY Procedure

You can specify $h$, $k$, and $\delta$ with the options H=, K=, and DELTA= or with the variables _H_, _K_, and _DELTA_ in a LIMITS= data set. The reference value $k$ is optional, and its default value is $k = \frac{|\delta|}{2}$, referred to as the central reference value.

**Error Probability Approach**

This approach is available only for two-sided cusum schemes. Values of $\alpha$ (the probability of incorrectly signaling the occurrence of a shift) and $\beta$ (the probability of failing to detect a shift) are specified, and $h$ and $k$ are computed from $\alpha$ and $\beta$ as described in “Defining the V-Mask for a Two-Sided Cusum Scheme” on page 429. The error probability approach interprets the cusum as a sequence of reversed sequential probability ratio tests. Refer to Johnson (1961), Johnson and Leone (1962, 1974), van Dobben de Bruyn (1968), Montgomery (1991), and Wadsworth and others (1986).

Although the error probability method is intuitively appealing, the actual error probabilities achieved may not be close to those specified since the V-mask does not provide for an acceptance region. This has been pointed out by various authors, including Johnson (1961) and van Dobben de Bruyn (1968). If you follow this approach, it is recommended that you examine the average run lengths for the cusum scheme (these are tabulated by the TABLESUMMARY option and are saved in OUTLIMITS= data sets).

You can specify $\alpha$ and $\beta$ with the ALPHA= and BETA= options or with the variables _ALPHA_ and _BETA_ in a LIMITS= data set. It is not necessary to specify $\beta$, and the interpretation of $\beta$ is somewhat questionable. The SIGMAS= option is an alternative to the ALPHA= option, and the variable _SIGMAS_ is an alternative to the variable _ALPHA_ (if you specify the READSIGMAS option).

**Economic Design**

The parameters $n$, $h$, and $k$ are chosen so that the long-run average cost of the cusum scheme is minimized. Refer to Chiu (1974), Montgomery (1980), Svoboda (1991), and Ho and Case (1994) for reviews of the literature on economic design. This approach typically requires numerical optimization techniques, which are available in SAS/IML software and in the NLP procedure in SAS/OR software.

You can pass the optimal parameters to the CUSUM procedure as values of the variables _LIMITN_, _H_, and _K_ in a LIMITS= data set.
Table 12.17. Average Run Lengths for One-Sided V-Mask Cusum Charts as a Function of $h$, $k$, and $\delta$.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$\delta$ (shift in mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h$</td>
<td>$k$</td>
</tr>
<tr>
<td>2.50</td>
<td>0.25</td>
</tr>
<tr>
<td>4.00</td>
<td>0.25</td>
</tr>
<tr>
<td>6.00</td>
<td>0.25</td>
</tr>
<tr>
<td>8.00</td>
<td>0.25</td>
</tr>
<tr>
<td>10.00</td>
<td>0.25</td>
</tr>
<tr>
<td>2.00</td>
<td>0.50</td>
</tr>
<tr>
<td>3.00</td>
<td>0.50</td>
</tr>
<tr>
<td>4.00</td>
<td>0.50</td>
</tr>
<tr>
<td>5.00</td>
<td>0.50</td>
</tr>
<tr>
<td>6.00</td>
<td>0.50</td>
</tr>
<tr>
<td>1.50</td>
<td>0.75</td>
</tr>
<tr>
<td>2.00</td>
<td>0.75</td>
</tr>
<tr>
<td>3.00</td>
<td>0.75</td>
</tr>
<tr>
<td>3.75</td>
<td>0.75</td>
</tr>
<tr>
<td>4.50</td>
<td>0.75</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>1.50</td>
<td>1.00</td>
</tr>
<tr>
<td>2.00</td>
<td>1.00</td>
</tr>
<tr>
<td>2.50</td>
<td>1.00</td>
</tr>
<tr>
<td>3.00</td>
<td>1.00</td>
</tr>
<tr>
<td>3.75</td>
<td>1.00</td>
</tr>
<tr>
<td>4.50</td>
<td>1.00</td>
</tr>
<tr>
<td>0.70</td>
<td>1.50</td>
</tr>
<tr>
<td>1.10</td>
<td>1.50</td>
</tr>
<tr>
<td>1.50</td>
<td>1.50</td>
</tr>
<tr>
<td>1.90</td>
<td>1.50</td>
</tr>
<tr>
<td>2.30</td>
<td>1.50</td>
</tr>
</tbody>
</table>
Table 12.18. Average Run Lengths for Two-Sided V-Mask Cusum Charts as a Function of $h$, $k$, and $\delta$.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$\delta$ (shift in mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h$</td>
<td>$k$</td>
</tr>
<tr>
<td>2.50</td>
<td>0.25</td>
</tr>
<tr>
<td>4.00</td>
<td>0.25</td>
</tr>
<tr>
<td>6.00</td>
<td>0.25</td>
</tr>
<tr>
<td>8.00</td>
<td>0.25</td>
</tr>
<tr>
<td>10.00</td>
<td>0.25</td>
</tr>
<tr>
<td>2.00</td>
<td>0.50</td>
</tr>
<tr>
<td>3.00</td>
<td>0.50</td>
</tr>
<tr>
<td>4.00</td>
<td>0.50</td>
</tr>
<tr>
<td>5.00</td>
<td>0.50</td>
</tr>
<tr>
<td>6.00</td>
<td>0.50</td>
</tr>
<tr>
<td>1.50</td>
<td>0.75</td>
</tr>
<tr>
<td>2.25</td>
<td>0.75</td>
</tr>
<tr>
<td>3.00</td>
<td>0.75</td>
</tr>
<tr>
<td>3.75</td>
<td>0.75</td>
</tr>
<tr>
<td>4.50</td>
<td>0.75</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>1.50</td>
<td>1.00</td>
</tr>
<tr>
<td>2.00</td>
<td>1.00</td>
</tr>
<tr>
<td>2.50</td>
<td>1.00</td>
</tr>
<tr>
<td>3.00</td>
<td>1.00</td>
</tr>
<tr>
<td>3.50</td>
<td>1.00</td>
</tr>
<tr>
<td>0.70</td>
<td>1.50</td>
</tr>
<tr>
<td>1.10</td>
<td>1.50</td>
</tr>
<tr>
<td>1.50</td>
<td>1.50</td>
</tr>
<tr>
<td>1.90</td>
<td>1.50</td>
</tr>
<tr>
<td>2.30</td>
<td>1.50</td>
</tr>
</tbody>
</table>
Chapter 12. Details

Cusum Charts Compared with Shewhart Charts

Although cusum charts and Shewhart charts are both used to detect shifts in the process mean, there are important differences in the two methods.

- Each point on a Shewhart chart is based on information for a single subgroup sample or measurement. Each point on a cusum chart is based on information from all samples (measurements) up to and including the current sample (measurement).

- On a Shewhart chart, upper and lower control limits are used to decide whether a point signals an out-of-control condition. On a cusum chart, the limits take the form of a decision interval or a V-mask.

- On a Shewhart chart, the control limits are commonly computed as 3σ limits. On a cusum chart, the limits are determined from average run length specifications, specified error probabilities, or an economic design.

A cusum chart offers several advantages over a Shewhart chart.

- A cusum chart is more efficient for detecting small shifts in the process mean, in particular, shifts of 0.5 to 2 standard deviations from the target mean (refer to Montgomery 1991). Lucas (1976) noted that “a V-mask designed to detect a 1σ shift will detect it about four times as fast as a competing Shewhart chart.”

- Shifts in the process mean are visually easy to detect on a cusum chart since they produce a change in the slope of the plotted points. The point at which the slope changes is the point at which the shift has occurred.

These advantages are not as pronounced if the Shewhart chart is augmented by the tests for special causes described by Nelson (1984, 1985). Also see Chapter 45, “Tests for Special Causes.” Moreover,

- cusum schemes are more complicated to design.

- a cusum chart can be slower to detect large shifts in the process mean.

- it can be difficult to interpret point patterns on a cusum chart since the cusums are correlated.

Methods for Estimating the Standard Deviation

It is recommended practice to provide a stable estimate or standard value for σ with either the SIGMA0= option or the variable _STDDEV_ in a LIMITS= data set. However, if such a value is not available, you can compute an estimate $\hat{\sigma}$ from the data, as described in this section.
Part 2. The CAPABILITY Procedure

This section provides formulas for various methods used to estimate the standard deviation $\sigma$. One method is applicable with individual measurements, and three are applicable with subgrouped data. The methods can be requested with the SMETHOD= option.

**Method for Individual Measurements**

When the cumulative sums are calculated from individual observations $x_1, x_2, \ldots, x_N$ rather than subgroup samples of two or more observations, the CUSUM procedure estimates $\sigma$ as $\sqrt{\sigma^2}$, where

$$\sigma^2 = \frac{1}{2(N-1)} \sum_{i=1}^{N-1} (x_{i+1} - x_i)^2$$

where $N$ is the number of observations. Wetherill (1977) states that the estimate of the variance is biased if the measurements are autocorrelated.

Note that you can compute alternative estimates (for instance, robust estimates or estimates based on variance components models) by analyzing the data with SAS modeling procedures or your own DATA step program. Such estimates can be passed to the CUSUM procedure as values of the variable _STDDEV_ in a LIMITS= data set.

**NOWEIGHT Method for Subgroup Samples**

This method is the default for cusum charts for subgrouped data. The estimate is

$$\hat{\sigma} = \frac{(s_1/c_4(n_1)) + \cdots + (s_N/c_4(n_N))}{N}$$

where $n_i$ is the sample size of the $i$th subgroup, $N$ is the number of subgroups for which $n_i \geq 2$, $s_i$ is the sample standard deviation of the observations $x_{i1}, \ldots, x_{in_i}$ in the $i$th subgroup.

$$s_i = \sqrt{(1/(n_i - 1)) \sum_{j=1}^{n_i} (x_{ij} - \bar{X}_i)^2}$$

and

$$c_4(n_i) = \frac{\Gamma(n_i/2)\sqrt{2/(n_i - 1)}}{\Gamma((n_i - 1)/2)}$$

where $\Gamma(\cdot)$ denotes the gamma function, and $\bar{X}_i$ denotes the $i$th subgroup mean. A subgroup standard deviation $s_i$ is included in the calculation only if $n_i \geq 2$. If the observations are normally distributed, then the expected value of $s_i$ is

$$E(s_i) = c_4(n_i)\sigma$$

Thus, $\hat{\sigma}$ is the unweighted average of $N$ unbiased estimates of $\sigma$. This method is described in the ASTM Manual on Presentation of Data and Control Chart Analysis.
**MVLUE Method for Subgroup Samples**

If you specify SMETHOD=MVLUE, a minimum variance linear unbiased estimate (MVLUE) is computed, as introduced by Burr (1969, 1976). This estimate is a weighted average of unbiased estimates of $\sigma$ of the form

$$s_i/c_4(n_i)$$

where

- $s_i$ is the standard deviation of the $i^{th}$ subgroup.
- $c_4(n_i)$ is the unbiasing factor defined previously.
- $n_i$ is the $i^{th}$ subgroup sample size, $i = 1, 2, \ldots, N$.
- $N$ is the number of subgroups for which $n_i \geq 2$.

The estimate is

$$\hat{\sigma} = \frac{h_1 s_1/c_4(n_1) + \cdots + h_N s_N/c_4(n_N)}{h_1 + \cdots + h_N}$$

where $h_i = c_4^2(n_i)/(1 - c_4^2(n_i))$. A subgroup standard deviation $s_i$ is included in the calculation only if $n_i \geq 2$.

The MVLUE assigns greater weight to estimates of $\sigma$ from subgroups with larger sample sizes and is intended for situations where the subgroup sample sizes vary. If the subgroup sample sizes are constant, the MVLUE reduces to the default estimate (NOWEIGHT).

**RMSDF Method for Subgroup Samples**

If you specify SMETHOD=RMSDF, a weighted root-mean-square estimate is computed:

$$\hat{\sigma} = \frac{\sqrt{(n_1 - 1)s_1^2 + \cdots + (n_N - 1)s_N^2}}{c_4(n)\sqrt{n_1 + \cdots + n_N - N}}$$

where

- $n_i$ is the sample size of the $i^{th}$ subgroup.
- $N$ is the number of subgroups for which $n_i \geq 2$.
- $s_i$ is the sample standard deviation of the $i^{th}$ subgroup.
- $c_4(n_i)$ is the unbiasing factor defined previously.
- $n$ is equal to $(n_1 + \cdots + n_N) - (N - 1)$.

The weights in the root-mean-square expression are the degrees of freedom $n_i - 1$. A subgroup standard deviation $s_i$ is included in the calculation only if $n_i \geq 2$. 
If the unknown standard deviation $\sigma$ is constant across subgroups, the root-mean-square estimate is more efficient than the minimum variance linear unbiased estimate. However, as noted by Burr (1969), “the constancy of $\sigma$ is the very thing under test,” and if $\sigma$ varies across subgroups, the root-mean-square estimate tends to be more inflated than the MVLUE.

Output Data Sets

OUTLIMITS= Data Set

When you save the parameters for the cusum scheme in an OUTLIMITS= data set, the following variables are included:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>ALPHA</em></td>
<td>probability ($\alpha$) of Type 1 error</td>
</tr>
<tr>
<td><em>ARLIN</em></td>
<td>average run length for zero shift</td>
</tr>
<tr>
<td><em>ARLOUT</em></td>
<td>average run length for shift of $\delta$</td>
</tr>
<tr>
<td><em>BETA</em></td>
<td>probability ($\beta$) of Type 2 error</td>
</tr>
<tr>
<td><em>DELTA</em></td>
<td>shift ($\delta$) to be detected</td>
</tr>
<tr>
<td><em>H</em></td>
<td>decision interval $h$ for one-sided scheme; distance $h$ between origin and upper arm V-mask for two-sided scheme</td>
</tr>
<tr>
<td><em>HSTART</em></td>
<td>headstart value</td>
</tr>
<tr>
<td><em>INDEX</em></td>
<td>optional identifier for cusum parameters (if the OUTINDEX= option is specified)</td>
</tr>
<tr>
<td><em>K</em></td>
<td>reference value $k$ for one-sided scheme; slope of lower V-mask arm for two-sided scheme</td>
</tr>
<tr>
<td><em>LIMITN</em></td>
<td>nominal sample size for cusum scheme</td>
</tr>
<tr>
<td><em>MEAN</em></td>
<td>estimated process mean ($\bar{X}$)</td>
</tr>
<tr>
<td><em>MU0</em></td>
<td>target mean $\mu_0$</td>
</tr>
<tr>
<td><em>ORIGIN</em></td>
<td>origin of V-mask</td>
</tr>
<tr>
<td><em>SCHEME</em></td>
<td>type of scheme (<strong>ONESIDED</strong> or <strong>TWOSIDED</strong>)</td>
</tr>
<tr>
<td><em>SIGMAS</em></td>
<td>$2^{1-\alpha/2}$</td>
</tr>
<tr>
<td><em>STDDEV</em></td>
<td>estimated or known standard deviation ($\sigma$ or $\sigma_0$)</td>
</tr>
<tr>
<td><em>SUBGRP</em></td>
<td>subgroup-variable specified in XCHART statement</td>
</tr>
<tr>
<td><em>TYPE</em></td>
<td>type (<strong>ESTIMATE</strong> or <strong>STANDARD</strong>) of <em>STDDEV</em></td>
</tr>
<tr>
<td><em>VAR</em></td>
<td>process specified in XCHART statement</td>
</tr>
</tbody>
</table>

Notes:

1. If the subgroup sample sizes vary, the special missing value $V$ is assigned to the variable _LIMITN_.
2. If a V-mask is specified with SIGMAS=$k$, _ALPHA_ is computed as $\alpha = 2(1 - \Phi(k))$, where $\Phi(\cdot)$ is the standard normal distribution function.
3. If a V-mask is specified with ALPHA=$\alpha$, _SIGMAS_ is computed as $k = \Phi^{-1}(1 - \alpha/2)$, where $\Phi^{-1}$ is the inverse standard normal distribution function.
4. BY variables are saved in the OUTLIMITS= data set.

The OUTLIMITS= data set contains one observation for each process specified in the XCHART statement. For an example, see “Saving Cusum Scheme Parameters” on page 405.

**OUTHISTORY= Data Set**

When you save subgroup summary statistics in an OUTHISTORY= data set, the following variables are included:

- the subgroup-variable
- a subgroup mean variable named by process suffixed with X
- a subgroup sample size variable named by process suffixed with N
- a subgroup standard deviation variable named by process suffixed with S
- a cusum variable named by process suffixed with C

Given a process that is eight characters long, the procedure first shortens the prefix to its first four characters and its last three characters and then adds the suffix. For example, the procedure shortens the process DIAMETER to DIAMTER before adding the suffixes.

Variables containing subgroup summary statistics are created for each process specified in the XCHART statement. For example, consider the following statements:

```sas
proc cusum data=steel limits=stparm;
  xchart (width diameter)*lot / outhistory=summary;
run;
```

The data set SUMMARY would contain nine variables named LOT, WIDTHX, WIDTHS, WIDTHN, WIDTHC, DIAMTERX, DIAMTERS, DIAMTERN, and DIAMTERC.

Additionally, if specified, the following variables are included:

- BY variables
- block-variables
- symbol-variable
- ID variables
- _PHASE_. (if the OUTPHASE= option is specified)

For an example creating an OUTHISTORY= data set, see “Saving Summary Statistics” on page 401.

**OUTTABLE= Data Set**

The OUTTABLE= data set saves subgroup means, subgroup sample sizes, cusums, and cusum limits. The following variables are included:
### Variable Description

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>CUSUM</em></td>
<td>cumulative sum</td>
</tr>
<tr>
<td><em>EXLIM</em></td>
<td>decision interval or V-mask arm exceeded</td>
</tr>
<tr>
<td><em>H</em></td>
<td>decision interval</td>
</tr>
<tr>
<td><em>MASKL</em></td>
<td>lower arm of V-mask</td>
</tr>
<tr>
<td><em>MASKU</em></td>
<td>upper arm of V-mask</td>
</tr>
<tr>
<td>subgroup</td>
<td>values of the subgroup variable</td>
</tr>
<tr>
<td><em>SUBN</em></td>
<td>subgroup sample size</td>
</tr>
<tr>
<td><em>SUBX</em></td>
<td>subgroup mean</td>
</tr>
<tr>
<td><em>SUBS</em></td>
<td>subgroup standard deviation</td>
</tr>
<tr>
<td><em>VAR</em></td>
<td>process specified in XCHART statement</td>
</tr>
</tbody>
</table>

In addition, the following variables are saved if specified:

- BY variables
- block-variables
- ID variables
- _PHASE_ (if the READPHASES= option is specified)
- _TREND_ (if the TRENDVAR= option is specified)
- symbol-variable

Note that the variables _VAR_ and _EXLIM_ are character variables of length eight. The variable _PHASE_ is a character variable of length 16.

### ODS Tables

The following table summarizes the ODS tables that you can request with the XCHART statement.

**Table 12.20. ODS Tables Produced with the XCHART Statement**

<table>
<thead>
<tr>
<th>Table Name</th>
<th>Description</th>
<th>Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>CompCusum</td>
<td>computational form of the cusum scheme</td>
<td>TABLEALL, TABLECOMP</td>
</tr>
<tr>
<td>Parameters</td>
<td>cusum parameters and computed average run lengths</td>
<td>TABLEALL, TABLESUMMARY</td>
</tr>
<tr>
<td>XCHART</td>
<td>cusum chart summary statistics</td>
<td>TABLEALL, TABLECHART, TABLEOUT</td>
</tr>
</tbody>
</table>

### Input Data Sets

**DATA= Data Set**

You can read raw data (measurements) from a DATA= data set specified in the PROC CUSUM statement. Each process specified in the XCHART statement must be a SAS variable in the DATA= data set. The values of this variable are typically measurements of a quality characteristic taken on items in subgroup samples indexed by the values of the subgroup variable. The subgroup-variable specified in the XCHART
statement must also be a SAS variable in the DATA= data set. Other variables that can be read from a DATA= data set include

- `_PHASE_` (if the READPHASES= option is specified)
- `block-variables`
- `symbol-variable`
- `BY variables`
- `ID variables`

Each observation in a DATA= data set should contain a raw measurement for each process and a value for the subgroup variable. If the \(i\)th subgroup contains \(n_i\) items, there should be \(n_i\) consecutive observations for which the value of the subgroup variable is the index of the \(i\)th subgroup. For example, if each of 30 subgroup samples contains five items, the DATA= data set should contain 150 observations.

By default, the CUSUM procedure reads all of the observations in a DATA= data set. However, if the DATA= data set includes the variable `_PHASE_`, you can read selected groups of observations (referred to as `phases`) by specifying the READPHASES= option in the XCHART statement.

For an example of a DATA= data set, see “Creating a V-Mask Cusum Chart from Raw Data” on page 396.

**LIMITS= Data Set**

You can read cusum scheme parameters from a LIMITS= data set specified in the PROC CUSUM statement.* As an alternative to specifying the parameters with options, a LIMITS= data set provides the following advantages: it facilitates reusing a permanently saved set of parameters, reading a distinct set of parameters for each process specified in the XCHART statement, and keeping track of multiple sets of parameters for the same process over time.

The LIMITS= data set can be an OUTLIMITS= data set that was created in a previous run of the CUSUM procedure. Such data sets always contain the variables required for a LIMITS= data set; consequently, this is the easiest way to construct a LIMITS= data set.

A LIMITS= data set can also be created directly using a DATA step. The variables required for the data set depend on the type of cusum scheme and how the scheme is specified. The following restrictions apply:

- The variables `_VAR_`, `_SUBGRP_`, `_DELTA_`, and `_MU0_` are required.
- For a one-sided cusum scheme, `_H_` is required.
- For a two-sided cusum scheme, one of the following three variables is required: `_ALPHA_`, `_H_`, or `_SIGMAS_`.
- If you plan to use the READINDEX= option, the variable `_INDEX_` is required; otherwise, it is optional.

*If you are using Release 6.09 or an earlier release, you must also specify the READLIMITS or READINDEX= option in the XCHART statement.
Part 2. The CAPABILITY Procedure

- For a one-sided scheme, the variable _SCHEME_ is required; otherwise, it is optional.
- If you want to provide a value for the process standard deviation \( \sigma \), the variable _STDDEV_ is required; otherwise, it is optional.

Variable names in a LIMITS= data set are predefined; the procedure reads only variables with these predefined names. With the exception of BY variables, all names start and end with an underscore. In addition, note the following:

- The variables _VAR_, _SUBGRP_, _TYPE_, and _SCHEME_ must be character variables of length eight. The variable _INDEX_ must be a character variable of length 16.
- The variable _TYPE_ is a bookkeeping variable that uses the values ESTIMATE and STANDARD to record whether the value of _STDDEV_ represents an estimate or standard (known) value.
- BY variables are required if specified with a BY statement.

For an example of reading control limit information from a LIMITS= data set, see “Reading Cusum Scheme Parameters” on page 408.

**HISTORY= Data Set**

Instead of reading raw data from a DATA= data set, you can read subgroup summary statistics from a HISTORY= data set specified in the PROC CUSUM statement. This allows you to reuse OUTHISTORY= data sets that have been created in previous runs of the CUSUM, MACONTROL, or SHEWHART procedures or to read output data sets created with SAS summarization procedures such as PROC MEANS. A HISTORY= data set must contain the following variables:

- subgroup-variable
- subgroup mean variable for each process
- subgroup standard deviation variable for each process
- subgroup sample size variable for each process

The names of the subgroup mean, subgroup standard deviation, and subgroup sample size variables must be the process concatenated with the special suffix characters X, S, and N respectively.

For example, consider the following statements:

```sas
proc cusum history=steel limits=stlparm;
  xchart (weight yldstren)*batch;
run;
```

The data set STEEL must contain the variables BATCH, WEIGHTX, WEIGHTS, WEIGHTN, YLDSRENX, YLDSRENS, and YLDSRENN.

Note that if you specify a process that is eight characters long, the procedure uses a prefix formed from the first four characters and the last three characters of the process.

Other variables that can be read from a HISTORY= data set include
Chapter 12. Details

- `_PHASE_` (if the READPHASES= option is specified)
- `block-variables`
- `symbol-variable`
- `BY` variables
- `ID` variables

By default, the CUSUM procedure reads all of the observations in a HISTORY= data set. However, if the HISTORY= data set includes the variable `_PHASE_`, you can read selected groups of observations (referred to as phases) by specifying the READPHASES= option.

For an example of reading summary information from a HISTORY= data set, see “Creating a V-Mask Cusum Chart from Subgroup Summary Data” on page 399.

### Missing Values

An observation read from a DATA= or HISTORY= data set is not analyzed if the value of the subgroup variable is missing. For a particular process variable, an observation read from a DATA= data set is not analyzed if the value of the process variable is missing. Missing values of process variables generally lead to unequal subgroup sample sizes. For a particular process variable, an observation read from a HISTORY= data set is not analyzed if the values of any of the corresponding summary variables are missing.
Part 2. The CAPABILITY Procedure

Examples

This section provides advanced examples of the XCHART statement.

Example 12.1. Cusum and Standard Deviation Charts

When you are working with subgrouped data, it can be helpful to accompany a cusum chart for means with a Shewhart $s$ chart for monitoring the variability of the process. This example creates this combination for the variable WEIGHT in the data set OIL (see “Creating a V-Mask Cusum Chart from Raw Data” on page 396).

The first step is to create a one-sided cusum chart for means that detects a shift of one standard error ($\delta = -1$) below the target mean.

```sas
proc cusum data=oil;
  xchart weight*hour /
    nochart /* suppress display */
    mu0 = 8.100 /* target mean */
    sigma0 = 0.050 /* known standard deviation */
    delta = -1 /* shift to be detected */
    h = 3 /* decision interval */
    k = 0.5 /* reference value */
    scheme=onesided /* one-sided scheme */
    outtable = tabcusum /* save the results */
      ( drop = _var_ _subn_ _subx_ _exlim_
      rename = ( _cusum_ = _subx_ _h_ = _uclx_ ) )
  ;
run;
```

The results are saved in an OUTTABLE= data set named TABCUSUM. The cusum variable (CUSUM_) and the decision interval variable (H_) are renamed to SUBX_ and UCLX_ so that they can later be read by the SHEWHART procedure.

The next step is to construct a Shewhart $\bar{X}$ and $s$ chart for WEIGHT and save the results in a data set named TABXSCHT.

```sas
proc shewhart data=oil;
  xschart weight*hour /
    nochart /* suppress display */
    outtable = tabxscht /* save the results */
      ( drop = _subx_ _uclx_ );
run;
```

Note that the variables SUBX_ and UCLX_ are dropped from TABXSCHT.

The third step is to merge the data sets TABCUSUM and TABXSCHT.

```sas
data taball;
```
merge tabxscht tabcusum; by hour;
   _mean_ = _uclx_ * 0.5;
   _lclx_ = 0.0;
run;

proc print;
run;

The variable _LCLX_ is assigned the role of the lower limit for the cusums, and
the variable _MEAN_ is assigned a dummy value. Now, TABALL, which is listed
in Output 12.1.1, has the structure required for a TABLE= data set used with the
XSCHART statement in the SHEWHART procedure (see “TABLE= Data Set” on
page 1698 in Chapter 41, “XSCHART Statement.”).

Output 12.1.1.  Listing of the Data Set TABALL

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The final step is to use the SHEWHART procedure to read TABALL as a TABLE=
data set and to display the cusum and s charts.

```
title 'Cusum Chart for Mean and s chart';
symbol v=dot;
proc shewhart table=taball;
xchart weight * hour /
   ucllabel = 'h=3.0'
   nolimitslegend
   noctl
   split = '/'
   nolegend;
   label _subx_ = 'Lower Cusum/Std Dev';
run;
```

The central line for the primary (cusum) chart is suppressed with the NOCTL option,
and the default $3\sigma$ Limits legend is suppressed with the NOLIMITLEGEND option.
The charts are shown in Output 12.1.2.
The process variability is stable, and there is no signal of a downward shift in the process mean.
Chapter 12. Examples

Example 12.2. Upper and Lower One-Sided Cusum Charts

This example illustrates how to combine upper and lower one-sided cusum charts for means in the same display. As in the preceding example, OUTTABLE= data sets are created with the CUSUM procedure, and the display is created with the SHEWHART procedure.

The following statements analyze the variable WEIGHT in the data set OIL (see "Creating a V-Mask Cusum Chart from Raw Data" on page 396). The first step is to compute and save upper and lower one-sided cusums for shifts of one standard error in the positive and negative directions.

```
proc cusum data=oil;
  xchart weight*hour /
    nochart
    mu0 = 8.100 /* target mean */
    sigma0 = 0.050 /* known standard deviation */
    delta = 1 /* shift to be detected */
    h = 3 /* decision interval */
    k = 0.5 /* reference value */
    scheme=onesided
    outtable = tabupper
      ( drop = _subx_ _subs_ _exlim_
        rename = ( _cusum_ = _subx_ _h_ = _uclx_ ) )
  ;
  xchart weight*hour /
    nochart
    mu0 = 8.100 /* target mean */
    sigma0 = 0.050 /* known standard deviation */
    delta = -1 /* shift to be detected */
    h = 3 /* decision interval */
    k = 0.5 /* reference value */
    scheme=onesided
    outtable = tablower
      ( drop = _var_ _subn_ _subx_ _subs_ _exlim_
        rename = ( _cusum_ = _subs_ _h_ = _ucls_ ) )
  ;
run;
```

Next, the OUTTABLE= data sets are merged.

```
data tabboth;
  merge tabupper tablower; by hour;
  _mean_ = _uclx_ * 0.5;
  _s_ = _ucls_ * 0.5;
  _lclx_ = 0.0;
  _lcls_ = 0.0;
run;
```

The variables _LCLX_ and _UCLX_ are assigned lower limits of zero for the cusums, and the variables _MEAN_ and _S_ are assigned dummy values.
TABBOTH has the structure required for a TABLE= data set used with the XSCHART statement in the SHEWHART procedure (see “TABLE= Data Set” on page 1698 in Chapter 41, “XSCHART Statement.”).

The final step is to read TABBOTH as a TABLE= data set with the SHEWHART procedure.

```
    title 'Upper and Lower Cusums';
    symbol v=dot;
    proc shewhart table=tabboth;
      xchart weight*hour /
        nolimitslegend
        ucllabel = 'h=3.0'
        ucllabel2 = 'h=3.0'
        ypct1 = 50
        vref = 1 2
        vref2 = 1 2
        lvref = 2
        noctl
        noctl2
        split = '/'
        nolegend;
      label _subx_ = 'Upper Sum/Lower Sum';
    run;
```

The combined display is shown in Output 12.2.1. There is no evidence of a shift in either direction.

**Output 12.2.1.** Upper and Lower One-Sided Cusum Charts
Chapter 12. Examples

Example 12.3. Combined Shewhart–Cusum Scheme

Lucas (1982) introduced a combined Shewhart-cusum scheme that is illustrated in this example. Also refer to Ryan (1989). The data set used here is CANS, which is created in “Creating a One-Sided Cusum Chart with a Decision Interval” on page 402.

The first step is to compute and save one-sided cusums to detect a positive shift from the mean.

```sas
proc cusum data=cans;
  xchart weight*hour /
    nochart
    mu0 = 8.100 /* target mean */
    sigma0 = 0.050 /* known standard deviation */
    delta = 1 /* shift to be detected */
    h = 3 /* decision interval */
    k = 0.5 /* reference value */
    scheme = onesided /* onesided scheme */
    outtable = tabcus /* save the result */
      ( drop = _var_ _subn_ _exlim_
        rename = ( _cusum_ = _subr_ _h_ = _uclr_ )
      );
  run;
```

Note that a headstart value is not used here but can be specified with the HSTART= option. Several variables in the OUTTABLE= data set are dropped or renamed so that they can later be read by the SHEWHART procedure.

The next step is to construct a Shewhart chart (not shown) for individual measurements.

```sas
proc shewhart data=cans;
  irchart weight*hour /
    nochart
    mu0 = 8.100
    sigma0 = 0.050
    outtable = tabx
      ( drop = _subr_ _lclr_ _r_ _uclr_ );
  id comment;
  run;
```

By default, $3\sigma$ limits are computed, but the multiple of $\sigma$ can be modified with the SIGMAS= option. As before, the results are saved in an OUTTABLE= data set.

Next, the two OUTTABLE= data sets are merged.

```sas
data combine;
  merge tabx tabcus; by hour;
  _lclr_ = 0.0;
  _r_ = 0.5 * _uclr_;
  run;
```

The data set COMBINE has the structure required for a TABLE= data set used with the IRCHART statement in the SHEWHART procedure (see “TABLE= Data Set” on page 1247 in Chapter 31, “IRCHART Statement.”).
Finally, the combined scheme is displayed with the SHEWHART procedure.

```sas
symbol v=dot;
title 'Combined Shewhart-Cusum Analysis for Weight';
proc shewhart table=combine;
  irchart weight*hour /
    ypctl1 = 50
    noctl2
    uclabel2 = 'h=0.3'
    outlabel = ( comment )
    outlabel2 = ( comment )
    cframelab = ligr
    split = '/';
    label _subi_ = 'Shewhart/Cusum';
run;
```

The chart is shown in Output 12.3.1.

**Output 12.3.1. Combined Shewhart–Cusum Scheme**

Note that a shift is detected by the cusum scheme but not by the Shewhart chart. The point exceeding the decision interval is labeled with the variable COMMENT created in the data set CANS.

Lucas (1982) tabulates average run lengths for combined Shewhart-cusum schemes. The scheme used here has an ARL of 111.1 for \( \delta = 0 \) and an ARL of 6.322 for \( \delta = 1 \).