

Chapter 12

XCHART Statement

Chapter Table of Contents

OVERVIEW	361
GETTING STARTED	362
Creating a V-Mask Cusum Chart from Raw Data	362
Creating a V-Mask Cusum Chart from Subgroup Summary Data	365
Saving Summary Statistics	367
Creating a One-Sided Cusum Chart with a Decision Interval	368
Saving Cusum Scheme Parameters	371
Reading Cusum Scheme Parameters	373
SYNTAX	375
Summary of Options	376
Dictionary of Special Options	385
DETAILS	391
Basic Notation for Cusum Charts	391
Formulas for Cumulative Sums	392
Defining the Decision Interval for a One-Sided Cusum Scheme	394
Defining the V-Mask for a Two-Sided Cusum Scheme	395
Designing a Cusum Scheme	397
Cusum Charts Compared with Shewhart Charts	401
Methods for Estimating the Standard Deviation	401
Output Data Sets	404
ODS Tables	406
Input Data Sets	406
Missing Values	409
EXAMPLES	410
Example 12.1 Cusum and Standard Deviation Charts	410
Example 12.2 Upper and Lower One-Sided Cusum Charts	413
Example 12.3 Combined Shewhart–Cusum Scheme	414

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 δ to be detected
- specify the target mean μ_0
- specify a known (standard) value σ_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 375, and advanced examples are given in the “Examples” section on page 410.

Creating a V-Mask Cusum Chart from Raw Data

See CUSTWOS1
in the SAS/QC
Sample Library

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.

```
data oil;
  label hour = 'Hour';
  input hour @;
  do i=1 to 4;
    input weight @;
    output;
  end;
  drop i;
  datalines;
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.

hour	weight
1	8.024
1	8.135
1	8.151
1	8.065
2	7.971
2	8.165
2	8.077
2	8.157
3	8.125
3	8.031
3	8.198
3	8.050
4	8.123
.	.
.	.
.	.
12	8.018

Figure 12.1. Partial Listing of the Data Set OIL

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:

```

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.

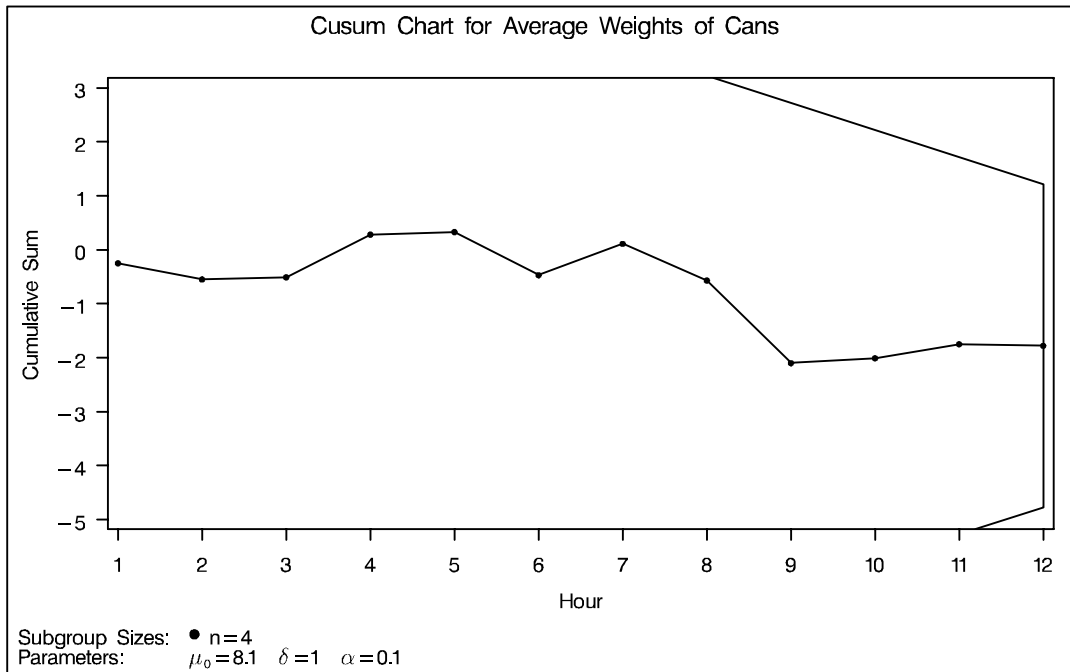


Figure 12.2. Two-Sided Cusum Chart with V-Mask

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 393.

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.

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.

See CUSTWOS1
in the SAS/QC
Sample Library

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 362):

```

data oilstat;
  input hour weightx weights weightn;
  datalines;
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.

Obs	hour	weightx	weights	weightn
1	1	8.0938	0.0596	4
2	2	8.0925	0.0902	4
3	3	8.1010	0.0763	4
4	4	8.1198	0.0256	4
5	5	8.1013	0.0265	4
6	6	8.0800	0.0756	4
7	7	8.1145	0.0372	4
8	8	8.0830	0.0593	4
9	9	8.0618	0.0057	4
10	10	8.1023	0.0465	4
11	11	8.1065	0.0405	4
12	12	8.0993	0.0561	4

Figure 12.3. Listing of the Data Set OILSTAT

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:

```
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 406 and “HISTORY= Data Set” on page 408.

Saving Summary Statistics

In this example, the CUSUM procedure is used to save summary statistics and cumsums 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 362) and create a summary data set named OILHIST:

See CUSTWOS1
in the SAS/QC
Sample Library

```
proc cusum data=oil;
  xchart weight*hour /
  nochart
  outhistory = oilhist
  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;

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.

Obs	hour	weightX	weights	weightC	weightN
1	1	8.0938	0.0596	-.2500	4
2	2	8.0925	0.0902	-.5500	4
3	3	8.1010	0.0763	-.5100	4
4	4	8.1198	0.0256	0.2800	4
5	5	8.1013	0.0265	0.3300	4
6	6	8.0800	0.0756	-.4700	4
7	7	8.1145	0.0372	0.1100	4
8	8	8.0830	0.0593	-.5700	4
9	9	8.0618	0.0057	-2.100	4
10	10	8.1023	0.0465	-2.010	4
11	11	8.1065	0.0405	-1.750	4
12	12	8.0993	0.0561	-1.780	4

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* WEIGHT specified in the 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 405.

Creating a One-Sided Cusum Chart with a Decision Interval

See CUSONES1
in the SAS/QC
Sample Library

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:*

```
data cans;
  length comment $16;
  label hour = 'Hour';
  input hour weight comment $16. ;
  datalines;
1  8.024
2  7.971
3  8.125
4  8.123
5  8.068
6  8.177  Pump Adjusted
7  8.229  Pump Adjusted
8  8.072
9  8.066
10 8.089
11 8.058
12 8.147
13 8.141
14 8.047
15 8.125
;
```

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 12.18 on page 399 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:

*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.

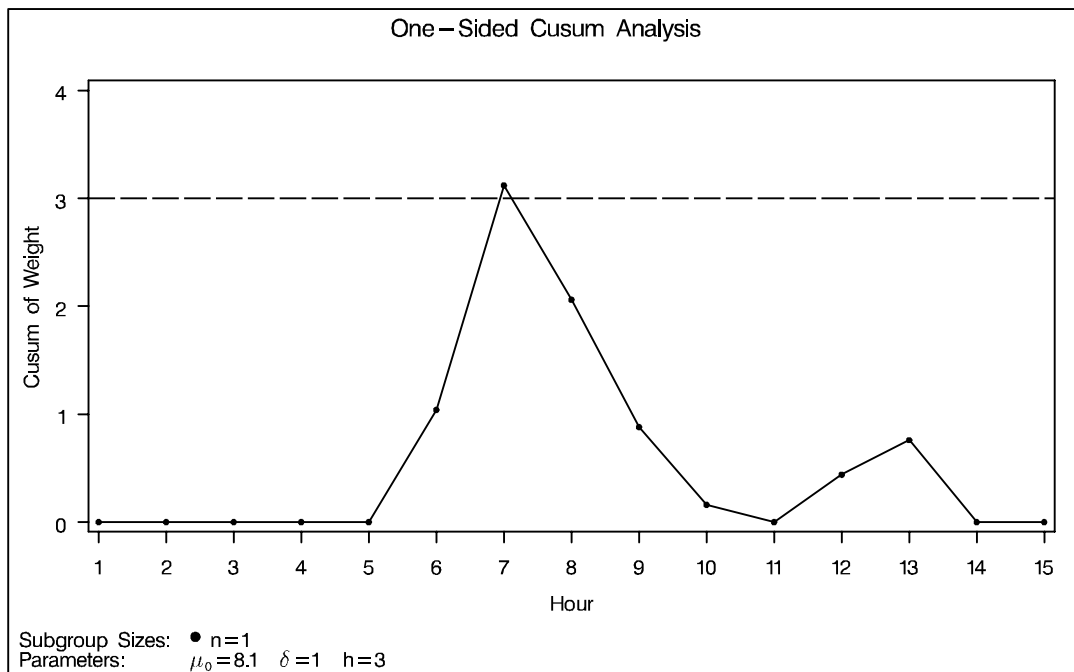


Figure 12.5. One-Sided Cusum Chart with Decision Interval

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 392.

Part 2. The CAPABILITY Procedure

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	
Cusum Parameters	
Process Variable	weight (Cusum of Weight)
Subgroup Variable	hour (Hour)
Scheme	One-Sided
Target Mean (μ_0)	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

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

Cumulative Sum Chart Summary for weight					
hour	Subgroup Sample Size	Individual Value	Cusum	Decision Interval	Decision Interval Exceeded
1	1	8.0240000	0.0000000	3.0000	
2	1	7.9710000	0.0000000	3.0000	
3	1	8.1250000	0.0000000	3.0000	
4	1	8.1230000	0.0000000	3.0000	
5	1	8.0680000	0.0000000	3.0000	
6	1	8.1770000	1.0400000	3.0000	
7	1	8.2290000	3.1200000	3.0000	Upper
8	1	8.0720000	2.0600000	3.0000	
9	1	8.0660000	0.8800000	3.0000	
10	1	8.0890000	0.1600000	3.0000	
11	1	8.0580000	0.0000000	3.0000	
12	1	8.1470000	0.4400000	3.0000	
13	1	8.1410000	0.7600000	3.0000	
14	1	8.0470000	0.0000000	3.0000	
15	1	8.1250000	0.0000000	3.0000	

Figure 12.7. Tabulation of One-Sided Chart

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

Computational Cumulative Sum for weight				
hour	Subgroup Sample Size	Individual Value	Upper Cusum	Number of Consecutive Upper Sums > 0
1	1	8.0240000	0.0000000	0
2	1	7.9710000	0.0000000	0
3	1	8.1250000	0.0000000	0
4	1	8.1230000	0.0000000	0
5	1	8.0680000	0.0000000	0
6	1	8.1770000	1.0400000	1
7	1	8.2290000	3.1200000	2
8	1	8.0720000	2.0600000	3
9	1	8.0660000	0.8800000	4
10	1	8.0890000	0.1600000	5
11	1	8.0580000	0.0000000	0
12	1	8.1470000	0.4400000	1
13	1	8.1410000	0.7600000	2
14	1	8.0470000	0.0000000	0
15	1	8.1250000	0.0000000	0

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

$$\begin{aligned}
 & \mu_0 + \sigma_0(N_7k + S_7)/(N_7\sqrt{n}) \\
 & = 8.10 + 0.05(2(0.5) + 3.12)/2 \\
 & = 8.203 \text{ ounces}
 \end{aligned}$$

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.

See CUSONES1
in the SAS/QC
Sample Library

```

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;

```

Part 2. The CAPABILITY Procedure

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. *

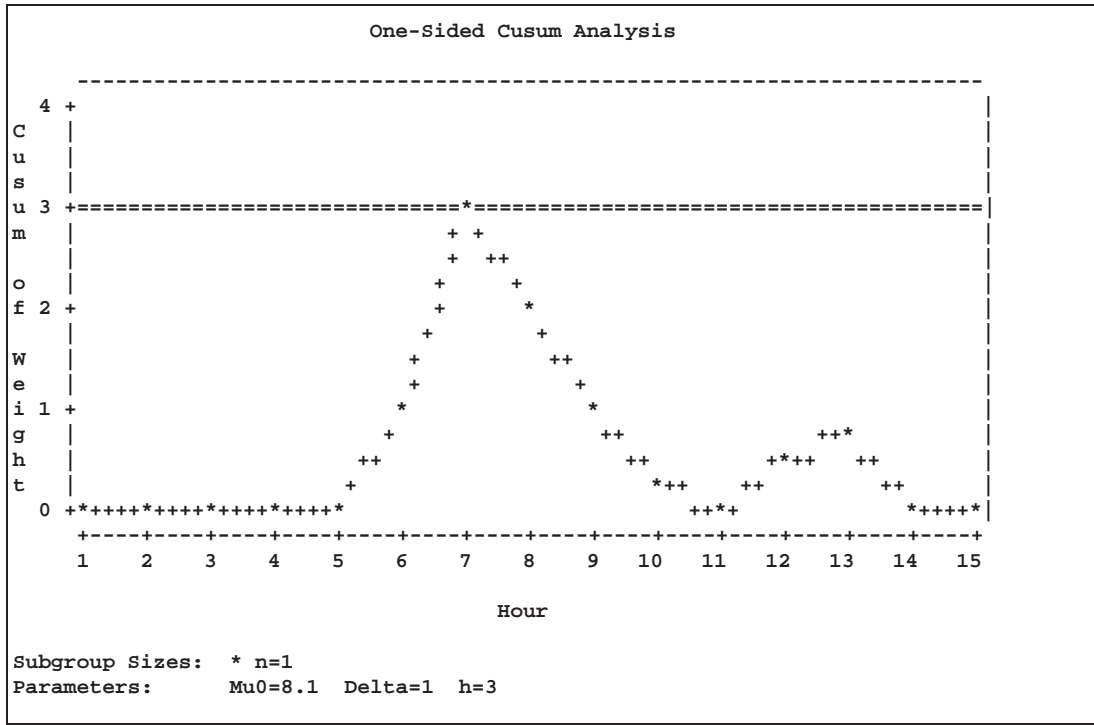


Figure 12.9. One-Sided Cusum Scheme with Decision Interval
The OUTLIMITS= data set is listed in Figure 12.10.

	S		L		S		S		A
	U		I		C		D		T
	B	T	M		H		E	M	D
	V	G	Y		I		E	M	L
O	A	R	P	T			M	U	T
b	R	P	E	N	H	K		E	O
s									
	1	weight	hour	STANDARD	1	3	0.5	ONESIDED	8.1
					1	8.09747	0.05	117.596	6.40391

Figure 12.10. Listing of the OUTLIMITS= Data Set CUSPARM

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 $\delta = 0$ and for $\delta = 1$.

*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.

The variables `_VAR_` and `_SUBGRP_` save the *process* and *subgroup-variable*. The variable `_TYPE_` is a bookkeeping variable that indicates whether the value of `_STDDEV_` is an estimate or a standard value.

For more information, see “OUTLIMITS= Data Set” on page 404.

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:

See CUSONES1
in the SAS/QC
Sample Library

```
data cans2;
  length pump $ 8;
  label hour = 'Hour';
  input hour weight pump $ 8. ;
  datalines;
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:

```
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

*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.

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 12.7 on page 379.

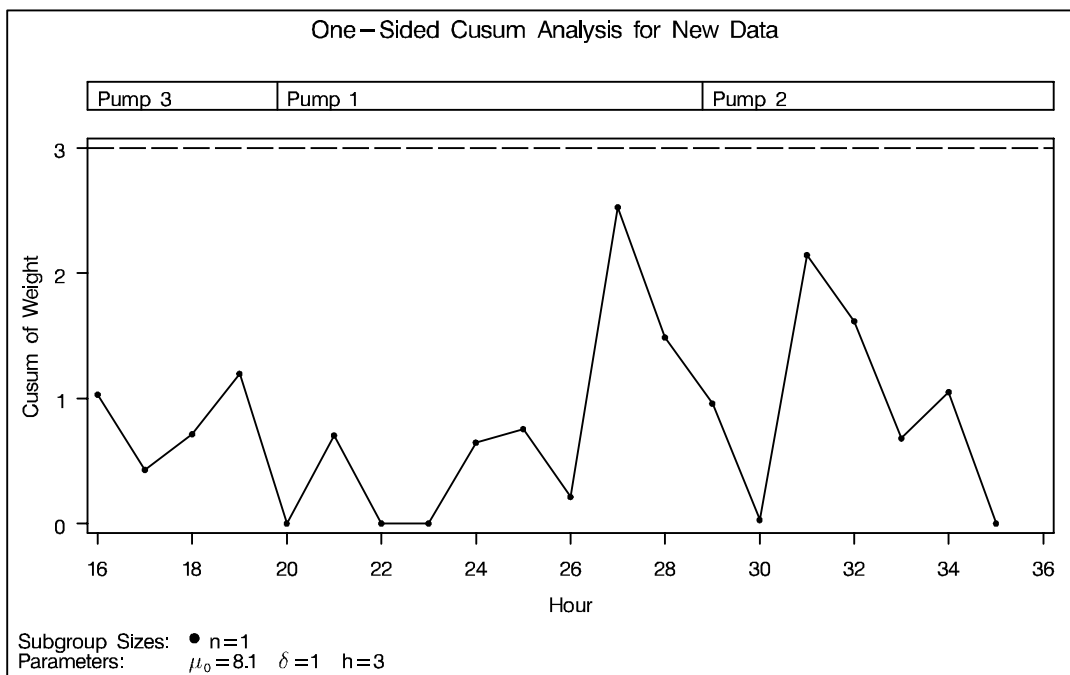


Figure 12.11. Cusum Chart with Decision Interval for New Data

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 407 for details concerning the variables that you must provide.

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=ONESIDED 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=ONESIDED MU0=target DELTA=shift H=h < options
>;
```

Note that the options SCHEME=ONESIDED, 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 362.
- 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 example, see “Creating a V-Mask Cusum Chart from Subgroup Summary Data” on page 365.

Part 2. The CAPABILITY Procedure

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 362 and 365, 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.

symbol-variable

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 SYMBOL n statements.

character

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 385. Options that are common to both the CUSUM and SHEWHART procedures are listed in Table 12.14 to Table 12.17. They are described in detail beginning on page 1613 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*

DELTA= <i>value</i>	specifies shift to be detected as a multiple of standard error
H= <i>value</i>	specifies decision interval h ($h > 0$) as a multiple of standard error
HEADSTART= <i>value</i>	specifies headstart value S_0 as a multiple of standard error
K= <i>value</i>	specifies reference value k ($k > 0$)
LIMITN= <i>n</i>	specifies fixed nominal sample size for cusum scheme
LIMITN=VARYING	specifies that cusums are to be computed for all subgroups regardless of sample size
MU0= <i>value</i>	specifies target μ_0 for mean
NOREADLIMITS	specifies that cusum parameters are not to be read from LIMITS= data set (Release 6.10 and later releases)
READINDEX='value'	reads cusum scheme parameters from a LIMITS= data set
READLIMITS	specifies that cusum parameters are to be read from LIMITS= data set (Release 6.09 and earlier releases)
SCHEME=ONESIDED	specifies a one-sided scheme
SHIFT= <i>value</i>	specifies shift to be detected in data units
SIGMA0= <i>value</i>	specifies standard (known) value σ_0 for process standard deviation

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

ALPHA= <i>value</i>	specifies probability of Type 1 error
BETA= <i>value</i>	specifies probability of Type 2 error
H= <i>value</i>	specifies vertical distance between V-mask origin and upper (or lower) arm
K= <i>value</i>	specifies slope of lower arm of V-mask
LIMITN= <i>n</i>	specifies fixed nominal sample size for cusum scheme
LIMITN=VARYING	specifies that cusums are to be computed for all subgroups regardless of sample size
NOREADLIMITS	specifies that cusum parameters are not to be read from LIMITS= data set (Release 6.10 and later releases)
READINDEX='value'	reads cusum scheme parameters from a LIMITS= data set
READLIMITS	specifies that cusum parameters are to be read from LIMITS= data set (Release 6.09 and earlier releases)
READSIGMAS	reads <code>_SIGMAS_</code> instead of <code>_ALPHA_</code> from LIMITS= data set when both variables are available
SIGMAS= <i>value</i>	specifies probability of Type 1 error as probability that standard normally distributed variable exceeds <i>value</i> in absolute value

*The options in these tables are described in the “Dictionary of Special Options” section on page 385.

Table 12.3. Options for Estimating Process Standard Deviation *

SMETHOD= <i>keyword</i>	specifies method for estimating process standard deviation σ
TYPE= <i>keyword</i>	identifies whether _STDDEV_ in OUTLIMITS= data set is an estimate or standard, and specifies value of _TYPE_ in OUTLIMITS= data set

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

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

Table 12.5. Tabulation Options*

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

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

*The options in these tables are described in the “Dictionary of Special Options” section on page 385.

Table 12.6. Axis and Axis Label Options[†]

CAXIS= <i>color</i>	specifies color for axis lines and tick marks
CFRAME= <i>color</i> (<i>color-list</i>)	specifies fill color(s) for frame for plot area
CTEXT= <i>color</i>	specifies color for tick mark values and axis labels
HAXIS= <i>values</i> AXIS <i>n</i>	specifies tick mark values for horizontal axis
HEIGHT= <i>value</i>	specifies height of axis label and axis legend text
HMINOR= <i>n</i>	specifies number of minor tick marks between major tick marks on horizontal axis
HOFFSET= <i>value</i>	specifies length of offset at both ends of horizontal axis
INTSTART= <i>value</i>	specifies first major tick mark value for numeric horizontal axis
NOHLABEL	suppresses label for horizontal axis
NOVANGLE	requests vertical axis labels that are strung out vertically
SKIPLABELS= <i>n</i>	specifies thinning factor for tick mark labels on horizontal axis
SPLIT='character'	specifies splitting character for axis labels
TURNHLABELS	requests horizontal axis labels that are strung out vertically
VAXIS= <i>value-list</i> <i>name</i>	specifies major tick mark values for vertical axis of cusum chart
VAXIS2= <i>value-list</i> <i>name</i>	specifies major tick mark values for vertical axis of trend chart
VMINOR= <i>n</i>	specifies number of minor tick marks between major tick marks on vertical axis
VOFFSET= <i>value</i>	specifies length of offset at both ends of vertical axis
WAXIS= <i>n</i>	specifies width of axis lines

Table 12.7. Block Variable Legend Options[†]

BLOCKLABELPOS= <i>keyword</i>	specifies position of label for <i>block-variable</i> legend
BLOCKLABTYPE= <i>value</i> <i>keyword</i>	specifies text size of <i>block-variable</i> legend
BLOCKPOS= <i>n</i>	specifies vertical position of <i>block-variable</i> legend
BLOCKREP	repeats identical consecutive labels in <i>block-variable</i> legend
CBLOCKLAB= <i>color</i>	specifies color for filling background in <i>block-variable</i> legend
CBLOCKVAR= <i>variable</i> (<i>variables</i>)	specifies one or more variables whose values are colors for filling background of <i>block-variable</i> legend

[†]The options in these tables are described in Chapter 46, "Dictionary of Options," of Part 9, "The SHEWHART Procedure."

Table 12.8. Reference Line Options[†]

CHREF= <i>color</i>	specifies color for HREF= and HREF2= lines
CVREF= <i>color</i>	specifies color for VREF= and VREF2= lines
HREF= <i>values</i> <i>SAS-data-set</i>	specifies reference lines perpendicular to horizontal axis on cusum chart
HREF2= <i>values</i> <i>SAS-data-set</i>	specifies reference lines perpendicular to horizontal axis on trend chart
HREFCHAR= <i>'character'</i>	specifies line character for HREF= and HREF2= lines
HREFDATA= <i>SAS-data-set</i>	specifies position of reference lines perpendicular to horizontal axis on cusum chart
HREF2DATA= <i>SAS-data-set</i>	specifies position of reference lines perpendicular to horizontal axis on trend chart
HREFLABELS= <i>'label1'...'labeln'</i>	specifies labels for HREF= lines
HREF2LABELS= <i>'label1'...'labeln'</i>	specifies labels for HREF2= lines
HREFLABPOS= <i>n</i>	specifies position of HREFLABELS= and HREF2LABELS= labels
LHREF= <i>linetype</i>	specifies line type for HREF= and HREF2= lines
LVREF= <i>linetype</i>	specifies line type for VREF= and VREF2= lines
NOBYREF	specifies that reference line information in a data set is to be applied uniformly to charts created for all BY groups
VREF= <i>values</i> <i>SAS-data-set</i>	specifies reference lines perpendicular to vertical axis on cusum chart
VREF2= <i>values</i> <i>SAS-data-set</i>	specifies reference lines perpendicular to vertical axis on trend chart
VREFCHAR= <i>'character'</i>	specifies line character for VREF= and VREF2= lines
VREFLABELS= <i>'label1'...'labeln'</i>	specifies labels for VREF= lines
VREF2LABELS= <i>'label1'...'labeln'</i>	specifies labels for VREF2= lines
VREFLABPOS= <i>n</i>	specifies position of VREFLABELS= and VREF2LABELS= labels

[†]The options in these tables are described in Chapter 46, "Dictionary of Options," of Part 9, "The SHEWHART Procedure."

Table 12.9. Options for Plotting and Labeling Points[†]

ALLLABEL=VALUE (<i>variable</i>)	labels every point on cusum chart
ALLLABEL2=VALUE (<i>variable</i>)	labels every point on trend chart
CCONNECT= <i>color</i>	specifies color for line segments that connect points on chart
CFRAMELAB= <i>color</i>	specifies fill color for frame around labeled points
CONNECTCHAR= ' <i>character</i> '	specifies character used to form line segments that connect points on chart
COUT= <i>color</i>	specifies color for connecting line segments that lie above the decision interval or outside the V-mask
COUTFILL= <i>color</i>	specifies color for areas between connected points and decision interval or V-mask
NOCONNECT	suppresses line segments that connect points on chart
NOTRENDCONNECT	suppresses line segments that connect points on trend chart
OUTLABEL=VALUE (<i>variable</i>)	labels points exceeding decision interval on one-sided chart
SYMBOLCHARS= ' <i>characters</i> '	specifies characters indicating <i>symbol-variable</i>
SYMBOLLEGEND= NONE <i>name</i>	specifies LEGEND statement for levels of <i>symbol-variable</i>
SYMBOLORDER= <i>keyword</i>	specifies order in which symbols are assigned for levels of <i>symbol-variable</i>
TURNALL	turns point labels so that they are strung out vertically

Table 12.10. Phase Options[†]

CPHASELEG= <i>color</i>	specifies text color for <i>phase</i> legend
OUTPHASE='string'	specifies value of <code>_PHASE_</code> in OUTHISTORY= data set
PHASEBREAK	disconnects last point in a <i>phase</i> from first point in next <i>phase</i>
PHASELABTYPE= <i>value</i> <i>keyword</i>	specifies text size of <i>phase</i> legend
PHASELEGEND	displays <i>phase</i> labels in a legend across top of chart
PHASEREF	delineates <i>phases</i> with vertical reference lines
READPHASES= ALL ' <i>label1</i> '...' <i>labeln</i> '	specifies <i>phases</i> to be read from input data set

[†]The options in these tables are described in Chapter 46, "Dictionary of Options," of Part 9, "The SHEWHART Procedure."

Table 12.11. Grid Options[†]

ENDGRID	adds grid after last plotted point
GRID	adds grid to chart
LENDGRID= <i>linetype</i>	specifies line type for grid requested with the ENDGRID option
LGRID= <i>linetype</i>	specifies line type for grid requested with the GRID option
WGRID= <i>n</i>	specifies width of grid lines

Table 12.12. Graphical Enhancement Options[†]

ANNOTATE= <i>SAS-data-set</i>	specifies annotate data set that adds features to cusum chart
ANNOTATE2= <i>SAS-data-set</i>	specifies annotate data set that adds features to trend chart
DESCRIPTION='string'	specifies string that appears in the description field of PROC GREPLAY master menu for cusum chart
FONT= <i>font</i>	specifies software font for labels and legends on charts
NAME='string'	specifies name that appears in the name field of the PROC GREPLAY master menu for cusum chart
PAGENUM='string'	specifies the form of the label used in pagination
PAGENUMPOS= <i>keyword</i>	specifies the position of the page number requested with the PAGENUM= option
WTREND= <i>n</i>	specifies width of line segments connecting points on trend chart

Table 12.13. Input Data Set Options[†]

MISSBREAK	specifies that observations with missing values are not to be processed
-----------	---

Table 12.14. Output Data Set Options[†]

OUTHISTORY= <i>SAS-data-set</i>	creates output data set containing subgroup summary statistics and cusums
OUTINDEX='string'	specifies value of <code>_INDEX_</code> in OUTLIMITS= data set
OUTLIMITS= <i>SAS-data-set</i>	creates output data set containing cusum parameters
OUTPHASE='string'	specifies value of <code>_PHASE_</code> in OUTHISTORY= data set or OUTTABLE= data set
OUTTABLE= <i>SAS-data-set</i>	creates output data set containing subgroup summary statistics, cusums, and decision interval or V-mask values

[†]The options in these tables are described in Chapter 46, "Dictionary of Options," of Part 9, "The SHEWHART Procedure."

Table 12.15. Plot Layout Options[†]

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

Table 12.16. Options for Interactive Control Charts[†]

HTML=(<i>variable</i>)	specifies a variable whose values are URLs to be associated with subgroups
HTML_LEGEND= <i>(variable)</i>	specifies a variable whose values are URLs to be associated with symbols in the symbol legend
WEBOUT= <i>SAS-data-set</i>	creates an OUTTABLE= data set with additional graphics coordinate data

[†]The options in these tables are described in Chapter 46, “Dictionary of Options,” of Part 9, “The SHEWHART Procedure.”

Table 12.17. Star Options[†]

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

[†]The options in these tables are described in Chapter 46, "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 α 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 395.

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 β 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 β 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.

<i>Graphics</i>

CLIMITS=*color*

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

<i>Graphics</i>

CMASK=*color*

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

<i>Graphics</i>

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.

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

Part 2. The CAPABILITY Procedure

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 δ of the process standard deviation σ or the standard error $\sigma_{\bar{X}}$, depending on whether δ 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 399 and 400.

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 392, and refer to Lucas and Crosier (1982), Ryan (1989), and Montgomery (1996).

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 399 and 400.

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=*n*

LIMITN=VARYING

specifies either a fixed or varying nominal sample size for the control limits. If you specify `LIMITN=n`, cusums are calculated and displayed only for those subgroups with a sample size equal to *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=*linetype*

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

Graphics

LMASK=*linetype*

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

Graphics

MU0=*value*

specifies the target mean μ_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 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 *t*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 `_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 Δ) and the DELTA= *value* (denoted by δ) is $\delta = \Delta / (\sigma / \sqrt{n})$, where σ is the process standard deviation.

SIGMA0=value

specifies a known standard deviation σ_0 for the process standard deviation σ . The *value* must be positive. By default, PROC CUSUM estimates σ from the data using the formulas given in “Methods for Estimating the Standard Deviation” on page 401. You can use the variable _STDDEV_ in a LIMITS= data set as an alternative to the SIGMA0= option.

SIGMAS=value

specifies the probability α 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 σ . Refer to Johnson and Leone (1962, 1974).

SMETHOD=NOWEIGHT | MVLUE | RMSDF

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

Keyword	Method for Estimating Standard Deviation
NOWEIGHT	estimates σ as an unweighted average of unbiased subgroup estimates of σ
MVLUE	calculates a minimum variance linear unbiased estimate for σ
RMSDF	calculates a root-mean square estimate for σ

For formulas, see “Methods for Estimating the Standard Deviation” on page 401.

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 392 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 δ . 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 404 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.

Graphics

WMASK=*linetype*

specifies the width (in pixels) of the V-mask arms. The default width is 1.

Graphics

Details

Basic Notation for Cusum Charts

The following notation is used in this chapter:

- μ denotes the mean of the population, also referred to as the *process mean* or the *process level*.
- μ_0 denotes the target mean (goal) for the population. Goel and Wu (1971) refer to μ_0 as the “acceptable quality level” and use the symbol μ_a instead. The symbol \bar{X}_0 is used for μ_0 in *Glossary and Tables for Statistical Quality Control*. You can provide μ_0 with the MU0= option or with the variable _MU0_ in a LIMITS= data set.
- σ denotes the population standard deviation. You can provide σ with the variable _STDDEV_ in a LIMITS= data set (where _TYPE_=STANDARD).
- σ_0 denotes a known standard deviation. You can provide σ_0 with the SIGMA0= option or the variable _STDDEV_ in a LIMITS= data set.
- $\hat{\sigma}$ denotes an estimate of σ . 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.
- δ denotes the shift in μ to be detected, expressed as a multiple of the standard deviation. You can provide δ with the DELTA= option or the variable _DELTA_ in a LIMITS= data set.
- Δ denotes the shift in μ 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 Δ ; for example, refer to Johnson and Leone (1962, 1974) and Wadsworth and others (1986). You can provide Δ 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 δ , since this generalizes to data with unequal subgroup sample sizes.

Formulas for Cumulative Sums

One-Sided Cusum Schemes

Positive Shifts

If the shift δ 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 $t=1, 2, \dots, n$, 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

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

the sequence S_t cumulates deviations in the subgroup means greater than k standard errors from μ_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 (1996).

Negative Shifts

If the shift δ 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 $t=1, 2, \dots, n$, 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

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

the sequence S_t cumulates the absolute value of deviations in the subgroup means less than k standard errors from μ_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 (1996). Note that S_t is always positive and h is always positive, regardless of whether δ 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 (1996) 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 $t=1, 2, \dots, n$. 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 μ_0 .

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

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

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 (1996), Wadsworth and others (1986), and *ASQC Glossary and Tables for Statistical Quality Control*. If the subgroup sample sizes are constant ($= n$) and if you specify the DATAUNITS option in the XCHART statement, the 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 LIMITN= option or with 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 .

If the process is in control and the mean μ is at or near the target μ_0 , the points will not exhibit a trend since positive and negative displacements from μ_0 tend to cancel each other. If μ shifts in the positive direction, the points exhibit an upward trend, and if μ 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 H= option in the XCHART statement or with the variable _H_ in a LIMITS= data set. The decision interval is displayed as a horizontal line on the cusum chart, as illustrated in Figure 12.12.

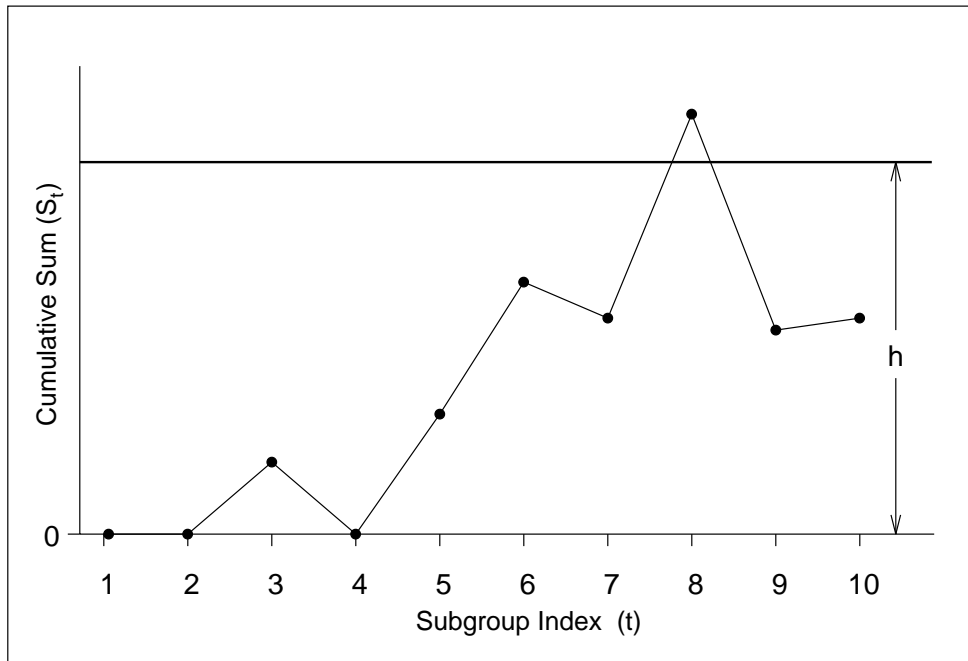


Figure 12.12. Decision Interval

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.

- θ , 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 (1996).
- 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.

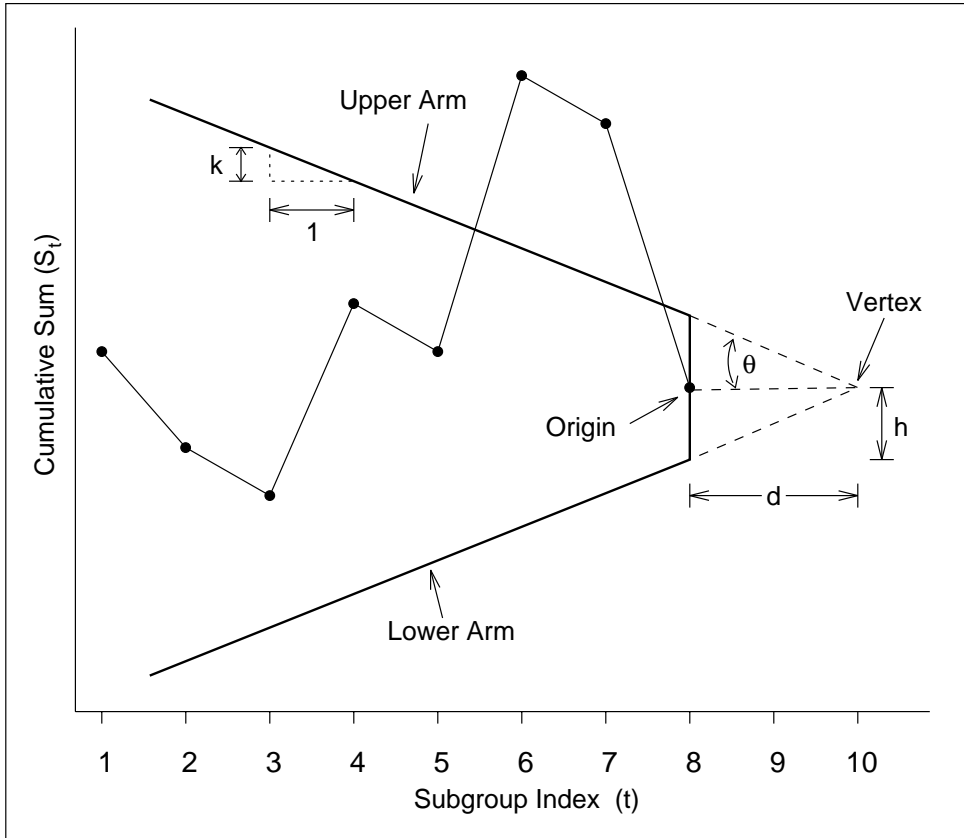


Figure 12.13. V-Mask Parameters

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 α and β (or simply α) 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 α and β , h and k are computed using the formulas

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

$$k = |\delta|/2$$

If you provide α but not β , h and k are computed using the formulas

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

$$k = |\delta|/2$$

In the preceding equations, the error probability α 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 δ 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 μ shifts away from μ_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 (1996), and Wadsworth and others (1986).

For one-sided charts, average run lengths are tabulated as a function of h , k , and δ in Table 12.18 on page 399. No headstart is assumed in this table. For two-sided charts, average run lengths are tabulated as a function of h , k , and δ in Table 12.19 on page 400, which is formatted similarly to Table 2 given by Lucas (1976).

The ARLs in Table 12.18 and Table 12.19 were calculated with the DATA step function CUSUMARL, which is described on page 1849. 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 .

Part 2. The CAPABILITY Procedure

You can specify h , k , and δ 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 $k = |\delta|/2$, referred to as the *central reference value*.

Error Probability Approach

This approach is available only for two-sided cusum schemes. Values of α (the probability of incorrectly signaling the occurrence of a shift) and β (the probability of failing to detect a shift) are specified, and h and k are computed from α and β as described in “Defining the V-Mask for a Two-Sided Cusum Scheme” on page 395. 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 (1996), 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 α and β with the ALPHA= and BETA= options or with the variables _ALPHA_ and _BETA_ in a LIMITS= data set. It is not necessary to specify β , and the interpretation of β 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 12.18. Average Run Lengths for One-Sided V-Mask Cusum Charts as a Function of h , k , and δ .

Parameters		δ (shift in mean)										
h	k	0.00	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00	4.00	5.00
2.50	0.25	27.27	13.43	7.96	5.42	4.06	2.71	2.06	1.68	1.42	1.11	1.01
4.00	0.25	77.08	26.68	13.29	8.38	6.06	3.91	2.93	2.38	2.05	1.61	1.23
6.00	0.25	350.80	51.34	20.90	12.37	8.73	5.51	4.07	3.26	2.74	2.13	1.90
8.00	0.25	736.78	84.00	28.76	16.37	11.39	7.11	5.21	4.15	3.48	2.67	2.14
10.00	0.25	2071.51	124.66	36.71	20.37	14.06	8.71	6.36	5.04	4.20	3.20	2.65
2.00	0.50	38.55	18.19	10.00	6.32	4.45	2.74	1.99	1.58	1.32	1.07	1.01
3.00	0.50	117.60	39.47	17.35	9.68	6.40	3.75	2.68	2.12	1.77	1.31	1.07
4.00	0.50	335.37	77.08	26.68	13.29	8.38	4.75	3.34	2.62	2.19	1.71	1.31
5.00	0.50	930.89	141.69	38.01	17.05	10.38	5.75	4.01	3.11	2.57	2.01	1.69
6.00	0.50	2553.11	250.80	51.34	20.90	12.37	6.75	4.68	3.62	2.98	2.24	1.95
1.50	0.75	42.57	21.09	11.59	7.09	4.78	2.73	1.90	1.48	1.24	1.04	1.00
2.25	0.75	139.71	51.46	22.38	11.66	7.13	3.73	2.51	1.91	1.56	1.16	1.02
3.00	0.75	442.80	117.60	39.47	17.35	9.68	4.73	3.12	2.36	1.93	1.41	1.11
3.75	0.75	1375.71	258.96	65.65	24.16	12.37	5.73	3.71	2.79	2.27	1.72	1.31
4.50	0.75	4251.69	559.95	105.12	32.09	15.15	6.73	4.31	3.21	2.59	1.97	1.60
1.00	1.00	35.29	19.22	11.21	7.03	4.75	2.63	1.78	1.38	1.17	1.02	1.00
1.50	1.00	93.85	42.57	21.09	11.59	7.09	3.50	2.24	1.66	1.34	1.07	1.01
2.00	1.00	258.67	94.34	38.55	18.19	10.00	4.45	2.74	1.99	1.58	1.16	1.02
2.50	1.00	716.00	205.97	68.19	27.27	13.43	5.42	3.25	2.34	1.85	1.31	1.07
3.00	1.00	1962.79	442.80	117.60	39.47	17.35	6.40	3.75	2.68	2.12	1.52	1.16
3.50	1.00	5341.40	943.73	199.57	55.69	21.76	7.39	4.25	3.01	2.37	1.73	1.31
0.70	1.50	67.72	36.03	20.26	12.07	7.63	3.66	2.18	1.55	1.25	1.04	1.00
1.10	1.50	184.28	86.36	42.72	22.50	12.74	5.17	2.80	1.86	1.43	1.08	1.01
1.50	1.50	549.69	221.49	93.85	42.57	21.09	7.09	3.50	2.24	1.66	1.16	1.02
1.90	1.50	1762.09	595.61	210.95	80.54	34.26	9.38	4.26	2.64	1.92	1.29	1.05
2.30	1.50	5897.30	1638.15	476.90	151.04	54.47	12.00	5.03	3.04	2.20	1.45	1.12

Table 12.19. Average Run Lengths for Two-Sided V-Mask Cusum Charts as a Function of h , k , and δ .

Parameters		δ (shift in mean)										
h	k	0.00	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00	4.00	5.00
2.50	0.25	13.64	11.22	7.67	5.38	4.06	2.71	2.06	1.68	1.42	1.11	1.01
4.00	0.25	38.54	24.71	13.20	8.38	6.06	3.91	2.93	2.38	2.05	1.61	1.23
6.00	0.25	125.40	50.33	20.89	12.37	8.73	5.51	4.07	3.26	2.74	2.13	1.90
8.00	0.25	368.39	83.63	28.76	16.37	11.39	7.11	5.21	4.15	3.48	2.67	2.14
10.00	0.25	1035.75	124.55	36.71	20.37	14.06	8.71	6.36	5.04	4.20	3.20	2.65
2.00	0.50	19.27	15.25	9.63	6.27	4.44	2.74	1.99	1.58	1.32	1.07	1.01
3.00	0.50	58.80	36.24	17.20	9.67	6.40	3.75	2.68	2.12	1.77	1.31	1.07
4.00	0.50	167.68	74.22	26.63	13.29	8.38	4.75	3.34	2.62	2.19	1.71	1.31
5.00	0.50	465.44	139.49	38.00	17.05	10.38	5.75	4.01	3.11	2.57	2.01	1.69
6.00	0.50	1276.55	249.26	51.34	20.90	12.37	6.75	4.68	3.62	2.98	2.24	1.95
1.50	0.75	21.28	17.22	11.01	7.00	4.77	2.73	1.90	1.48	1.24	1.04	1.00
2.25	0.75	69.85	45.97	22.04	11.63	7.13	3.73	2.51	1.91	1.56	1.16	1.02
3.00	0.75	221.40	110.95	39.31	17.34	9.68	4.73	3.12	2.36	1.93	1.41	1.11
3.75	0.75	687.85	251.56	65.58	24.16	12.37	5.73	3.71	2.79	2.27	1.72	1.31
4.50	0.75	2125.85	552.11	105.09	32.09	15.15	6.73	4.31	3.21	2.59	1.97	1.60
1.00	1.00	17.65	15.03	10.39	6.88	4.72	2.63	1.78	1.38	1.17	1.02	1.00
1.50	1.00	46.92	35.70	20.31	11.49	7.07	3.50	2.24	1.66	1.34	1.07	1.01
2.00	1.00	129.34	84.00	37.93	18.14	10.00	4.45	2.74	1.99	1.58	1.16	1.02
2.50	1.00	358.00	191.48	67.76	27.25	13.43	5.42	3.25	2.34	1.85	1.31	1.07
3.00	1.00	981.39	423.29	117.32	39.47	17.35	6.40	3.75	2.68	2.12	1.52	1.16
3.50	1.00	2670.70	917.89	199.40	55.69	21.76	7.39	4.25	3.01	2.37	1.73	1.31
0.70	1.50	33.86	28.41	18.90	11.84	7.59	3.66	2.18	1.55	1.25	1.04	1.00
1.10	1.50	92.14	71.41	40.91	22.29	12.71	5.17	2.80	1.86	1.43	1.08	1.01
1.50	1.50	274.84	191.58	91.58	42.39	21.07	7.09	3.50	2.24	1.66	1.16	1.02
1.90	1.50	881.05	536.07	208.31	80.41	34.25	9.38	4.26	2.64	1.92	1.29	1.05
2.30	1.50	2948.65	1523.15	474.09	150.96	54.47	12.00	5.03	3.04	2.20	1.45	1.12

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 1996). 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 48, “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 σ . 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

$$x_1, x_2, \dots, x_N$$

rather than subgroup samples of two or more observations, the CUSUM procedure estimates σ as $\sqrt{\hat{\sigma}^2}$, where

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

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

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

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}, \dots, 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 σ . 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 σ 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, $i = 1, 2, \dots, N$.
N	is the number of subgroups for which $n_i \geq 2$.

The estimate is

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

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 σ 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:

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

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 $(n_1 + \dots + n_N) - (N - 1)$.

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 σ 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 σ is the very thing under test,” and if σ 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 12.20. OUTLIMITS= Data Set

Variable	Description
ALPHA	probability (α) of Type 1 error
ARLIN	average run length for zero shift
ARLOUT	average run length for shift of δ
BETA	probability (β) of Type 2 error
DELTA	shift (δ) to be detected
H	decision interval h for one-sided scheme; distance h between origin and upper arm V-mask for two-sided scheme
HSTART	headstart value
INDEX	optional identifier for cusum parameters (if the OUTINDEX= option is specified)
K	reference value k for one-sided scheme; slope of lower V-mask arm for two-sided scheme
LIMITN	nominal sample size for cusum scheme
MEAN	estimated process mean ($\bar{\bar{X}}$)
MU0	target mean μ_0
ORIGIN	origin of V-mask
SCHEME	type of scheme (ONESIDED or TWOSIDED)
SIGMAS	$z_{1-\alpha/2}$
STDDEV	estimated or known standard deviation ($\hat{\sigma}$ or σ_0)
SUBGRP	<i>subgroup-variable</i> specified in XCHART statement
TYPE	type (ESTIMATE or STANDARD) of _STDDEV_
VAR	<i>process</i> specified in XCHART statement

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= α , _SIGMAS_ is computed as $k = \Phi^{-1}(1 - \alpha/2)$, where Φ^{-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 371.

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:

```
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 367.

OUTTABLE= Data Set

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

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

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 12.21. ODS Tables Produced with the XCHART Statement

Table Name	Description	Options
CompCusum	computational form of the cusum scheme	TABLEALL, TABLECOMP
Parameters	cusum parameters and computed average run lengths	TABLEALL, TABLESUMMARY
XCHART	cusum chart summary statistics	TABLEALL, TABLECHART, TABLEOUT

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 362.

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.
- 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 σ , the variable `_STDDEV_` 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

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 373.

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:

```
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

- `_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 365.

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.

Examples

This section provides advanced examples of the XCHART statement.

Example 12.1. Cusum and Standard Deviation Charts

See CUSXS
in the SAS/QC
Sample Library

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 362).

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.

```
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 _LCLX_ 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.

```
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.

```

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 1580 in Chapter 44, “`XSCHART` Statement”).

Output 12.1.1. Listing of the Data Set TABALL

		S L							E				
		I	I			E			X				
		G	M	S	L	M	X	L	S				
	V	h	M	I	U	C	E	L	C	U			
O	A	o	A	T	B	L	A	I	L	B			
b	R	u	S	N	N	X	N	M	S	S			
s		r											
1	weight	1	3	4	4	0	1.5	0	0.059640	0.049943	0.11317	0.00	3
2	weight	2	3	4	4	0	1.5	0	0.090220	0.049943	0.11317	0.00	3
3	weight	3	3	4	4	0	1.5	0	0.076346	0.049943	0.11317	0.00	3
4	weight	4	3	4	4	0	1.5	0	0.025552	0.049943	0.11317	0.00	3
5	weight	5	3	4	4	0	1.5	0	0.026500	0.049943	0.11317	0.00	3
6	weight	6	3	4	4	0	1.5	0	0.075617	0.049943	0.11317	0.30	3
7	weight	7	3	4	4	0	1.5	0	0.037242	0.049943	0.11317	0.00	3
8	weight	8	3	4	4	0	1.5	0	0.059290	0.049943	0.11317	0.18	3
9	weight	9	3	4	4	0	1.5	0	0.005737	0.049943	0.11317	1.21	3
10	weight	10	3	4	4	0	1.5	0	0.046522	0.049943	0.11317	0.62	3
11	weight	11	3	4	4	0	1.5	0	0.040542	0.049943	0.11317	0.00	3
12	weight	12	3	4	4	0	1.5	0	0.056103	0.049943	0.11317	0.00	3

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.

```

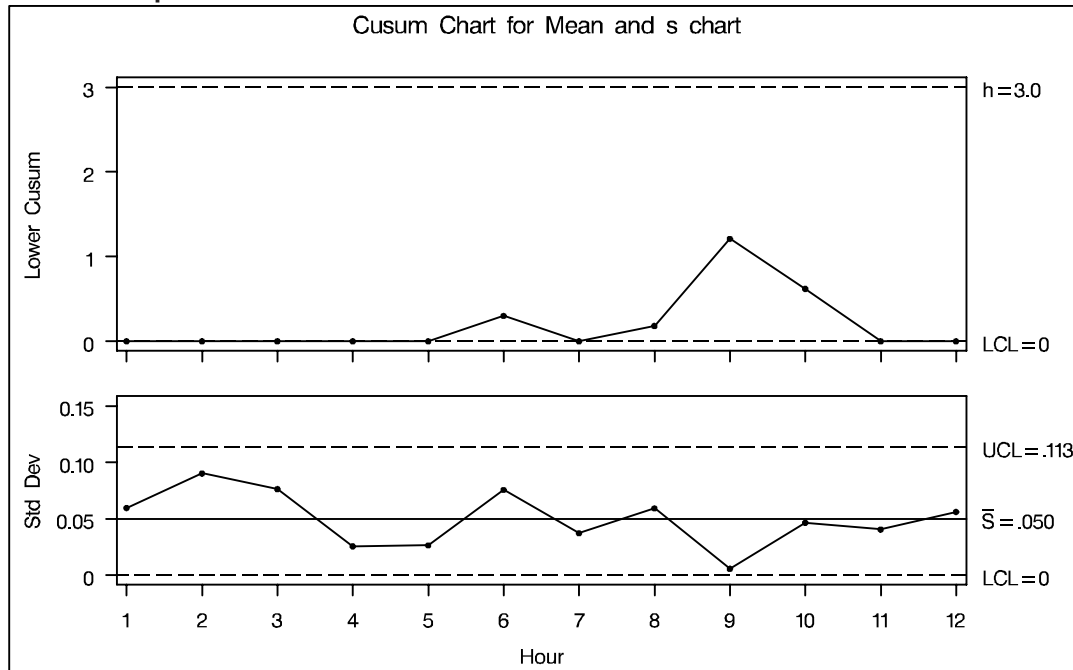
title 'Cusum Chart for Mean and s chart';
symbol v=dot;
proc shewhart table=taball;
  xschart weight * hour /
    ucllabel = 'h=3.0'
    nolimitslegend
    noctl
    split = '/'
    nolegend ;
  label _subx_ = 'Lower Cusum/Std Dev';
run;

```

Part 2. The CAPABILITY Procedure

The central line for the primary (cusum) chart is suppressed with the NOCTL option, and the default 3σ Limits legend is suppressed with the NOLIMITLEGEND option. The charts are shown in Output 12.1.2.

Output 12.1.2. Combined Cusum Chart and s Chart



The process variability is stable, and there is no signal of a downward shift in the process mean.

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.

See CUSUPLO
in the SAS/QC
Sample Library

The following statements analyze the variable WEIGHT in the data set OIL (see “Creating a V-Mask Cusum Chart from Raw Data” on page 362). 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. Now,

Part 2. The CAPABILITY Procedure

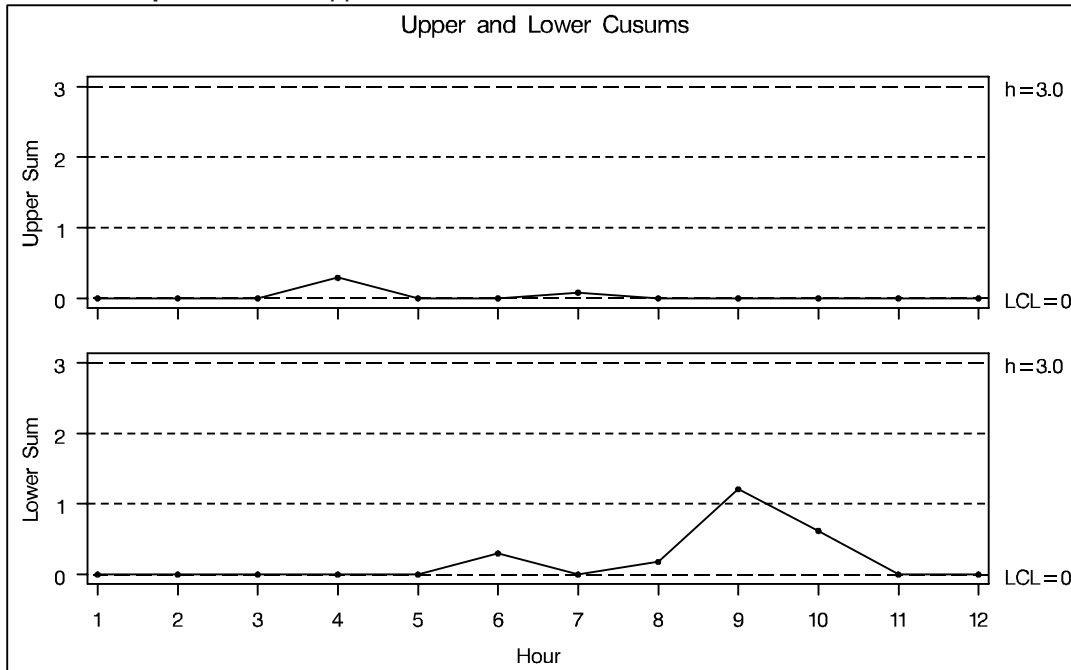
TABBOTH has the structure required for a TABLE= data set used with the XSCHART statement in the SHEWHART procedure (see “TABLE= Data Set” on page 1580 in Chapter 44, “XSCHART Statement”).

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=tabbboth;
  xschart weight*hour /
    nolimitslegend
    ucllabel = 'h=3.0'
    ucllabel2 = 'h=3.0'
    ypct1    = 50
    vref     = 1 2
    vref2    = 1 2
    lvref    = 2
    noct1
    noct12
    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



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 368.

See CUSCOMB
in the SAS/QC
Sample Library

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

```
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.

```
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, 3σ limits are computed, but the multiple of σ 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.

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

Part 2. The CAPABILITY Procedure

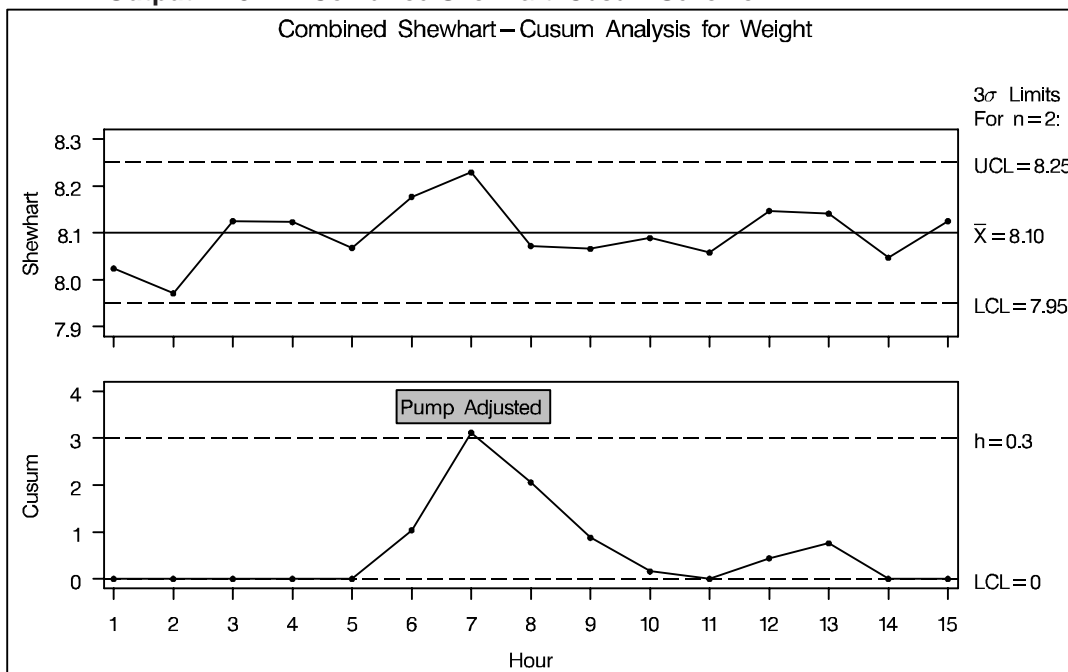
The data set COMBINE has the structure required for a TABLE= data set used with the IRCHART statement in the SHEWHART procedure (see “TABLE= Data Set” on page 1171 in Chapter 34, “IRCHART Statement”).

Finally, the combined scheme is displayed with the SHEWHART procedure.

```
symbol v=dot;
title 'Combined Shewhart-Cusum Analysis for Weight';
proc shewhart table=combine;
  irchart weight*hour /
    ypct1      = 50
    noct12
    ucllabel2  = '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 $\delta = 1$.

The correct bibliographic citation for this manual is as follows: SAS Institute Inc., *SAS/QC[®] User's Guide, Version 8*, Cary, NC: SAS Institute Inc., 1999. 1994 pp.

SAS/QC[®] User's Guide, Version 8

Copyright © 1999 SAS Institute Inc., Cary, NC, USA.

ISBN 1-58025-493-4

All rights reserved. Printed in the United States of America. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, by any form or by any means, electronic, mechanical, photocopying, or otherwise, without the prior written permission of the publisher, SAS Institute Inc.

U.S. Government Restricted Rights Notice. Use, duplication, or disclosure of the software by the government is subject to restrictions as set forth in FAR 52.227-19 Commercial Computer Software-Restricted Rights (June 1987).

SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513.

1st printing, October 1999

SAS[®] and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute in the USA and other countries. [®] indicates USA registration.

IBM[®], ACF/VTAM[®], AIX[®], APPN[®], MVS/ESA[®], OS/2[®], OS/390[®], VM/ESA[®], and VTAM[®] are registered trademarks or trademarks of International Business Machines Corporation. [®] indicates USA registration.

Other brand and product names are registered trademarks or trademarks of their respective companies.

The Institute is a private company devoted to the support and further development of its software and related services.