Chapter 43
XRCHART Statement

Chapter Table of Contents

OVERVIEW 1501

GETTING STARTED 1502
Creating Charts for Means and Ranges from Raw Data 1502
Creating Charts for Means and Ranges from Summary Data 1505
Saving Summary Statistics 1507
Saving Control Limits 1508
Reading Preestablished Control Limits 1512

SYNTAX 1514
Summary of Options 1515

DETAILS 1525
Constructing Charts for Means and Ranges 1525
Output Data Sets 1527
ODS Tables 1530
Input Data Sets 1530
Subgroup Variables 1534
Methods for Estimating the Standard Deviation 1536
Capability Indices 1537
Axis Labels 1538
Missing Values 1539

EXAMPLES 1540
Example 43.1 Applying Tests for Special Causes 1540
Example 43.2 Specifying Standard Values for the Process Mean and Standard Deviation 1543
Example 43.3 Working with Unequal Subgroup Sample Sizes 1544

Chapter 43
XRCHART Statement

Overview

The XRCHART statement creates $\bar{X}$ and R charts for subgroup means and ranges, which are used to analyze the central tendency and variability of a process.

You can use options in the XRCHART statement to

- compute control limits from the data based on a multiple of the standard error of the plotted means and ranges or as probability limits
- tabulate subgroup sample sizes, subgroup means, subgroup ranges, control limits, and other information
- save control limits in an output data set
- save subgroup sample sizes, subgroup means, and subgroup ranges in an output data set
- read preestablished control limits from a data set
- apply tests for special causes (also known as runs tests and Western Electric rules)
- specify a method for estimating the process standard deviation
- specify a known (standard) process mean and standard deviation for computing control limits
- display distinct sets of control limits for data from successive time phases
- add block legends and symbol markers to reveal stratification in process data
- superimpose stars at points to represent related multivariate factors
- clip extreme points to make the charts more readable
- display vertical and horizontal reference lines
- control axis values and labels
- control layout and appearance of the chart

Getting Started

This section introduces the XRCHART statement with simple examples that illustrate commonly used options. Complete syntax for the XRCHART statement is presented in the "Syntax" section on page 1514, and advanced examples are given in the "Examples" section on page 1540.

Creating Charts for Means and Ranges from Raw Data

See SHWXR1 in the SAS/QC Sample Library

In the manufacture of silicon wafers, batches of five wafers are sampled, and their diameters are measured in millimeters. The following statements create a SAS data set named WAFERS, which contains the measurements for 25 batches:

```
data wafers;
   input batch @;
   do i=1 to 5;
      input diamtr @;
      output;
      end;
   drop i;
datalines;
 1  35.00 34.99 34.99 34.98 35.00
 2  35.01 34.99 34.99 34.98 35.00
 3  34.99 35.00 35.00 35.00 35.00
 4  35.01 35.00 34.99 34.99 35.00
 5  35.00 34.99 34.98 34.99 35.00
 6  34.99 34.99 35.00 35.00 35.00
 7  35.01 34.98 35.00 35.00 34.99
 8  35.00 35.00 34.99 34.98 34.99
 9  34.99 34.98 34.98 35.01 35.00
10  34.99 35.00 35.01 34.99 35.01
11  35.01 35.00 35.00 34.98 34.99
12  34.99 34.99 35.00 34.98 35.01
13  35.01 34.99 34.98 34.99 34.99
14  35.00 35.00 34.99 35.01 34.99
15  34.98 34.99 34.99 34.98 35.00
16  34.99 35.00 35.00 35.01 35.00
17  34.98 34.98 34.99 34.99 34.98
18  35.01 35.02 35.00 34.98 35.00
19  34.99 34.98 35.00 34.99 34.98
20  34.99 35.00 35.00 34.99 34.99
21  35.00 34.99 34.99 34.98 35.00
22  35.00 35.00 35.01 35.00 35.00
23  35.02 35.00 34.98 35.02 35.00
24  35.00 35.00 34.99 35.01 34.98
25  34.99 34.99 34.99 35.00 35.00
;
```

The following statements use the PRINT procedure to list the data set WAFERS. A portion of this listing is shown in Figure 43.1.

```
title 'The Data Set WAFERS';
proc print data=wafers noobs;
run;
```

The Data Set WAFERS

batch	diamtr
1	35.00
1	34.99
1	34.99
1	34.98
1	35.00
2	35.01
2	34.99
2	34.99
2	34.98
2	35.00
3	34.99
3	35.00
3	35.00
3	35.00
3	35.00
.	.
.	.
.	.

Figure 43.1. Partial Listing of the Data Set WAFERS

The data set WAFERS is said to be in "strung-out" form since each observation contains the batch number and diameter measurement for a single wafer. The first five observations contain the diameters for the first batch, the second five observations contain the diameters for the second batch, and so on. Because the variable BATCH classifies the observations into rational subgroups, it is referred to as the *subgroup-variable*. The variable DIAMTR contains the wafer diameter measurements and is referred to as the *process variable* (or *process* for short).

You can use $\bar{X}$ and R charts to determine whether the manufacturing process is in control. The following statements create the $\bar{X}$ and R charts shown in Figure 43.2:

```
title 'Mean and Range Charts for Diameters';
symbol v=dot;
proc shewhart data=wafers;
   xrchart diamtr*batch;
run;
```

This example illustrates the basic form of the XRCHART statement. After the keyword XRCHART, which specifies the type of control chart to display, you specify the *process* to analyze (in this case, DIAMTR) followed by an asterisk and the *subgroup-variable* (BATCH).

The input data set is specified with the DATA= option in the PROC SHEWHART statement.

If you use a graphics device, the SYMBOL statement specifies the symbol to plot the points. For more information on the SYMBOL statement, refer to *SAS/GRAPH Software: Reference*.

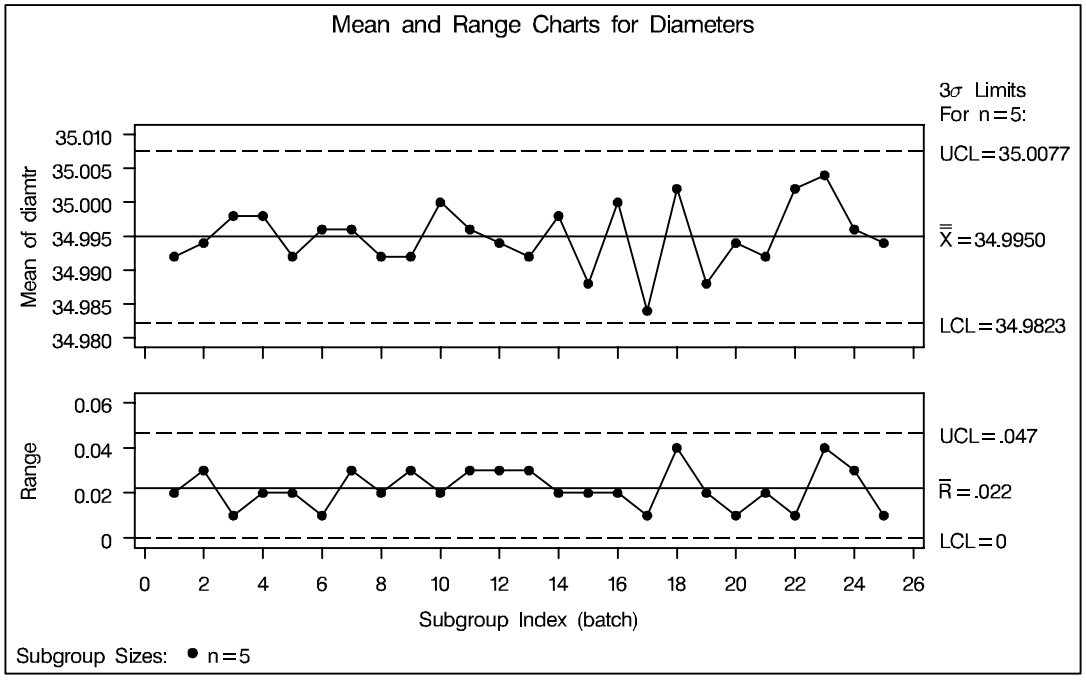


Figure 43.2. $\bar{X}$ and R Charts for Wafer Diameter Data

Each point on the $\bar{X}$ chart represents the average (mean) of the measurements for a particular batch. For instance, the mean plotted for the first batch is

$$\frac{35.00 + 34.99 + 34.99 + 34.98 + 35.00}{5} = 34.992$$

Each point on the R chart represents the range of the measurements for a particular batch. For instance, the range plotted for the first batch is $35.00 - 34.98 = 0.02$.

By default, the control limits shown are 3σ limits estimated from the data; the formulas for the limits are given in Table 43.22 on page 1526. You can also read control limits from an input data set; see "Reading Preestablished Control Limits" on page 1512.

Since all the points lie within the control limits, it can be concluded that the process is in statistical control. For computational details, see "Constructing Charts for Means and Ranges" on page 1525. For more details on reading raw data, see "DATA= Data Set" on page 1530.

Creating Charts for Means and Ranges from Summary Data

See SHWXR1 in the SAS/QC Sample Library

The previous example illustrates how you can create $\bar{X}$ and R charts using raw data (process measurements). However, in many applications, the data are provided as subgroup means and ranges. This example illustrates how you can use the XR-CHART statement with data of this type.

The following data set (WAFERSUM) provides the data from the preceding example in summarized form:

```
data wafersum;
   input batch diamtrx diamtrr;
   diamtrn = 5;
datalines;
 1  34.992  0.02
 2  34.994  0.03
 3  34.998  0.01
 4  34.998  0.02
 5  34.992  0.02
 6  34.996  0.01
 7  34.996  0.03
 8  34.992  0.02
 9  34.992  0.03
10  35.000  0.02
11  34.996  0.03
12  34.994  0.03
13  34.992  0.03
14  34.998  0.02
15  34.988  0.02
16  35.000  0.02
17  34.984  0.01
18  35.002  0.04
19  34.988  0.02
20  34.994  0.01
21  34.992  0.02
22  35.002  0.01
23  35.004  0.04
24  34.996  0.03
25  34.994  0.01
;
```

A partial listing of the data set WAFERSUM is shown in Figure 43.3.

Summary Data Set for Wafer Diameters

batch	diamtrx	diamtrr	diamtrn
1	34.992	0.02	5
2	34.994	0.03	5
3	34.998	0.01	5
.	.	.	.
.	.	.	.
.	.	.	.

Figure 43.3. The Summary Data Set WAFERSUM

In this data set, there is exactly one observation for each subgroup (note that the subgroups are still indexed by BATCH). The variable DIAMTRX contains the subgroup means, the variable DIAMTRR contains the subgroup ranges, and the variable DIAMTRN contains the subgroup sample sizes (these are all equal to five).

You can read this data set by specifying it as a HISTORY= data set in the PROC SHEWHART statement, as follows:

```
title 'Mean and Range Charts for Diameters';
proc shewhart history=wafersum lineprinter;
   xrchart diamtr*batch='*';
run;
```

The charts are shown in Figure 43.4. Since the LINEPRINTER option is specified in the PROC SHEWHART statement, line printer output is produced. * The asterisk (*) specified in single quotes after the *subgroup-variable* indicates the character used to plot the points. This character must follow an equal sign.

Note that DIAMTR is *not* the name of a SAS variable in the data set WAFERSUM but is, instead, the common prefix for the names of the three SAS variables DIAMTRX, DIAMTRR, and DIAMTRN. The suffix characters *X*, *R*, and *N* indicate *mean*, *range*, and *sample size*, respectively. Thus, you can specify three subgroup summary variables in the HISTORY= data set with a single name (DIAMTR), which is referred to as the *process*. The name BATCH specified after the asterisk is the name of the *subgroup-variable*.

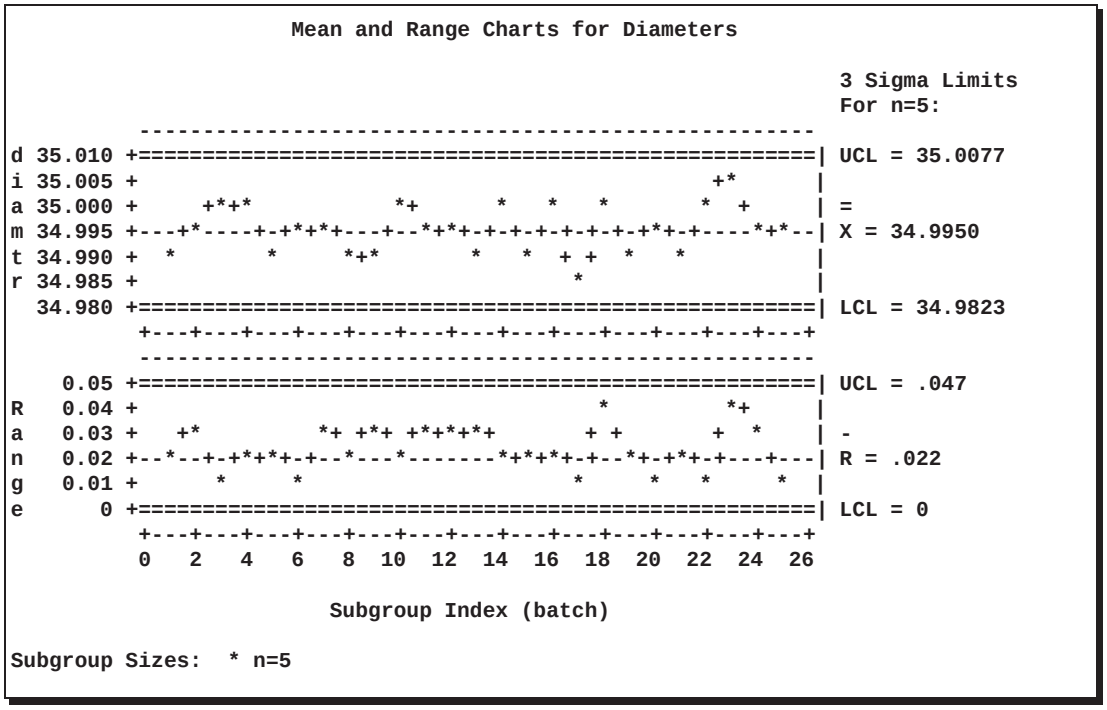


Figure 43.4. $\bar{X}$ and R Charts from Summary Data

*In Release 6.12 and previous releases of SAS/QC software, the keyword GRAPHICS was required in the PROC SHEWHART 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.

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

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

Furthermore, the names of the subgroup mean, range, and sample size variables must begin with the *process* name specified in the XRCHART statement and end with the special suffix characters *X*, *R*, and *N*, respectively. If the names do not follow this convention, you can use the RENAME option to rename the variables for the duration of the SHEWHART procedure step. Suppose that, instead of the variables DIAMTRX, DIAMTRR, and DIAMTRN, the data set WAFERSUM contained summary variables named MEANS, RANGES, and SIZES. The following statements would temporarily rename MEANS, RANGES, and SIZES to DIAMTRX, DIAMTRR, and DIAMTRN, respectively:

```
proc shewhart
   history=wafersum (rename=(means  = diamtrx
                             ranges = diamtrr
                             sizes  = diamtrn ));
   xrchart diamtr*batch='*';
run;
```

In summary, the interpretation of *process* depends on the input data set:

- 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 of the variables containing the summary statistics.

For more information, see "HISTORY= Data Set" on page 1531.

Saving Summary Statistics

In this example, the XRCHART statement is used to create a summary data set that can be read later by the SHEWHART procedure (as in the preceding example). The following statements read measurements from the data set WAFERS and create a summary data set named WAFRHIST:

See SHWXR1 in the SAS/QC Sample Library

```
title 'Summary Data Set for Wafer Diameters';
proc shewhart data=wafers;
   xrchart diamtr*batch / outhistory = wafrhist
                          nochart;
run;
```

The OUTHISTORY= option names the output data set, and the NOCHART option suppresses the display of the charts. Options such as OUTHISTORY= and

NOCHART are specified after the slash (/) in the XRCHART statement. A complete list of options is presented in the "Syntax" section on page 1514.

Figure 43.5 contains a partial listing of WAFRHIST.

Summary Data Set for Wafer Diameters

batch	diamtrX	diamtrR	diamtrN
1	34.992	0.02	5
2	34.994	0.03	5
3	34.998	0.01	5
4	34.998	0.02	5
5	34.992	0.02	5
.	.	.	.
.	.	.	.
.	.	.	.

Figure 43.5. The Summary Data Set WAFRHIST

There are four variables in the data set WAFRHIST.

- BATCH contains the subgroup index.
- DIAMTRX contains the subgroup means.
- DIAMTRR contains the subgroup ranges.
- DIAMTRN contains the subgroup sample sizes.

Note that the summary statistic variables are named by adding the suffix characters *X*, *R*, and *N* to the *process* DIAMTR specified in the XRCHART statement. In other words, the variable naming convention for OUTHISTORY= data sets is the same as that for HISTORY= data sets.

For more information, see "OUTHISTORY= Data Set" on page 1528.

Saving Control Limits

See SHWXR1 in the SAS/QC Sample Library

You can save the control limits for $\bar{X}$ and R charts in a SAS data set; this enables you to apply the control limits to future data (see "Reading Preestablished Control Limits" on page 1512) or modify the limits with a DATA step program.

The following statements read measurements from the data set WAFERS (see page 1502) and save the control limits displayed in Figure 43.2 in WAFERLIM:

```
title 'Control Limits for Wafer Diameters';
proc shewhart data=wafers;
   xrchart diamtr*batch / outlimits = waferlim
                          nochart;
run;
```

The OUTLIMITS= option names the data set containing the control limits, and the NOCHART option suppresses the display of the charts. The data set WAFERLIM is listed in Figure 43.6.

Control Limits for Wafer Diameters

VAR	_SUBGRP_	_TYPE_	_LIMITN_	_ALPHA_	_SIGMAS_	_LCLX_
diamtr	batch	ESTIMATE	5	.002699796	3	34.9823

MEAN	_UCLX_	_LCLR_	_R_	_UCLR_	_STDDEV_
34.9950	35.0077	0	0.022	0.046519	.009458586

Figure 43.6. The Data Set WAFERLIM Containing Control Limit Information

The data set WAFERLIM contains one observation with the limits for *process* DIAMTR. The variables _LCLX_ and _UCLX_ contain the lower and upper control limits for the $\bar{X}$ chart. The variables _LCLR_ and _UCLR_ contain the lower and upper control limits for the R chart. The variable _MEAN_ contains the central line for the $\bar{X}$ chart, and the variable _R_ contains the central line for the R chart. The value of _MEAN_ is an estimate of the process mean, and the value of _STDDEV_ is an estimate of the process standard deviation σ. The value of _LIMITN_ is the nominal sample size associated with the control limits, and the value of _SIGMAS_ is the multiple of σ associated with the control limits. The variables _VAR_ and _SUBGRP_ are bookkeeping variables that save the *process* and *subgroup-variable*. The variable _TYPE_ is a bookkeeping variable that indicates whether the values of _MEAN_ and _STDDEV_ are estimates or standard values.

You can save process capability indices in an OUTLIMITS= data set if you provide specification limits with the LSL= and USL= options. This is illustrated by the following statements:

```
title 'Control Limits and Capability Indices';
proc shewhart data=wafers;
   xrchart diamtr*batch / outlimits = wafrlim2
                          usl       = 35.03
                          lsl       = 34.97
                          nochart;
run;
```

The data set WAFRLIM2 is listed in Figure 43.7.

Control Limits and Capability Indices

VAR	_SUBGRP_	_TYPE_	_LIMIT_	_ALPHAS_	_SIGMAS_	_LCLX_	_MEANX_	_UCLX_	_CLR_	_R_
diamtr	batch	ESTIMATE	5	.002699796	3	34.9823	34.9950	35.0077	0	0.022

UCLR	_STDDEV_	_LSL_	_USL_	_CP_	_CPL_	_CPU_	_CPK_
0.046519	.009458586	34.97	35.03	1.05724	0.87962	1.23486	0.87962

Figure 43.7. The Data Set WAFRLIM2 Containing Process Capability Indices

The variables _CP_, _CPL_, _CPU_, and _CPK_ contain the process capability indices. It is reasonable to compute capability indices, since Figure 43.2 indicates that the wafer process is in statistical control. However, it is recommended that you also check for normality of the data. You can use the CAPABILITY procedure for this purpose.

For more information, see “OUTLIMITS= Data Set” on page 1527.

You can create an output data set containing both control limits and summary statistics with the OUTTABLE= option, as illustrated by the following statements:

```
title 'Summary Statistics and Control Limit Information';
proc shewhart data=wafers;
   xrchart diamtr*batch / outtable=wtable
                          nochart;
run;
```

The data set WTABLE is listed in Figure 43.8.

Summary Statistics and Control Limit Information

VAR	batch	_SIGMA_	_MIN_	_SUBN_	_LCLX_	_SUBX_	_MEAN_	_UCLX_	_EXLIM_	_LCLR_	_SUBR_	_UCLR_	_EXLIM_
diamtr	1	3	5	5	34.9823	34.992	34.9950	35.0077	0	0.02	0.022	0.046519	
diamtr	2	3	5	5	34.9823	34.994	34.9950	35.0077	0	0.03	0.022	0.046519	
diamtr	3	3	5	5	34.9823	34.998	34.9950	35.0077	0	0.01	0.022	0.046519	
diamtr	4	3	5	5	34.9823	34.998	34.9950	35.0077	0	0.02	0.022	0.046519	
diamtr	5	3	5	5	34.9823	34.992	34.9950	35.0077	0	0.02	0.022	0.046519	
diamtr	6	3	5	5	34.9823	34.996	34.9950	35.0077	0	0.01	0.022	0.046519	
diamtr	7	3	5	5	34.9823	34.996	34.9950	35.0077	0	0.03	0.022	0.046519	
diamtr	8	3	5	5	34.9823	34.992	34.9950	35.0077	0	0.02	0.022	0.046519	
diamtr	9	3	5	5	34.9823	34.992	34.9950	35.0077	0	0.03	0.022	0.046519	
diamtr	10	3	5	5	34.9823	35.000	34.9950	35.0077	0	0.02	0.022	0.046519	
diamtr	11	3	5	5	34.9823	34.996	34.9950	35.0077	0	0.03	0.022	0.046519	
diamtr	12	3	5	5	34.9823	34.994	34.9950	35.0077	0	0.03	0.022	0.046519	
diamtr	13	3	5	5	34.9823	34.992	34.9950	35.0077	0	0.03	0.022	0.046519	
diamtr	14	3	5	5	34.9823	34.998	34.9950	35.0077	0	0.02	0.022	0.046519	
diamtr	15	3	5	5	34.9823	34.988	34.9950	35.0077	0	0.02	0.022	0.046519	
diamtr	16	3	5	5	34.9823	35.000	34.9950	35.0077	0	0.02	0.022	0.046519	
diamtr	17	3	5	5	34.9823	34.984	34.9950	35.0077	0	0.01	0.022	0.046519	
diamtr	18	3	5	5	34.9823	35.002	34.9950	35.0077	0	0.04	0.022	0.046519	
diamtr	19	3	5	5	34.9823	34.988	34.9950	35.0077	0	0.02	0.022	0.046519	
diamtr	20	3	5	5	34.9823	34.994	34.9950	35.0077	0	0.01	0.022	0.046519	
diamtr	21	3	5	5	34.9823	34.992	34.9950	35.0077	0	0.02	0.022	0.046519	
diamtr	22	3	5	5	34.9823	35.002	34.9950	35.0077	0	0.01	0.022	0.046519	
diamtr	23	3	5	5	34.9823	35.004	34.9950	35.0077	0	0.04	0.022	0.046519	
diamtr	24	3	5	5	34.9823	34.996	34.9950	35.0077	0	0.03	0.022	0.046519	
diamtr	25	3	5	5	34.9823	34.994	34.9950	35.0077	0	0.01	0.022	0.046519	

Figure 43.8. The Data Set WTABLE

This data set contains one observation for each subgroup sample. The variables _SUBX_, _SUBR_, and _SUBN_ contain the subgroup means, subgroup ranges, and subgroup sample sizes. The variables _LCLX_ and _UCLX_ contain the lower and upper control limits for the $\bar{X}$ chart. The variables _LCLR_ and _UCLR_ contain the lower and upper control limits for the R chart. The variable _MEAN_ contains the central line of the $\bar{X}$ chart, and the variable _R_ contains the central line of the R chart. The variables _VAR_ and BATCH contain the *process* name and values of the *subgroup-variable*, respectively. For more information, see "OUTTABLE= Data Set" on page 1529.

An OUTTABLE= data set can be read later as a TABLE= data set. For example, the following statements read WTABLE and display $\bar{X}$ and R charts identical to those in Figure 43.2:

```
title 'Mean and Range Charts for Diameters';
proc shewhart table=wtable;
   xrchart diamtr*batch;
run;
```

Because the SHEWHART procedure simply displays the information read from a TABLE= data set, you can use TABLE= data sets to create specialized control charts (see Chapter 49, “Specialized Control Charts”).

For more information, see “TABLE= Data Set” on page 1532.

Reading Preestablished Control Limits

See SHWXR1 in the SAS/QC Sample Library

In the previous example, the OUTLIMITS= data set saved control limits computed from the measurements in WAFERS. This example shows how these limits can be applied to new data provided in the following data set:

```
data wafers2;
   input batch @;
   do i=1 to 5;
      input diamtr @;
      output;
      end;
   drop i;
   datalines;
26 34.99 34.99 35.00 34.99 35.00
27 34.99 35.01 34.98 34.98 34.97
28 35.00 34.99 34.99 34.99 35.01
29 34.98 34.96 34.98 34.98 34.99
30 34.98 35.00 34.98 34.98 34.99
31 35.00 35.00 34.99 35.01 35.01
32 35.00 34.99 34.98 34.98 35.00
33 34.98 35.00 34.99 35.00 35.01
34 35.00 34.97 35.00 34.99 35.01
35 34.99 34.99 34.98 34.99 34.98
36 35.01 34.98 34.99 34.99 35.00
37 35.01 34.99 34.97 34.98 35.00
38 34.98 34.99 35.00 34.98 35.00
39 34.99 34.99 34.99 34.99 35.01
40 34.99 35.01 35.00 35.01 34.99
41 34.99 35.00 34.99 34.98 34.99
42 35.00 34.99 34.98 34.99 35.00
43 34.99 34.98 34.98 34.99 34.99
44 35.00 35.00 34.98 35.00 34.99
45 34.99 34.99 35.00 34.99 34.99
;
```

The following statements create $\bar{X}$ and R charts for the data in WAFERS2 using the control limits in WAFERLIM:

```
title 'Mean and Range Charts for Diameters';
proc shewhart data=wafers2 limits=waferlim;
   xrchart diamtr*batch;
run;
```

The charts are shown in Figure 43.9.

The LIMITS= option in the PROC SHEWHART statement specifies the data set containing the control limits. By default,* this information is read from the first observation in the LIMITS= data set for which

- the value of _VAR_ matches the *process* name DIAMTR
- the value of _SUBGRP_ matches the *subgroup-variable* name BATCH

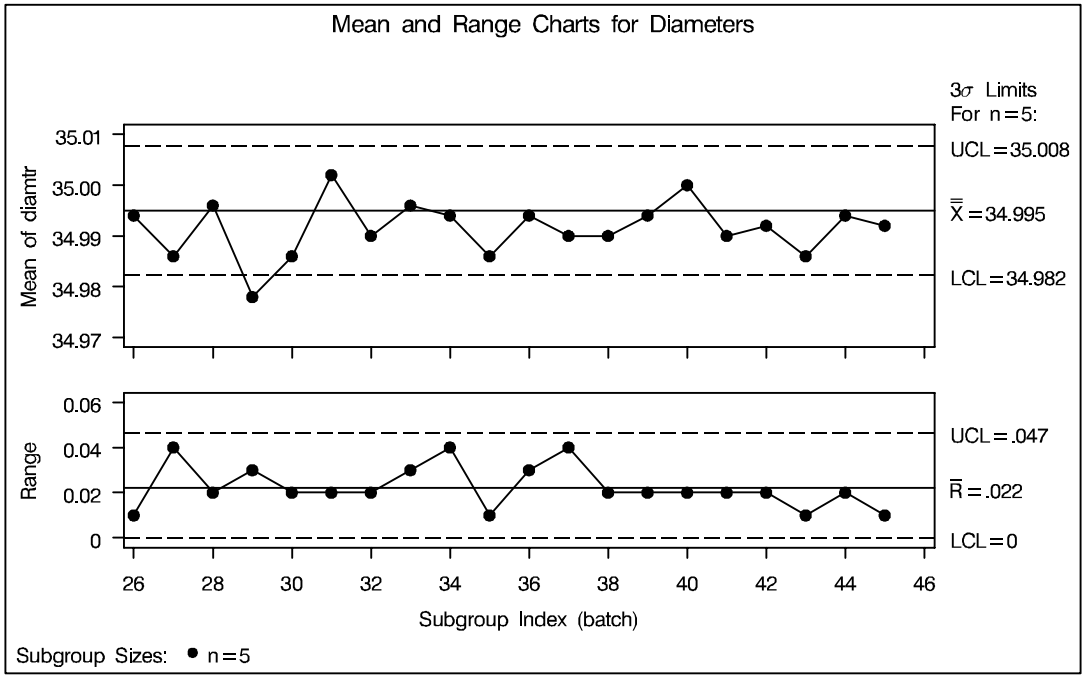


Figure 43.9. $\bar{X}$ and R Charts for Second Set of Wafer Data

Note that the mean diameter of the 29^{th} batch lies below the lower control limit in the $\bar{X}$ chart, signaling a special cause of variation.

In this example, the LIMITS= data set was created in a previous run of the SHEWHART procedure. You can also create a LIMITS= data set with the DATA step. See "LIMITS= Data Set" on page 1531 for details concerning the variables that you must provide.

*In Release 6.09 and in earlier releases, it is also necessary to specify the READLIMITS option to read control limits from a LIMITS= data set.

Syntax

The basic syntax for the XRCHART statement is as follows:

XRCHART *process******subgroup-variable* **;**

The general form of this syntax is as follows:

XRCHART (*processes***)***subgroup-variable* <**(***block-variables* **)** >
< =*symbol-variable* | =*'character'* > < **/** *options* >**;**

You can use any number of XRCHART statements in the SHEWHART procedure. The components of the XRCHART 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 SHEWHART 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 Charts for Means and Ranges from Raw Data" on page 1502.
- 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 Charts for Means and Ranges from Summary Data" on page 1505.
- If summary data and control limits are read from a TABLE= data set, *process* must be the value of the variable _VAR_ in the TABLE= data set. For an example, see "Saving Control Limits" on page 1508.

A *process* is required. If you specify more than one *process*, enclose the list in parentheses. For example, the following statements request distinct $\bar{X}$ and R charts for WEIGHT, LENGTH, and WIDTH:

```
proc shewhart data=measures;
   xrchart (weight length width)*day;
run;
```

subgroup-variable
is the variable that identifies subgroups in the data. The *subgroup-variable* is required. In the preceding XRCHART statement, DAY is the subgroup variable. For details, see "Subgroup Variables" on page 1534.

block-variables
are optional 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 "Displaying Stratification in Blocks of Observations" on page 1684 for an example.

symbol-variable
is an optional variable whose levels (unique values) determine the symbol marker or character used to plot the means and ranges.

- If you produce a chart on a line printer, an 'A' is displayed for the points corresponding to the first level of the *symbol-variable*, a 'B' is displayed for the points corresponding to the second level, and so on.
- If you produce a chart on a graphics device, distinct symbol markers are displayed for points corresponding to the various levels of the *symbol-variable*. You can specify the symbol markers with SYMBOLn statements. See "Displaying Stratification in Levels of a Classification Variable" on page 1683 for an example.

character
specifies a plotting character for charts produced on line printers. For example, the following statements create $\bar{X}$ and R charts using an asterisk (*) to plot the points:

```
proc shewhart data=values;
   xrchart weight*day='*';
run;
```

options
enhance the appearance of the charts, request additional analyses, save results in data sets, and so on. The "Summary of Options" section, which follows, lists all options by function. Chapter 46, "Dictionary of Options," describes each option in detail.

Summary of Options

The following tables list the XRCHART statement options by function. For complete descriptions, see Chapter 46, "Dictionary of Options."

Table 43.1. Tabulation Options

TABLE	creates a basic table of subgroup means, subgroup ranges, subgroup sample sizes, and control limits
TABLEALL	is equivalent to the options TABLE, TABLECENTRAL, TABLEID, TABLELEGEND, TABLEOUTLIM, and TABLETESTS
TABLECENTRAL	augments basic table with values of central lines
TABLEID	augments basic table with columns for ID variables
TABLELEGEND	augments basic table with legend for tests for special causes
TABLEOUTLIM	augments basic table with columns indicating control limits exceeded
TABLETESTS	augments basic table with columns indicating which tests for special causes are positive

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

Table 43.2. Options for Specifying Tests for Special Causes

NO3SIGMACHECK	allows tests to be applied with control limits other than 3σ limits
TESTS=*value-list*| *customized-pattern-list*	specifies tests for special causes for the $\bar{X}$ chart
TESTS2=*value-list*| *customized-pattern-list*	specifies tests for special causes for the R chart
TEST2RUN=*n*	specifies length of pattern for Test 2
TEST3RUN=*n*	specifies length of pattern for Test 3
TESTACROSS	applies tests across *phase* boundaries
TESTLABEL='*label*'| (*variable*)|*keyword*	provides labels for points where test is positive
TESTLABEL*n*='*label*'	specifies label for n^{th} test for special causes
TESTNMETHOD= STANDARDIZE	applies tests to standardized chart statistics
TESTOVERLAP	performs tests on overlapping patterns of points
ZONELABELS	adds labels A, B, and C to zone lines for $\bar{X}$ chart
ZONE2LABELS	adds labels A, B, and C to zone lines for R chart
ZONES	adds lines to $\bar{X}$ chart delineating zones A, B, and C
ZONES2	adds lines to R chart delineating zones A, B, and C
ZONEVALPOS=*n*	specifies position of ZONEVALUES and ZONE2VALUESlabels
ZONEVALUES	labels $\bar{X}$ chart zone lines with their values
ZONE2VALUES	labels R chart zone lines with their values

Table 43.3. Graphical Options for Displaying Tests for Special Causes

CTESTS=*color*| *test-color-list*	specifies color for labels used to identify points where test is positive
CZONES=*color*	specifies color for lines and labels delineating zones A, B, and C
LABELFONT=*font*	specifies software font for labels at points where test is positive (alias for the TESTFONT= option)
LABELHEIGHT=*value*	specifies height of labels at points where test is positive (alias for the TESTHEIGHT= option)
LTESTS=*linetype*	specifies type of line connecting points where test is positive
LZONES=*linetype*	specifies line type for lines delineating zones A, B, and C
TESTFONT=*font*	specifies software font for labels at points where test is positive
TESTHEIGHT=*value*	specifies height of labels at points where test is positive

Table 43.4. Line Printer Options for Displaying Tests for Special Causes

TESTCHAR='*character*'	specifies character for line segments that connect any sequence of points for which a test for special causes is positive
ZONECHAR='*character*'	specifies character for lines that delineate zones for tests for special causes

Table 43.5. Reference Line Options

CHREF=*color*	specifies color for HREF= and HREF2= lines
CVREF=*color*	specifies color for VREF= and VREF2= lines
HREF=*values*| *SAS-data-set*	specifies position of reference lines perpendicular to horizontal axis on $\bar{X}$ chart
HREF2=*values*| *SAS-data-set*	specifies position of reference lines perpendicular to horizontal axis on R chart
HREFCHAR='*character*'	specifies line character for HREF= and HREF2= lines
HREFDATA= *SAS-data-set*	specifies position of reference lines perpendicular to horizontal axis on $\bar{X}$ chart
HREF2DATA= *SAS-data-set*	specifies position of reference lines perpendicular to horizontal axis on R chart
HREFLABELS= *'label1'...'labeln'*	specifies labels for HREF= lines
HREF2LABELS= *'label1'...'labeln'*	specifies labels for HREF2= lines
HREFLABPOS=*n*	specifies position of HREFLABELS= and HREF2LABELS= labels
LHREF=*linetype*	specifies line type for HREF= and HREF2= lines
LVREF=*linetype*	specifies line type for VREF= and VREF2= lines
NOBYREF	specifies that reference line information in a data set applies uniformly to charts created for all BY groups
VREF=*values*| *SAS-data-set*	specifies position of reference lines perpendicular to vertical axis on $\bar{X}$ chart
VREF2=*values*| *SAS-data-set*	specifies position of reference lines perpendicular to vertical axis on R chart
VREFCHAR='*character*'	specifies line character for VREF= and VREF2= lines
VREFLABELS= *'label1'...'labeln'*	specifies labels for VREF= lines
VREF2LABELS= *'label1'...'labeln'*	specifies labels for VREF2= lines
VREFLABPOS=*n*	specifies position of VREFLABELS= and VREF2LABELS= labels

Table 43.6. Block Variable Legend Options

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

Table 43.7. Axis and Axis Label Options

CAXIS=*color*	specifies color for axis lines and tick marks
CFRAME=*color*|(*color-list*)	specifies fill colors for frame for plot area
CTEXT=*color*	specifies color for tick mark values and axis labels
HAXIS=*values*|AXISn	specifies major tick mark values for horizontal axis
HEIGHT=*value*	specifies height of axis label and axis legend text
HMINOR=*n*	specifies number of minor tick marks between major tick marks on horizontal axis
HOFFSET=*value*	specifies length of offset at both ends of horizontal axis
INTSTART=*value*	specifies first major tick mark value for numeric horizontal axis
NOHLABEL	suppresses label for horizontal axis
NOTICKREP	specifies that only the first occurrence of repeated, adjacent subgroup values is to be labeled on horizontal axis
NOTRUNC	suppresses vertical axis truncation at zero applied by default to R chart
NOVANGLE	requests vertical axis labels that are strung out vertically
SKIPHLABELS=*n*	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=*values*|AXISn	specifies major tick mark values for vertical axis of $\bar{X}$ chart
VAXIS2=*values*|AXISn	specifies major tick mark values for vertical axis of R chart
VMINOR=n	specifies number of minor tick marks between major tick marks on vertical axis
VOFFSET=*value*	specifies length of offset at both ends of vertical axis
VZERO	forces origin to be included in vertical axis for primary chart
VZERO2	forces origin to be included in vertical axis for secondary chart
WAXIS=*n*	specifies width of axis lines

Table 43.8. Specification Limit Options

CIINDICES=(ALPHA=*value* TYPE=*keyword*)	specifies α value and type for computing capability index confidence limits
LSL=*value-list*	specifies list of lower specification limits
TARGET=*value-list*	specifies list of target values
USL=*value-list*	specifies list of upper specification limits

Table 43.9. Options for Specifying Control Limits

ALPHA=*value*	requests probability limits for control charts
LIMITN=*n*|VARYING	specifies either nominal sample size for fixed control limits or varying limits
NOREADLIMITS	computes control limits for each *process* from the data rather than from a LIMITS= data set (Release 6.10 and later releases)
READALPHA	reads _ALPHA_ instead of _SIGMAS_ from a LIMITS= data set
READINDEXES=ALL| '*label1*'...'*labeln*'	reads multiple sets of control limits for each *process* from a LIMITS= data set
READLIMITS	reads single set of control limits for each *process* from a LIMITS= data set (Release 6.09 and earlier releases)
SIGMAS=k	specifies width of control limits in terms of multiple k of standard error of plotted means and ranges

Table 43.10. Graphical Enhancement Options

ANNOTATE=*SAS-data-set*	specifies annotate data set that adds features to $\bar{\bar{X}}$ chart
ANNOTATE2=*SAS-data-set*	specifies annotate data set that adds features to R chart
DESCRIPTION='*string*'	specifies string that appears in the description field of the PROC GREPLAY master menu for $\bar{\bar{X}}$ chart
DESCRIPTION2='*string*'	specifies string that appears in the description field of the PROC GREPLAY master menu for R chart
FONT=*font*	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 $\bar{\bar{X}}$ chart
NAME2='*string*'	specifies name that appears in the name field of the PROC GREPLAY master menu for R chart
PAGENUM='*string*'	specifies the form of the label used in pagination
PAGENUMPOS= *keyword*	specifies the position of the page number requested with the PAGENUM= option

Table 43.11. Options for Displaying Control Limits

CINFILL=*color*	specifies color for area inside control limits
CLIMITS=*color*	specifies color of control limits, central line, and related labels
LCLLABEL='*label*'	specifies label for lower control limit on $\bar{X}$ chart
LCLLABEL2='*label*'	specifies label for lower control limit on R chart
LIMLABSUBCHAR='*character*'	specifies a substitution character for labels provided as quoted strings; the character is replaced with the value of the control limit
LLIMITS=*linetype*	specifies line type for control limits
NDECIMAL=*n*	specifies number of digits to right of decimal place in default labels for control limits and central line on $\bar{X}$ chart
NDECIMAL2=*n*	specifies number of digits to right of decimal place in default labels for control limits and central line on R chart
NOCTL	suppresses display of central line on $\bar{X}$ chart
NOCTL2	suppresses display of central line on R chart
NOLCL	suppresses display of lower control limit on $\bar{X}$ chart
NOLCL2	suppresses display of lower control limit on R chart
NOLIMITLABEL	suppresses labels for control limits and central lines
NOLIMITS	suppresses display of control limits
NOLIMITSFRAME	suppresses default frame around control limit information when multiple sets of control limits are read from a LIMITS= data set
NOLIMITSLEGEND	suppresses legend for control limits
NOLIMIT0	suppresses display of zero lower control limit on R chart
NOUCL	suppresses display of upper control limit on $\bar{X}$ chart
NOUCL2	suppresses display of upper control limit on R chart
RSYMBOL='*string*'|*keyword*	specifies label for central line on R chart
UCLLABEL='*string*'	specifies label for upper control limit on $\bar{X}$ chart
UCLLABEL2='*string*'	specifies label for upper control limit on R chart
WLIMITS=*n*	specifies width for control limits and central line
XSYMBOL='*string*'|*keyword*	specifies label for central line on $\bar{X}$ chart

Table 43.12. Grid Options

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

Table 43.13. Options for Plotting and Labeling Points

ALLLABEL=VALUE|(*variable*)	labels every point on $\bar{X}$ chart
ALLLABEL2=VALUE|(*variable*)	labels every point on R chart
CCONNECT=*color*	specifies color for line segments that connect points on chart
CFRAMELAB=*color*	specifies fill color for frame around labeled points
CNEEDLES=*color*	specifies color for needles that connect points to central line
CONNECTCHAR='*character*'	specifies character used to form line segments that connect points on chart
COUT=*color*	specifies color for portions of line segments that connect points outside control limits
COUTFILL=*color*	specifies color for shading areas between the connected points and control limits outside the limits
NEEDLES	connects points to central line with vertical needles
NOCONNECT	suppresses line segments that connect points on chart
OUTLABEL=VALUE|(*variable*)	labels points outside control limits on $\bar{X}$ chart
OUTLABEL2=VALUE|(*variable*)	labels points outside control limits on R chart
SYMBOLCHARS='*characters*'	specifies characters indicating *symbol-variable*
SYMBOLLEGEND=NONE|*name*	specifies LEGEND statement for levels of *symbol-variable*
SYMBOLORDER=*keyword*	specifies order in which symbols are assigned for levels of *symbol-variable*
TURNALL|TURNOUT	turns point labels so that they are strung out vertically

Table 43.14. Clipping Options

CCLIP=*color*	specifies color for plot symbol for clipped points
CLIPCHAR='*character*'	specifies plot character for clipped points
CLIPFACTOR=*value*	determines extent to which extreme points are clipped
CLIPLEGEND='*string*'	specifies text for clipping legend
CLIPLEGPOS=*keyword*	specifies position of clipping legend
CLIPSUBCHAR='*character*'	specifies substitution character for CLIPLEGEND= text
CLIPSYMBOL=*symbol*	specifies plot symbol for clipped points
CLIPSYMBOLHT=*value*	specifies symbol marker height for clipped points

Table 43.15. Phase Options

CPHASEBOX=*color*	specifies color for box enclosing all plotted points for a phase
CPHASEBOX-CONNECT=*color*	specifies color for line segments connecting adjacent enclosing boxes
CPHASEBOXFILL=*color*	specifies fill color for box enclosing all plotted points for a phase
CPHASELEG=*color*	specifies text color for *phase* legend
CPHASEMEAN-CONNECT=*color*	specifies color for line segments connecting average value points within a phase
NOPHASEFRAME	suppresses default frame for *phase* legend
OUTPHASE='*string*'	specifies value of _PHASE_ in the OUTHISTORY= data set
PHASEBREAK	disconnects last point in a *phase* from first point in next *phase*
PHASELABTYPE=*value*|*keyword*	specifies text size of *phase* legend
PHASELEGEND	displays *phase* labels in a legend across top of charts
PHASELIMITS	labels control limits for each phase, provided they are constant within that phase
PHASEMEANSYMBOL=*symbol*	specifies symbol marker for average of values within a phase
PHASEREF	delineates *phases* with vertical reference lines
READPHASES= ALL|'*label1*'...'*labeln*'	specifies *phases* to be read from an input data set

Table 43.16. Input Data Set Options

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

Table 43.17. Output Data Set Options

OUTHISTORY=*SAS-data-set*	creates output data set containing subgroup summary statistics
OUTINDEX='*string*'	specifies value of _INDEX_ in the OUTLIMITS= data set
OUTLIMITS=*SAS-data-set*	creates output data set containing control limits
OUTTABLE=*SAS-data-set*	creates output data set containing subgroup summary statistics and control limits

Table 43.18. Process Mean and Standard Deviation Options

MU0=*value*	specifies known (standard) value μ_0 for process mean μ
SIGMA0=*value*	specifies known (standard) value σ_0 for process standard deviation σ
SMETHOD=*keyword*	specifies method for estimating process standard deviation σ
TYPE=*keyword*	identifies parameters as estimates or standard values and specifies value of _TYPE_ in the OUTLIMITS= data set

Table 43.19. Plot Layout Options

ALLN	plots means and ranges for all subgroups
BILEVEL	creates control charts using half-screens and half-pages
EXCHART	creates control charts for a *process* only when exceptions occur
INTERVAL=*keyword*	specifies natural time interval between consecutive subgroup positions when time, date, or datetime format is associated with a numeric subgroup variable
MAXPANELS=*n*	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 charts
NOCHART2	suppresses creation of R chart
NOFRAME	suppresses frame for plot area
NOLEGEND	suppresses legend for subgroup sample sizes
NPANELPOS=*n*	specifies number of subgroup positions per panel on each chart
REPEAT	repeats last subgroup position on panel as first subgroup position of next panel
SEPARATE	displays $\bar{X}$ and R charts on separate screens or pages
TOTPANELS=*n*	specifies number of pages or screens to be used to display chart
YPCT1=*value*	specifies length of vertical axis on $\bar{X}$ chart as a percentage of sum of lengths of vertical axes for $\bar{X}$ and R charts
ZEROSTD	displays $\bar{X}$ and R charts regardless of whether $\hat{\sigma} = 0$

Table 43.20. Options for Interactive Control Charts

HTML=(*variable*)	specifies a variable whose values are URLs to be associated with subgroups
HTML_LEGEND=(*variable*)	specifies a variable whose values are URLs to be associated with symbols in the symbol legend
TESTURLS=*SAS-data-set*	associates URLs with tests for special causes
WEBOUT=*SAS-data-set*	creates an OUTTABLE= data set with additional graphics coordinate data

Table 43.21. Star Options

CSTARCIRCLES=*color*	specifies color for STARCIRCLES= circles
CSTARFILL=*color*| *(variable)*	specifies color for filling stars
CSTAROUT=*color*	specifies outline color for stars exceeding inner or outer circles
CSTARS=*color*| *(variable)*	specifies color for outlines of stars
LSTARCIRCLES= *linetypes*	specifies line types for STARCIRCLES= circles
LSTARS=*linetype*|*(variable)*	specifies line types for outlines of STARVERTICES= stars
STARBDRADIUS=*value*	specifies radius of outer bound circle for vertices of stars
STARCIRCLES=*value-list*	specifies reference circles for stars
STARINRADIUS=*value*	specifies inner radius of stars
STARLABEL=*keyword*	specifies vertices to be labeled
STARLEGEND=*keyword*	specifies style of legend for star vertices
STARLEGENDLAB=*'label'*	specifies label for STARLEGEND= legend
STAROUTRADIUS=*value*	specifies outer radius of stars
STARSPEC=*value*| *SAS-data-set*	specifies method used to standardize vertex variables
STARSTART=*value*	specifies angle for first vertex
STARTYPE=*keyword*	specifies graphical style of star
STARVERTICES= *variable*|*(variables)*	superimposes star at each point on $\bar{X}$ chart
WSTARCIRCLES=*n*	specifies width of STARCIRCLES= circles
WSTARS=*n*	specifies width of STARVERTICES= stars

Details

Constructing Charts for Means and Ranges

The following notation is used in this section:

μ	process mean (expected value of the population of measurements)
σ	process standard deviation (standard deviation of the population of measurements)
$\bar{X}_i$	mean of measurements in i^{th} subgroup
R_i	range of measurements in i^{th} subgroup
n_i	sample size of i^{th} subgroup
N	number of subgroups
$\overline{\overline{X}}$	weighted average of subgroup means
$d_2(n)$	expected value of the range of n independent normally distributed variables with unit standard deviation
$d_3(n)$	standard error of the range of n independent observations from a normal population with unit standard deviation
z_p	$100p^{\text{th}}$ percentile of the standard normal distribution
$D_p(n)$	$100p^{\text{th}}$ percentile of the distribution of the range of n independent observations from a normal population with unit standard deviation

Plotted Points

Each point on the $\bar{X}$ chart indicates the value of a subgroup mean ($\bar{X}_i$). For example, if the tenth subgroup contains the values 12, 15, 19, 16, and 14, the mean plotted for this subgroup is

$$\bar{X}_{10} = \frac{12 + 15 + 19 + 16 + 14}{5} = 15.2$$

Each point on the R chart indicates the value of a subgroup range (R_i). For example, the range plotted for the tenth subgroup is $R_{10} = 19 - 12 = 7$.

Central Lines

On an $\bar{X}$ chart, by default, the central line indicates an estimate of μ, which is computed as

$$\hat{\mu} = \overline{\overline{X}} = \frac{n_1\bar{X}_1 + \cdots + n_N\bar{X}_N}{n_1 + \cdots + n_N}$$

If you specify a known value (μ_0) for μ, the central line indicates the value of μ_0.

On an R chart, by default, the central line for the i^{th} subgroup indicates an estimate for the expected value of R_i, which is computed as $d_2(n_i)\hat{\sigma}$, where $\hat{\sigma}$ is an estimate of σ. If you specify a known value (σ_0) for σ, the central line indicates the value of $d_2(n_i)\sigma_0$. Note that the central line varies with n_i.

Control Limits

You can compute the limits in the following ways:

- as a specified multiple (k) of the standard errors of $\bar{X}_i$ and R_i above and below the central line. The default limits are computed with $k = 3$ (these are referred to as 3σ *limits*).
- as probability limits defined in terms of α, a specified probability that $\bar{X}_i$ or R_i exceeds the limits

The following table provides the formulas for the limits:

Table 43.22. Limits for $\bar{X}$ and R Charts

Control Limits	
$\bar{X}$ Chart	LCL = lower limit = $\overline{\overline{X}} - k\hat{\sigma}/\sqrt{n_i}$ UCL = upper limit = $\overline{\overline{X}} + k\hat{\sigma}/\sqrt{n_i}$
R Chart	LCL = lower limit = $\max(d_2(n_i)\hat{\sigma} - kd_3(n_i)\hat{\sigma}, 0)$ UCL = upper limit = $d_2(n_i)\hat{\sigma} + kd_3(n_i)\hat{\sigma}$

Probability Limits	
$\bar{X}$ Chart	LCL = lower limit = $\overline{\overline{X}} - z_{\alpha/2}(\hat{\sigma}/\sqrt{n_i})$ UCL = upper limit = $\overline{\overline{X}} + z_{\alpha/2}(\hat{\sigma}/\sqrt{n_i})$
R Chart	LCL = lower limit = $D_{\alpha/2}\hat{\sigma}$ UCL = upper limit = $D_{1-\alpha/2}\hat{\sigma}$

The formulas for R charts assume that the data are normally distributed. If standard values μ_0 and σ_0 are available for μ and σ, respectively, replace $\overline{\overline{X}}$ with μ_0 and $\hat{\sigma}$ with σ_0 in Table 43.22. Note that the limits vary with n_i and that the probability limits for R_i are asymmetric around the central line.

You can specify parameters for the limits as follows:

- Specify k with the SIGMAS= option or with the variable _SIGMAS_ in a LIMITS= data set.
- Specify α with the ALPHA= option or with the variable _ALPHA_ in a LIMITS= data set.
- Specify a constant nominal sample size $n_i \equiv n$ for the control limits with the LIMITN= option or with the variable _LIMITN_ in a LIMITS= data set.
- Specify μ_0 with the MU0= option or with the variable _MEAN_ in a LIMITS= data set.
- Specify σ_0 with the SIGMA0= option or with the variable _STDDEV_ in a LIMITS= data set.

Output Data Sets

OUTLIMITS= Data Set

The OUTLIMITS= data set saves control limits and control limit parameters. The following variables are saved:

Table 43.23. OUTLIMITS= Data Set

Variable	Description
ALPHA	probability (α) of exceeding limits
_CP_	capability index C_p
CPK	capability index C_{pk}
CPL	capability index CPL
CPM	capability index C_{pm}
CPU	capability index CPU
INDEX	optional identifier for the control limits specified with the OUTINDEX= option
LCLR	lower control limit for subgroup range
LCLX	lower control limit for subgroup mean
LIMITN	nominal sample size associated with the control limits
LSL	lower specification limit
MEAN	process mean ($\overline{\overline{X}}$ or μ_0)
_R_	value of central line on R chart
SIGMAS	multiple (k) of standard error of $\bar{X}_i$ or R_i
STDDEV	process standard deviation ($\hat{\sigma}$ or σ_0)
SUBGRP	*subgroup-variable* specified in the XRCHART statement
TARGET	target value
TYPE	type (estimate or standard value) of _MEAN_ and _STDDEV_
UCLR	upper control limit for subgroup range
UCLX	upper control limit for subgroup mean
USL	upper specification limit
VAR	*process* specified in the XRCHART statement

Notes:

1. If the control limits vary with subgroup sample size, the special missing value *V* is assigned to the variables _LIMITN_, _LCLX_, _UCLX_, _LCLR_, _R_, and _UCLR_.

2. If the limits are defined in terms of a multiple k of the standard errors of $\bar{X}_i$ and R_i, the value of _ALPHA_ is computed as $\alpha = 2(1 - \Phi(k))$, where $\Phi(\cdot)$ is the standard normal distribution function.

3. If the limits are probability limits, the value of _SIGMAS_ is computed as $k = \Phi^{-1}(1 - \alpha/2)$, where Φ^{-1} is the inverse standard normal distribution function.

4. The variables _CP_, _CPK_, _CPL_, _CPU_, _LSL_, and _USL_ are included only if you provide specification limits with the LSL= and USL= options. The variables _CPM_ and _TARGET_ are included if, in addition, you

provide a target value with the TARGET= option. See “Capability Indices” on page 1537 for computational details.

5. Optional BY variables are saved in the OUTLIMITS= data set.

The OUTLIMITS= data set contains one observation for each *process* specified in the XRCHART statement. For an example, see “Saving Control Limits” on page 1508.

OUTHISTORY= Data Set

The OUTHISTORY= data set saves subgroup summary statistics. The following variables are saved:

- the *subgroup-variable*
- a subgroup mean variable named by *process* suffixed with *X*
- a subgroup range variable named by *process* suffixed with *R*
- a subgroup sample size variable named by *process* suffixed with *N*

Given a *process* name that contains eight characters, the procedure first shortens the name to its first four characters and its last three characters, and then it adds the suffix. For example, the procedure shortens the *process* DIAMETER to DIAMTER before adding the suffix.

Variables containing subgroup means, ranges, and sample sizes are created for each *process* specified in the XRCHART statement. For example, consider the following statements:

```
proc shewhart data=steel;
   xrchart (width diameter)*lot / outhistory=summary;
run;
```

The data set SUMMARY contains variables named LOT, WIDTHX, WIDTHR, WIDTHN, DIAMTERX, DIAMTERR, and DIAMTERN.

Additionally, the following variables, if specified, are included:

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

For an example of an OUTHISTORY= data set, see “Saving Summary Statistics” on page 1507.

OUTTABLE= Data Set

The OUTTABLE= data set saves subgroup summary statistics, control limits, and related information. The following variables are saved:

Variable	Description
ALPHA	probability (α) of exceeding control limits
EXLIM	control limit exceeded on $\bar{X}$ chart
EXLIMR	control limit exceeded on R chart
LCLR	lower control limit for range
LCLX	lower control limit for mean
LIMITN	nominal sample size associated with the control limits
MEAN	process mean
_R_	average range
SIGMAS	multiple (k) of the standard error associated with control limits
subgroup	values of the subgroup variable
SUBN	subgroup sample size
SUBR	subgroup range
SUBX	subgroup mean
TESTS	tests for special causes signaled on $\bar{X}$ chart
TESTS2	tests for special causes signaled on R chart
UCLR	upper control limit for range
UCLX	upper control limit for mean
VAR	*process* specified in the XRCHART statement

In addition, the following variables, if specified, are included:

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

Notes:

1. Either the variable _ALPHA_ or the variable _SIGMAS_ is saved, depending on how the control limits are defined (with the ALPHA= or SIGMAS= options, respectively, or with the corresponding variables in a LIMITS= data set).

2. The variable _TESTS_ is saved if you specify the TESTS= option. The k^{th} character of a value of _TESTS_ is k if Test k is positive at that subgroup. For example, if you request all eight tests and Tests 2 and 8 are positive for a given subgroup, the value of _TESTS_ has a 2 for the second character, an 8 for the eighth character, and blanks for the other six characters.

3. The variable _TESTS2_ is saved if you specify the TESTS2= option.

4. The variables _VAR_, _EXLIM_, _EXLIMR_, _TESTS_, and _TESTS2_ are character variables of length 8. The variable _PHASE_ is a character variable of length 16. All other variables are numeric.

For an example, see “Saving Control Limits” on page 1508.

ODS Tables

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

Table 43.24. ODS Tables Produced with the XRCHART Statement

Table Name	Description	Options
XRCHART	$\bar{X}$ and R chart summary statistics	TABLE, TABLEALL, TABLEC, TABLEID, TABLELEG, TABLEOUT, TABLETESTS
Tests	descriptions of tests for special causes requested with the TESTS= option for which at least one positive signal is found	TABLEALL, TABLELEG

Input Data Sets

DATA= Data Set

You can read raw data (process measurements) from a DATA= data set specified in the PROC SHEWHART statement. Each *process* specified in the XRCHART statement must be a SAS variable in the DATA= data set. This variable provides measurements which must be grouped into subgroup samples indexed by the *subgroup-variable*. The *subgroup-variable*, which is specified in the XRCHART statement, must also be a SAS variable in the DATA= data set. Each observation in a DATA= data set must contain a value 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 subgroup contains five items and there are 30 subgroup samples, the DATA= data set should contain 150 observations.

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

By default, the SHEWHART procedure reads all 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 (for an example, see “Displaying Stratification in Phases” on page 1689).

For an example of a DATA= data set, see “Creating Charts for Means and Ranges from Raw Data” on page 1502.

LIMITS= Data Set

You can read preestablished control limits (or parameters from which the control limits can be calculated) from a LIMITS= data set specified in the PROC SHEWHART statement. For example, the following statements read control limit information from the data set CONLIMS:*

```
proc shewhart data=info limits=conlims;
   xrchart weight*batch;
run;
```

The LIMITS= data set can be an OUTLIMITS= data set that was created in a previous run of the SHEWHART procedure. Such data sets always contain the variables required for a LIMITS= data set. The LIMITS= data set can also be created directly using a DATA step. When you create a LIMITS= data set, you must provide one of the following:

- the variables _LCLX_, _MEAN_, _UCLX_, _LCLR_, _R_, and _UCLR_, which specify the control limits directly
- the variables _MEAN_ and _STDDEV_, which are used to calculate the control limits according to the equations in Table 43.22 on page 1526

In addition, note the following:

- The variables _VAR_ and _SUBGRP_ are required. These must be character variables of length 8.
- The variable _INDEX_ is required if you specify the READINDEX= option; this must be a character variable of length 16.
- The variables _LIMITN_, _SIGMAS_ (or _ALPHA_), and _TYPE_ are optional, but they are recommended to maintain a complete set of control limit information. The variable _TYPE_ must be a character variable of length 8; valid values are `ESTIMATE`, `STANDARD`, `STDMU`, and `STDSIGMA`.
- BY variables are required if specified with a BY statement.

For an example, see “Reading Preestablished Control Limits” on page 1512.

HISTORY= Data Set

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

A HISTORY= data set used with the XRCHART statement must contain the following variables:

*In Release 6.09 and in earlier releases, it is necessary to specify the READLIMITS option.

- the *subgroup-variable*
- a subgroup mean variable for each *process*
- a subgroup range variable for each *process*
- a subgroup sample size variable for each *process*

The names of the subgroup mean, subgroup range, and subgroup sample size variables must be the *process* name concatenated with the special suffix characters X, R, and N, respectively.

For example, consider the following statements:

```
proc shewhart history=summary;
   xrchart (weight yldstren)*batch;
run;
```

The data set SUMMARY must include the variables BATCH, WEIGHTX, WEIGHTR, WEIGHTN, YLDSRENX, YLDSRENR, and YLDSRENN.

Note that if you specify a *process* name that contains eight characters, the names of the summary variables must be formed from the first four characters and the last three characters of the *process* name, suffixed with the appropriate character.

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 SHEWHART procedure reads all the observations in a HISTORY= data set. However, if the data set includes the variable _PHASE_, you can read selected groups of observations (referred to as *phases*) by specifying the READPHASES= option (see “Displaying Stratification in Phases” on page 1689 for an example).

For an example of a HISTORY= data set, see “Creating Charts for Means and Ranges from Summary Data” on page 1505.

TABLE= Data Set

You can read summary statistics and control limits from a TABLE= data set specified in the PROC SHEWHART statement. This enables you to reuse an OUTTABLE= data set created in a previous run of the SHEWHART procedure or to read data sets created by other SAS procedures. Because the SHEWHART procedure simply displays the information read from a TABLE= data set, you can use TABLE= data sets to create specialized control charts. xamples are provided in Chapter 49, “Specialized Control Charts.”

The following table lists the variables required in a TABLE= data set used with the XRCHART statement:

Table 43.25. Variables Required in a TABLE= Data Set

Variable	Description
LCLR	lower control limit for range
LCLX	lower control limit for mean
LIMITN	nominal sample size associated with the control limits
MEAN	process mean
_R_	average range
subgroup-variable	values of the *subgroup-variable*
SUBN	subgroup sample size
SUBR	subgroup range
SUBX	subgroup mean
UCLR	upper control limit for range
UCLX	upper control limit for mean

Other variables that can be read from a TABLE= data set include

- *block-variables*
- *symbol-variable*
- BY variables
- ID variables
- _PHASE_ (if the READPHASES= option is specified). This variable must be a character variable of length 16.
- _TESTS_ (if the TESTS= option is specified). This variable is used to flag tests for special causes for subgroup means and must be a character variable of length 8.
- _TESTS2_ (if the TESTS2= option is specified). This variable is used to flag tests for special causes for subgroup ranges and must be a character variable of length 8.
- _VAR_. This variable is required if more than one *process* is specified or if the data set contains information for more than one *process*. This variable must be a character variable of length 8.

For an example of a TABLE= data set, see “Saving Control Limits” on page 1508.

Subgroup Variables

The values of the *subgroup-variable*, which is specified in the chart statement, indicate how the observations in the input data set (a DATA=, HISTORY=, or TABLE= data set) are arranged into rational subgroups.* Typically, the values of the *subgroup-variable* are one of the following:

- *indices* that give the order in which subgroup samples were collected (for example, 1, 2, 3, . . .). An unformatted numeric *subgroup-variable* is appropriate for this situation. For an example using this type of *subgroup-variable*, see "Creating Charts for Means and Ranges from Raw Data" on page 1502.
- the *dates* or *times* at which subgroup samples were collected (for example, 01JUN, 02JUN, 03JUN, . . .). A numeric *subgroup-variable* with a SAS date, time, or datetime format is appropriate for this situation. You can optionally associate a format with the *subgroup-variable* by using a FORMAT statement; refer to *SAS Language Reference: Dictionary* for details. For an example using this type of *subgroup-variable*, see Example 43.3 on page 1544.
- *labels* that uniquely identify subgroup samples (for example, LOT39, LOTX12, LOT43A). A character *subgroup-variable* (with or without a format) is appropriate for this situation. For an example using this type of *subgroup-variable*, see Example 43.1 on page 1540.

The values of the *subgroup-variable* also determine how the horizontal axis of the control chart is scaled and labeled.

The notion of a rational subgroup is fundamental to the application of a Shewhart chart. You should select your subgroups so that if special causes of variation are present, the opportunity for variation within subgroups is minimized while the opportunity for variation between subgroups is maximized. In other words, the conditions within a subgroup should be homogeneous. The reason for this requirement is that the construction of the control limits is based on within-subgroup variability. Refer to Montgomery (1996) and Wheeler and Chambers (1986) for approaches to rational subgrouping.

The selection of subgroups is both a practical and a statistical issue that requires knowledge of the process and the sampling or measurement procedure. The values of the subgroup-variable should reflect the selection of subgroups and should not be assigned arbitrarily. Incorrect subgrouping or assignment of subgroup-variable values can result in control limits that are too tight or too wide.

If the input data set is a HISTORY= or TABLE= data set, each observation represents a distinct subgroup, and, consequently, the observations within each BY group must have distinct subgroup variable values. Similarly, if the input data set is a DATA= data set and you are using the CCHART, IRCHART, NPCHART, PCHART, or UCHART

*This discussion also applies to the use of *subgroup-variables* in the CUSUM procedure and the MACONTROL procedure.

statement, each observation represents a distinct subgroup, and, consequently, the observations within each BY group must have distinct subgroup variable values. However, if the input data set is a DATA= data set and you are using the BOXCHART, MCHART, MRCHART, RCHART, SCHART, XCHART, XRCHART, or XSCHART statement, subgroups are identified by groups of consecutive observations with identical values of the subgroup-variable.

The order of the observations in the input data set and the scaling of the horizontal axis depend on the type of the subgroup-variable, which can be numeric or character.

Numeric Subgroup Variables

If the subgroup-variable is numeric, the observations must be sorted in increasing order of the values of the subgroup variable. If you use a BY statement, first sort by the BY variables and then by the subgroup variable.

The unformatted values of the subgroup-variable are used to scale the horizontal axis of the control chart, and the formatted values are used to label the major tick marks on the horizontal axis. As a result, the horizontal distance between two points corresponding to consecutive subgroups is proportional to the difference between their unformatted subgroup values.

If a DATE, DATETIME, WEEKDATE, or WORDDATE format is associated with the subgroup variable, the major tick mark labels are split and displayed in two levels to save space. You can override this default with the TURNHLABELS option (which turns the labels vertically) or with tick label options in an AXISn statement specified with the HAXIS= option.

Character Subgroup Variables

If the subgroup-variable is numeric, the order of the observations is not checked. The horizontal axis is scaled so that the subgroups are spaced uniformly. Formatted subgroup variable values are used to label the major tick marks.

You can use a character subgroup variable to avoid gaps between groups of points or time values on a control chart. You can also use a character subgroup variable to create a chart in which the order of the points depends only on the order in which the subgroups are arranged in the input data set.

You should verify the order of the observations in the input data set before using a character subgroup variable in conjunction with the TESTS= option. With the exception of Test 1, the tests for special causes are applicable only if the subgroups are provided in chronological order. See Chapter 48, “Tests for Special Causes,” for details.

To avoid collision of adjacent tick labels on the horizontal axis, the labels are thinned by default. You can override this default with the TURNHLABELS option or with tick label options in an AXISn statement specified with the HAXIS= option.

Methods for Estimating the Standard Deviation

When control limits are determined from the input data, two methods (referred to as default and MVLUE) are available for estimating σ.

Default Method

The default estimate for σ is

$$\hat{\sigma} = \frac{R_1/d_2(n_1) + \cdots + R_N/d_2(n_N)}{N}$$

where N is the number of subgroups for which $n_i \geq 2$, and R_i is the sample range of the observations $x_{i1}, \ldots, x_{in_i}$ in the i^{th} subgroup.

$$R_i = \max_{1 \leq j \leq n_i}(x_{ij}) - \min_{1 \leq j \leq n_i}(x_{ij})$$

A subgroup range R_i is included in the calculation only if $n_i \geq 2$. The unbiasing factor $d_2(n_i)$ is defined so that, if the observations are normally distributed, the expected value of R_i is $d_2(n_i)\sigma$. 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* (1976).

MVLUE Method

If you specify SMETHOD=MVLUE, a minimum variance linear unbiased estimate (MVLUE) is computed for σ. Refer to Burr (1969, 1976) and Nelson (1989, 1994). The MVLUE is a weighted average of N unbiased estimates of σ of the form $R_i/d_2(n_i)$, and it is computed as

$$\hat{\sigma} = \frac{f_1 R_1/d_2(n_1) + \cdots + f_N R_N/d_2(n_N)}{f_1 + \cdots + f_N}$$

where

$$f_i = \frac{[d_2(n_i)]^2}{[d_3(n_i)]^2}$$

A subgroup range R_i is included in the calculation only if $n_i \geq 2$, and N is the number of subgroups for which $n_i \geq 2$. The unbiasing factor $d_3(n_i)$ is defined so that, if the observations are normally distributed, the expected value of σ_{R_i} is $d_3(n_i)\sigma$. The MVLUE assigns greater weight to estimates of σ from subgroups with larger sample sizes, and it is intended for situations where the subgroup sample sizes vary. If the subgroup sample sizes are constant, the MVLUE reduces to the default estimate.

Capability Indices

This section provides formulas for process capability indices, which are saved in the OUTLIMITS= data set when you use the LSL= and USL= options to provide lower and upper specification limits (LSL and USL, respectively) for the *process*. The estimate $\hat{\sigma}$ is computed as described in the previous section, "Methods for Estimating the Standard Deviation."

The Index Cp

The process capability index C_p is computed as

$$C_p = (USL - LSL)/6\hat{\sigma}$$

If you do not specify both LSL and USL, the variable _CP_ is assigned a missing value.

The Index CPL

The process capability index *CPL* is computed as

$$CPL = (\overline{\overline{X}} - LSL)/3\hat{\sigma}$$

If you do not specify LSL, the variable _CPL_ is assigned a missing value.

The Index CPU

The process capability index *CPU* is computed as

$$CPU = (USL - \overline{\overline{X}})/3\hat{\sigma}$$

If you do not specify USL, the variable _CPU_ is assigned a missing value.

The Index Cpk

The process capability index C_{pk} is computed as

$$C_{pk} = \min(USL - \overline{\overline{X}}, \overline{\overline{X}} - LSL)/3\hat{\sigma}$$

If you specify only USL, the index C_{pk} is computed as

$$C_{pk} = (USL - \overline{\overline{X}})/3\hat{\sigma}$$

and if you specify only LSL, the index C_{pk} is computed as

$$C_{pk} = (\overline{\overline{X}} - LSL)/3\hat{\sigma}$$

The Index Cpm

The process capability index C_{pm} is computed as

$$C_{pm} = \frac{\min(T - LSL, USL - T)}{3\sqrt{\hat{\sigma}^2 + (\overline{\overline{X}} - T)^2}}$$

where T is the target value specified with the TARGET= option.

When a single specification limit (SL) and target are specified, C_{pm} is computed as

$$C_{pm} = \frac{|T - SL|}{3\sqrt{\hat{\sigma}^2 + (\bar{\bar{X}} - T)^2}}$$

You can also use the CAPABILITY procedure to compute a variety of capability indices. The SHEWHART procedure and the CAPABILITY procedure use the same formulas to calculate the indices, but they use different estimates for the process standard deviation σ.

- The SHEWHART procedure calculates $\hat{\sigma}$ from subgroup estimates of σ. For details, see the previous section, "Methods for Estimating the Standard Deviation."
- The CAPABILITY procedure calculates $\hat{\sigma}$ as the sample standard deviation of the entire sample. For details, see "Standard Deviation" on page 35.

Regardless of which method you use, you should verify that the process is in statistical control before interpreting the indices, and you should verify that the data are normally distributed. The CAPABILITY procedure provides a variety of statistical and graphical tests for checking normality.

Some references use different notation and names for capability indices. For example, the manual *Fundamental Statistical Process Control: Reference Manual* (1991) uses the term "process capability indices" for the indices listed in this section, and it uses the term "process performance indices" for the indices computed by the CAPABILITY procedure.

Axis Labels

You can specify axis labels by assigning labels to particular variables in the input data set, as summarized in the following table:

Axis	Input Data Set	Variable
Horizontal	all	*subgroup-variable*
Vertical ($\bar{X}$ chart)	DATA=	*process*
Vertical ($\bar{X}$ chart)	HISTORY=	subgroup mean variable
Vertical ($\bar{X}$ chart)	TABLE=	_SUBX_

You can specify distinct labels for the vertical axes of the $\bar{X}$ and R charts by breaking the vertical axis into two parts with a split character. Specify the split character with the SPLIT= option. The first part labels the vertical axis of the $\bar{X}$ chart, and the second part labels the vertical axis of the R chart.

For example, the following sets of statements specify the label *Avg Diameter in mm* for the vertical axis of the $\bar{X}$ chart and the label *Range in mm* for the vertical axis of the R chart:

```
proc shewhart data=wafers;
   xrchart diamtr*batch / split = '/' ;
   label diamtr = 'Avg Diameter in mm/Range in mm';
run;

proc shewhart history=wafersum;
   xrchart diamtr*batch / split = '/' ;
   label diamtrx = 'Avg Diameter in mm/Range in mm';
run;

proc shewhart table=wtable;
   xrchart diamtr*batch / split = '/' ;
   label _SUBX_ = 'Avg Diameter in mm/Range in mm';
run;
```

In this example, the label assignments are in effect only for the duration of the procedure step, and they temporarily override any permanent labels associated with the variables.

Missing Values

An observation read from a DATA=, HISTORY=, or TABLE= 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= or TABLE= data set is not analyzed if the values of any of the corresponding summary variables are missing.

Examples

The SHEWHART Procedure

This section provides advanced examples of the XRCHART statement.

Example 43.1. Applying Tests for Special Causes

See SHWXR2 in the SAS/QC Sample Library

This example illustrates how you can apply tests for special causes to make $\bar{X}$ and R charts more sensitive to special causes of variation.

The weight of a roll of tape is measured before and after an adhesive is applied. The difference in weight represents the amount of adhesive applied to the tape during the coating process. The following data set contains the average and the range of the adhesive amounts for 21 samples of five rolls:

```
data tape;
   input sample $ weightx weightr;
   weightn=5;
   label weightx = 'Average Adhesive Amount'
         sample  = 'Sample Code';
   datalines;
C9  1270  35
C4  1258  25
A7  1248  24
A1  1260  39
A5  1273  29
D3  1260  21
D6  1259  37
D1  1240  37
R4  1260  28
H7  1255  19
H2  1268  36
H6  1253  36
P4  1273  29
P9  1275  22
J7  1257  24
J2  1269  41
J3  1249  36
B2  1264  31
G4  1258  25
G6  1248  36
G3  1248  30
;
```

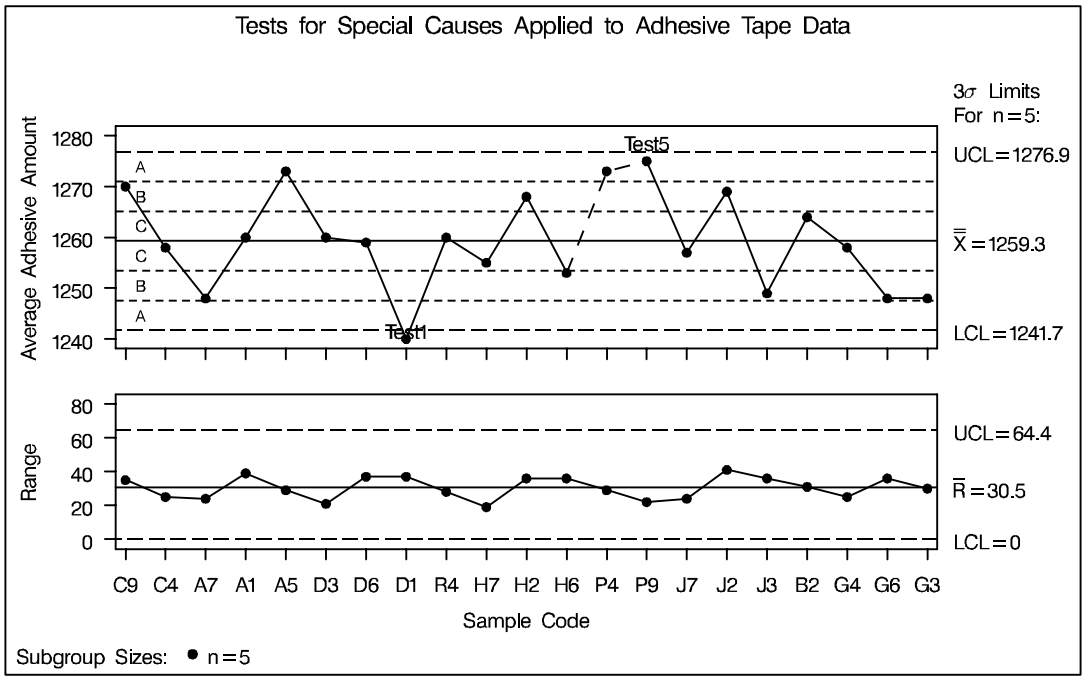
The following statements create $\bar{X}$ and R charts, apply several tests to the $\bar{X}$ chart, and tabulate the results:

```
title 'Tests for Special Causes Applied to Adhesive Tape Data';
symbol v=dot;
proc shewhart history=tape;
   xrchart weight*sample /
       tests  = 1 to 5
       tabletests
       zonelabels
       ltests = 20;
run;
```

The charts are shown in Output 43.1.1, and the table is shown in Output 43.1.2. The TESTS= option requests Tests 1, 2, 3, 4, and 5, which are described in Chapter 48, "Tests for Special Causes." The TABLETESTS option requests a basic table of subgroup statistics and control limits with a column indicating which subgroups tested positive for special causes.

The ZONELABELS option displays zone lines and zone labels on the $\bar{X}$ chart. The zones are used to define the tests. The LTESTS= option specifies the line type used to connect the points in a pattern for a test that is signaled.

Output 43.1.1. Tests for Special Causes Displayed on $\bar{X}$ and R Charts



Output 43.1.2. Tabular Form of $\bar{X}$ and R Charts

Tests for Special Causes Applied to Adhesive Tape Data

Means and Ranges Chart Summary for weight

sample	Subgroup Sample Size	---3 Sigma Limits with n=5 for Mean--- Lower Limit	Subgroup Mean	Upper Limit	Special Tests Signaled
C9	5	1241.7065	1270.0000	1276.8650	
C4	5	1241.7065	1258.0000	1276.8650	
A7	5	1241.7065	1248.0000	1276.8650	
A1	5	1241.7065	1260.0000	1276.8650	
A5	5	1241.7065	1273.0000	1276.8650	
D3	5	1241.7065	1260.0000	1276.8650	
D6	5	1241.7065	1259.0000	1276.8650	
D1	5	1241.7065	1240.0000	1276.8650	1
R4	5	1241.7065	1260.0000	1276.8650	
H7	5	1241.7065	1255.0000	1276.8650	
H2	5	1241.7065	1268.0000	1276.8650	
H6	5	1241.7065	1253.0000	1276.8650	
P4	5	1241.7065	1273.0000	1276.8650	
P9	5	1241.7065	1275.0000	1276.8650	5
J7	5	1241.7065	1257.0000	1276.8650	
J2	5	1241.7065	1269.0000	1276.8650	
J3	5	1241.7065	1249.0000	1276.8650	
B2	5	1241.7065	1264.0000	1276.8650	
G4	5	1241.7065	1258.0000	1276.8650	
G6	5	1241.7065	1248.0000	1276.8650	
G3	5	1241.7065	1248.0000	1276.8650	

Means and Ranges Chart Summary for weight

sample	-3 Sigma Limits with n=5 for Range- Lower Limit	Subgroup Range	Upper Limit
C9	0	35.000000	64.441879
C4	0	25.000000	64.441879
A7	0	24.000000	64.441879
A1	0	39.000000	64.441879
A5	0	29.000000	64.441879
D3	0	21.000000	64.441879
D6	0	37.000000	64.441879
D1	0	37.000000	64.441879
R4	0	28.000000	64.441879
H7	0	19.000000	64.441879
H2	0	36.000000	64.441879
H6	0	36.000000	64.441879
P4	0	29.000000	64.441879
P9	0	22.000000	64.441879
J7	0	24.000000	64.441879
J2	0	41.000000	64.441879
J3	0	36.000000	64.441879
B2	0	31.000000	64.441879
G4	0	25.000000	64.441879
G6	0	36.000000	64.441879
G3	0	30.000000	64.441879

Output 43.1.1 and Output 43.1.2 indicate that Test 1 is positive at sample D1 and Test 5 is positive at sample P9. Test 1 detects one point beyond Zone A (outside the control limits), and Test 5 detects two out of three points in a row in Zone A or beyond.

Example 43.2. Specifying Standard Values for the Process Mean and Standard Deviation

See SHWXR3 in the SAS/QC Sample Library

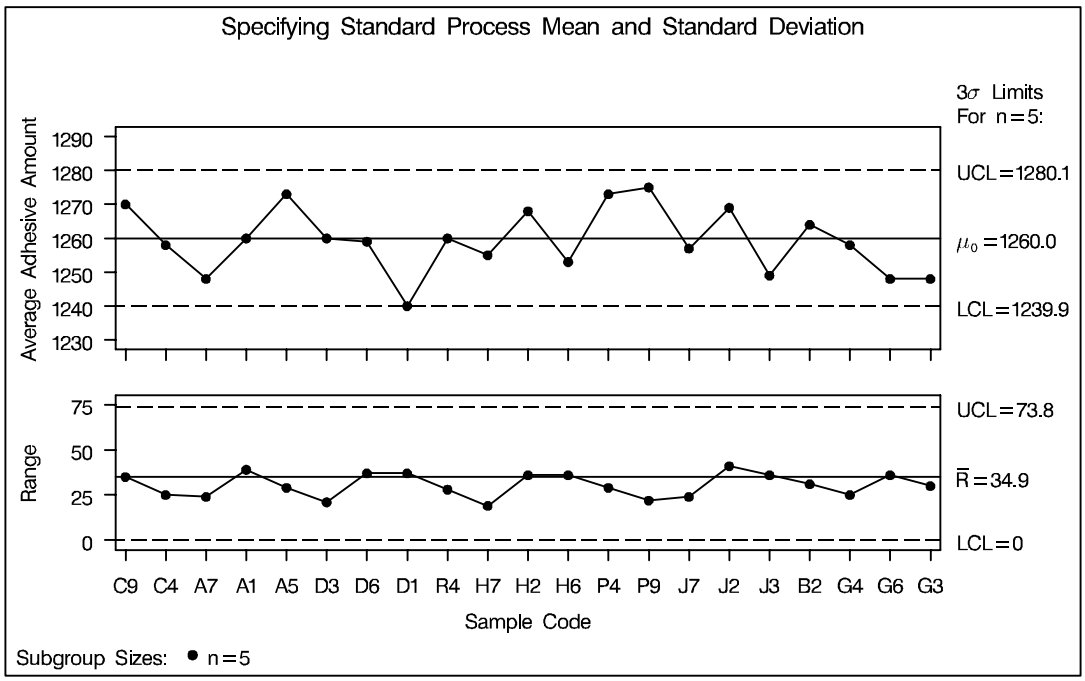
By default, the XRCHART statement estimates the process mean (μ) and standard deviation (σ) from the data, as in the previous example. However, there are applications in which standard values (μ_0 and σ_0) are available based, for instance, on previous experience or extensive sampling. You can specify these values with the MU0= and SIGMA0= options.

For example, suppose it is known that the adhesive coating process introduced in the previous example has a mean of 1260 and standard deviation of 15. The following statements specify these standard values:

```
title 'Specifying Standard Process Mean and Standard Deviation';
symbol v=dot;
proc shewhart history=tape;
   xrchart weight*sample /
       mu0    = 1260
       sigma0 = 15
       xsymbol= mu0;
run;
```

The XSYMBOL= option specifies the label for the central line on the $\bar{X}$ chart. The resulting $\bar{X}$ and R charts are shown in Output 43.2.1.

Output 43.2.1. Specifying Standard Values with MU0= and SIGMA0=



The central lines and control limits for both charts are determined using μ_0 and σ_0 (see the equations in Table 43.22 on page 1526). Output 43.2.1 indicates that the process is in statistical control.

You can also specify μ_0 and σ_0 with the variables _MEAN_ and _STDDEV_ in a LIMITS= data set, as illustrated by the following statements:

```
data tapelim;
   length _var_ _subgrp_ _type_ $8;
   _var_    = 'weight';
   _subgrp_ = 'sample';
   _type_   = 'STANDARD';
   _limitn_ = 5;
   _mean_   = 1260;
   _stddev_ = 15;

proc shewhart history=tape limits=tapelim;
   xrchart weight*sample / xsymbol=mu0;
run;
```

The variables _VAR_ and _SUBGRP_ are required, and their values must match the *process* and *subgroup-variable*, respectively, specified in the XRCHART statement. The bookkeeping variable _TYPE_ is not required, but it is recommended to indicate that the variables _MEAN_ and _STDDEV_ provide standard values rather than estimated values.

The resulting charts (not shown here) are identical to those shown in Output 43.2.1.

Example 43.3. Working with Unequal Subgroup Sample Sizes

See SHWXR4 in the SAS/QC Sample Library

The following data set (WIRE) contains breaking strength measurements recorded in pounds per inch for 25 samples from a metal wire manufacturing process. The subgroup sample sizes vary between 3 and 7.

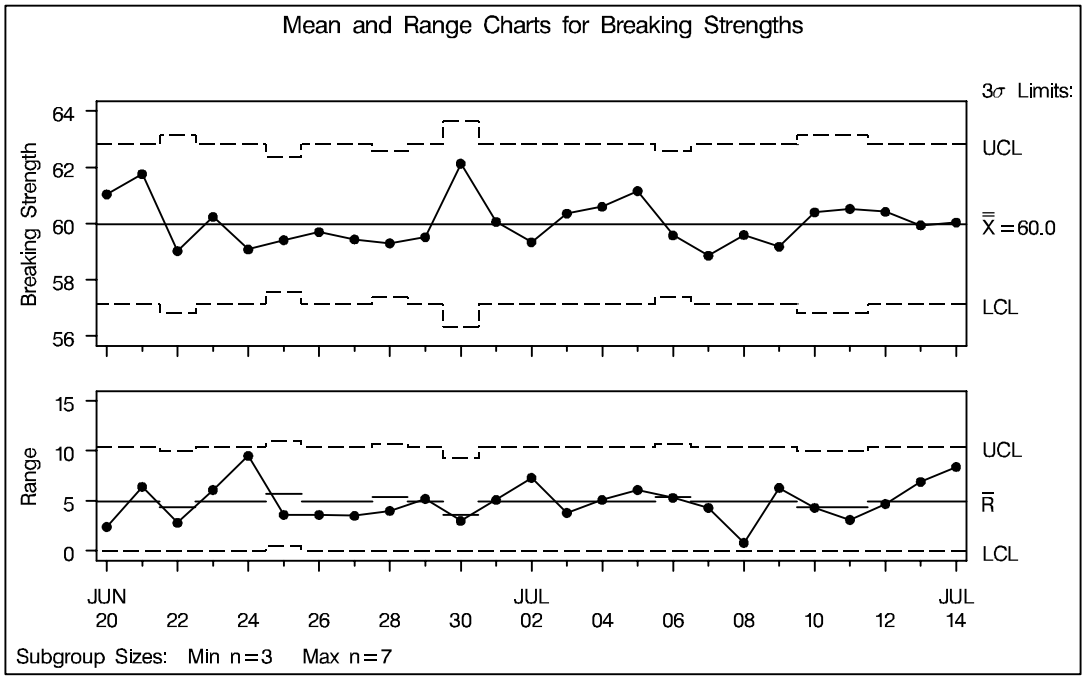
```
data wire;
   input day size @;
   informat day date7.;
   format day date7.;
   do i=1 to size;
      input brstr @@;
      output;
      end;
   drop i size;
   label brstr  ='Breaking Strength (lb/in)';
   datalines;
20JUN94 5 60.6 62.3 62.0 60.4 59.9
21JUN94 5 61.9 62.1 60.6 58.9 65.3
22JUN94 4 57.8 60.5 60.1 57.7
23JUN94 5 56.8 62.5 60.1 62.9 58.9
24JUN94 5 63.0 60.7 57.2 61.0 53.5
25JUN94 7 58.7 60.1 59.7 60.1 59.1 57.3 60.9
26JUN94 5 59.3 61.7 59.1 58.1 60.3
27JUN94 5 61.3 58.5 57.8 61.0 58.6
28JUN94 6 59.5 58.3 57.5 59.4 61.5 59.6
29JUN94 5 61.7 60.7 57.2 56.5 61.5
30JUN94 3 63.9 61.6 60.9
01JUL94 5 58.7 61.4 62.4 57.3 60.5
```

```
02JUL94 5 56.8 58.5 55.7 63.0 62.7
03JUL94 5 62.1 60.6 62.1 58.7 58.3
04JUL94 5 59.1 60.4 60.4 59.0 64.1
05JUL94 5 59.9 58.8 59.2 63.0 64.9
06JUL94 6 58.8 62.4 59.4 57.1 61.2 58.6
07JUL94 5 60.3 58.7 60.5 58.6 56.2
08JUL94 5 59.2 59.8 59.7 59.3 60.0
09JUL94 5 62.3 56.0 57.0 61.8 58.8
10JUL94 4 60.5 62.0 61.4 57.7
11JUL94 4 59.3 62.4 60.4 60.0
12JUL94 5 62.4 61.3 60.5 57.7 60.2
13JUL94 5 61.2 55.5 60.2 60.4 62.4
14JUL94 5 59.0 66.1 57.7 58.5 58.9
;
```

The following statements request $\bar{X}$ and R charts, shown in Output 43.3.1, for the strength measurements:

```
title 'Mean and Range Charts for Breaking Strengths';
symbol v=dot;
proc shewhart data=wire;
   xrchart brstr*day / nohlabel;
run;
```

Output 43.3.1. $\bar{X}$ and R Charts with Varying Subgroup Sample Sizes



Note that the central line on the R chart and the control limits on both charts vary with the subgroup sample size. The sample size legend in the lower left corner displays the minimum and maximum subgroup sample sizes.

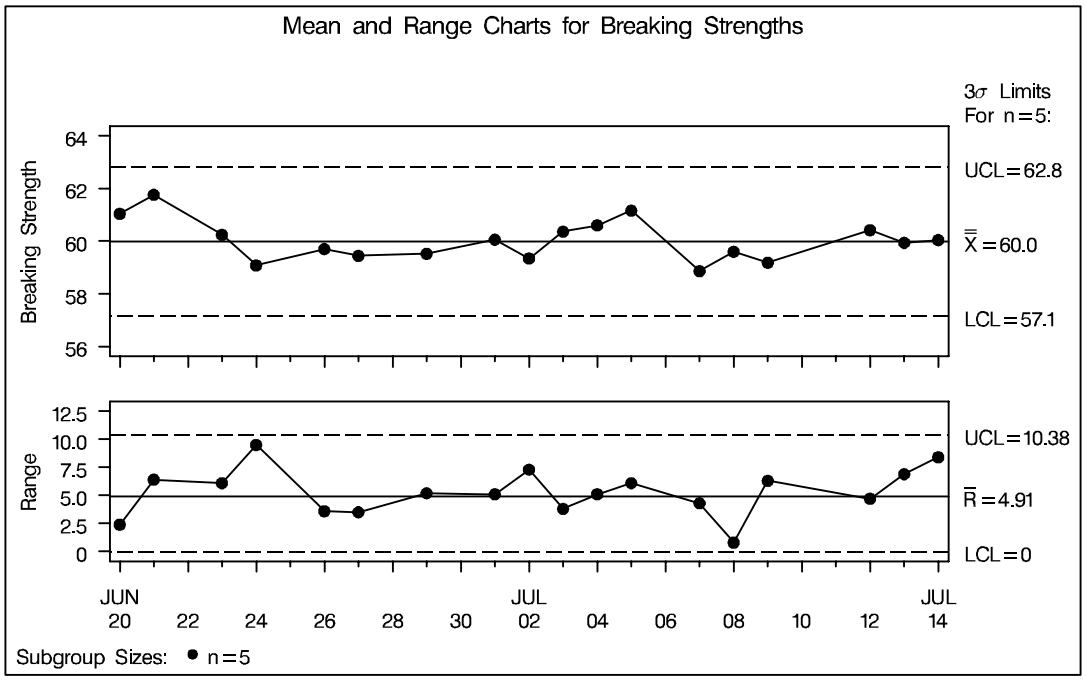
The XRCHART statement provides various options for working with unequal subgroup sample sizes. For example, you can use the LIMITN= option to specify a fixed

(nominal) sample size for computing control limits, as illustrated by the following statements:

```
proc shewhart data=wire;
   xrchart brstr*day / nohlabel
                       limitn=5;
run;
```

The resulting charts are shown in Output 43.3.2.

Output 43.3.2. Control Limits Based on Fixed Sample Size

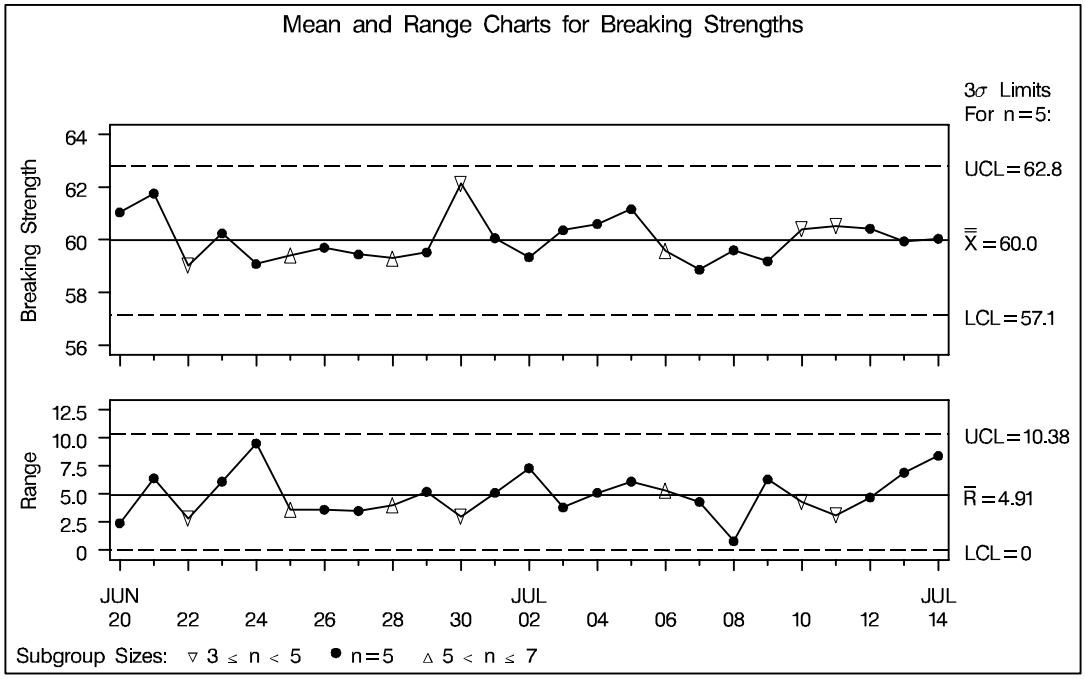


Note that the only points displayed on the chart are those corresponding to subgroups whose sample sizes match the nominal sample size of five. To plot points for all subgroups (regardless of subgroup sample size), you can specify the ALLN option, as follows:

```
proc shewhart data=wire;
   xrchart brstr*day / nohlabel
                       limitn=5
                       alln
                       nmarkers;
run;
```

The charts are shown in Output 43.3.3. The NMARKERS option requests special symbols to identify points for which the subgroup sample size differs from the nominal sample size.

Output 43.3.3. Displaying All Subgroups Regardless of Sample Size



You can use the SMETHOD= option to determine how the process standard deviation σ is to be estimated when the subgroup sample sizes vary. The default method computes $\hat{\sigma}$ as an unweighted average of subgroup estimates of σ. Specifying SMETHOD=MVLUE requests an estimate that assigns greater weight to estimates of σ from subgroups with larger sample sizes. For more information, see "Methods for Estimating the Standard Deviation" on page 1536.

The following statements apply both methods:

```
proc shewhart data=wire;
   xrchart brstr*day / outlimits = wlim1
                       outindex  = 'Default'
                       nochart;
   xrchart brstr*day / smethod   = mvlue
                       outlimits = wlim2
                       outindex  = 'MVLUE'
                       nochart;
run;


data wlimits;
   set wlim1 wlim2;
run;
```

The data set WLIMITS is listed in Output 43.3.4.

Output 43.3.4. Listing of the Data Set WLIMITS

```
                        The WLIMITS Data Set

      _                 _           _                         _
      S     _           L          _ S                        S
      U     I           I          A I _     _    _ _   _     T
  _   B     N     T     M          L G L     M    U L   U     D
  V   G     D     Y     I          P M C     E    C C   C     D
  A   R     E     P     T          H A L     A    L L _ L     E
  R   P     X     E     N          A S X     N    X R R R     V
  _   _     _     _     _          _ _ _     _    _ _ _ _     _

brstr day Default ESTIMATE V .002699796 3 V 59.9766 V V V V 2.11146
brstr day MVLUE   ESTIMATE V .002699796 3 V 59.9766 V V V V 2.11240
```

The variables in an OUTLIMITS= data set whose values vary with subgroup sample size are assigned the special missing value *V*. Consequently, the control limit variables (_LCLX_, _UCLX_, _LCLR_, and _UCLR_), as well as the variables _R_ and _LIMITN_, have this value.

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

ISBN 1–58025–493–4

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

1st printing, October 1999