

Chapter 44

XSCART Statement

Chapter Table of Contents

OVERVIEW	1551
GETTING STARTED	1552
Creating Charts for Means and Standard Deviations from Raw Data	1552
Creating Charts for Means and Standard Deviations from Summary Data	1555
Saving Summary Statistics	1557
Saving Control Limits	1558
Reading Preestablished Control Limits	1561
SYNTAX	1562
Summary of Options	1563
DETAILS	1573
Constructing Charts for Means and Standard Deviations	1573
Output Data Sets	1575
ODS Tables	1578
Input Data Sets	1578
Methods for Estimating the Standard Deviation	1582
Axis Labels	1583
Missing Values	1584
EXAMPLES	1585
Example 44.1 Specifying Probability Limits	1585
Example 44.2 Computing Subgroup Summary Statistics	1586
Example 44.3 Analyzing Nonnormal Process Data	1588

Chapter 44

XSCART Statement

Overview

The XSCART statement creates \bar{X} and s charts for subgroup means and standard deviations, which are used to analyze the central tendency and variability of a process.

You can use options in the XSCART statement to

- compute control limits from the data based on a multiple of the standard error of the plotted means and standard deviations or as probability limits
- tabulate subgroup sample sizes, subgroup means, subgroup standard deviations, control limits, and other information
- save control limits in an output data set
- save subgroup sample sizes, subgroup means, and subgroup standard deviations 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 XSCHEM statement with simple examples that illustrate commonly used options. Complete syntax for the XSCHEM statement is presented in the “Syntax” section on page 1562, and advanced examples are given in the “Examples” section on page 1585.

Creating Charts for Means and Standard Deviations from Raw Data

See SHWXS1
in the SAS/QC
Sample Library

A petroleum company uses a turbine to heat water into steam, which is then pumped into the ground to make oil more viscous and easier to extract. This process occurs 20 times daily, and the amount of power (in kilowatts) used to heat the water to the desired temperature is recorded. The following statements create a SAS data set named TURBINE, which contains the power output measurements for 20 days:

```

data turbine;
  informat day date7.;
  format day date5.;
  input day @;
  do i=1 to 10;
    input kwatts @;
    output;
  end;
  drop i;
  datalines;
04JUL94 3196 3507 4050 3215 3583 3617 3789 3180 3505 3454
04JUL94 3417 3199 3613 3384 3475 3316 3556 3607 3364 3721
05JUL94 3390 3562 3413 3193 3635 3179 3348 3199 3413 3562
05JUL94 3428 3320 3745 3426 3849 3256 3841 3575 3752 3347
06JUL94 3478 3465 3445 3383 3684 3304 3398 3578 3348 3369
06JUL94 3670 3614 3307 3595 3448 3304 3385 3499 3781 3711
07JUL94 3448 3045 3446 3620 3466 3533 3590 3070 3499 3457
07JUL94 3411 3350 3417 3629 3400 3381 3309 3608 3438 3567
08JUL94 3568 2968 3514 3465 3175 3358 3460 3851 3845 2983
08JUL94 3410 3274 3590 3527 3509 3284 3457 3729 3916 3633
09JUL94 3153 3408 3741 3203 3047 3580 3571 3579 3602 3335
09JUL94 3494 3662 3586 3628 3881 3443 3456 3593 3827 3573
10JUL94 3594 3711 3369 3341 3611 3496 3554 3400 3295 3002
10JUL94 3495 3368 3726 3738 3250 3632 3415 3591 3787 3478
11JUL94 3482 3546 3196 3379 3559 3235 3549 3445 3413 3859
11JUL94 3330 3465 3994 3362 3309 3781 3211 3550 3637 3626
12JUL94 3152 3269 3431 3438 3575 3476 3115 3146 3731 3171
12JUL94 3206 3140 3562 3592 3722 3421 3471 3621 3361 3370
13JUL94 3421 3381 4040 3467 3475 3285 3619 3325 3317 3472
13JUL94 3296 3501 3366 3492 3367 3619 3550 3263 3355 3510
14JUL94 3795 3872 3559 3432 3322 3587 3336 3732 3451 3215
14JUL94 3594 3410 3335 3216 3336 3638 3419 3515 3399 3709
15JUL94 3850 3431 3460 3623 3516 3810 3671 3602 3480 3388
15JUL94 3365 3845 3520 3708 3202 3365 3731 3840 3182 3677
16JUL94 3711 3648 3212 3664 3281 3371 3416 3636 3701 3385
16JUL94 3769 3586 3540 3703 3320 3323 3480 3750 3490 3395

```

```

17JUL94 3596 3436 3757 3288 3417 3331 3475 3600 3690 3534
17JUL94 3306 3077 3357 3528 3530 3327 3113 3812 3711 3599
18JUL94 3428 3760 3641 3393 3182 3381 3425 3467 3451 3189
18JUL94 3588 3484 3759 3292 3063 3442 3712 3061 3815 3339
19JUL94 3746 3426 3320 3819 3584 3877 3779 3506 3787 3676
19JUL94 3727 3366 3288 3684 3500 3501 3427 3508 3392 3814
20JUL94 3676 3475 3595 3122 3429 3474 3125 3307 3467 3832
20JUL94 3383 3114 3431 3693 3363 3486 3928 3753 3552 3524
21JUL94 3349 3422 3674 3501 3639 3682 3354 3595 3407 3400
21JUL94 3401 3359 3167 3524 3561 3801 3496 3476 3480 3570
22JUL94 3618 3324 3475 3621 3376 3540 3585 3320 3256 3443
22JUL94 3415 3445 3561 3494 3140 3090 3561 3800 3056 3536
23JUL94 3421 3787 3454 3699 3307 3917 3292 3310 3283 3536
23JUL94 3756 3145 3571 3331 3725 3605 3547 3421 3257 3574
;

```

A partial listing of TURBINE is shown in Figure 44.1.

Kilowatt Power Output Data		
Obs	day	kwatts
1	04JUL	3196
2	04JUL	3507
3	04JUL	4050
.	.	.
.	.	.
.	.	.
21	05JUL	3390
22	05JUL	3562
23	05JUL	3413
.	.	.
.	.	.
.	.	.
398	23JUL	3421
399	23JUL	3257
400	23JUL	3574

Figure 44.1. Partial Listing of the Data Set TURBINE

The data set is said to be in “strung-out” form since each observation contains the day and power output for a single heating. The first 20 observations contain the power outputs for the first day, the second 20 observations contain the power outputs for the second day, and so on. Because the variable DAY classifies the observations into rational subgroups, it is referred to as the *subgroup-variable*. The variable KWATTS contains the power output measurements and is referred to as the *process variable* (or *process* for short).

You can use \bar{X} and s charts to determine whether the heating process is in control. The following statements create the \bar{X} and s charts shown in Figure 44.2:

```

title 'Mean and Standard Deviation Charts for Power Output';
symbol v=dot;
proc shewhart data=turbine;
    xschart kwatts*day;
run;

```

This example illustrates the basic form of the XSCHEM statement. After the keyword XSCHEM, you specify the *process* to analyze (in this case KWATTS), followed by an asterisk and the *subgroup-variable* (DAY).

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

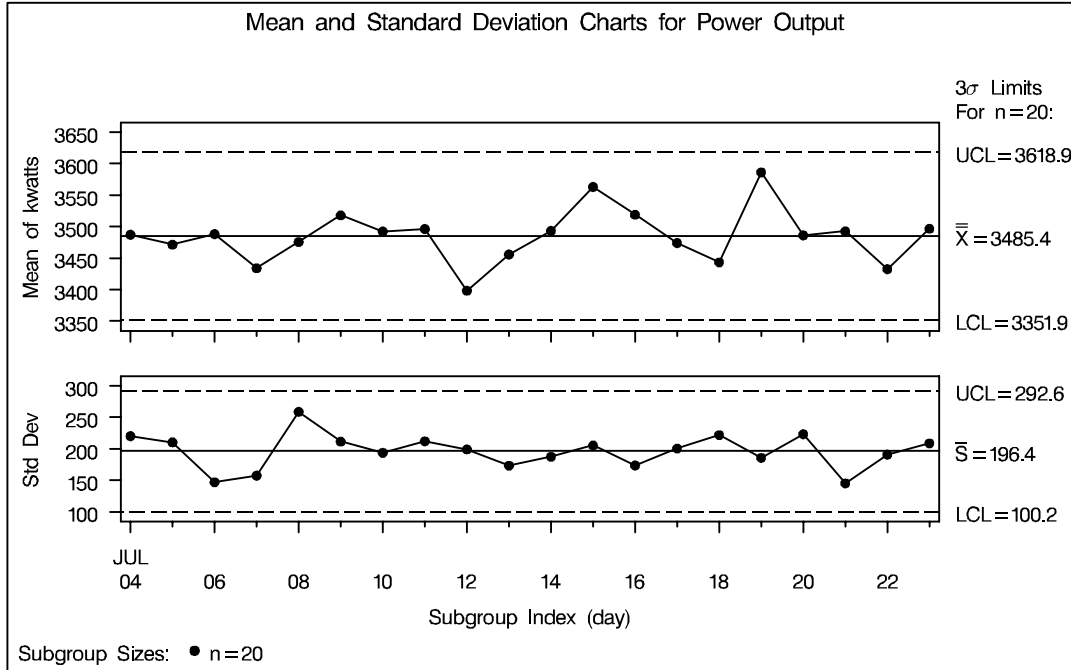


Figure 44.2. \bar{X} and s Charts for Power Output Data

Each point on the \bar{X} chart represents the mean of the measurements for a particular day. For instance, the mean plotted for the first day is $(3196 + 3507 + \dots + 3721) / 20 = 3487.4$.

Each point on the s chart represents the standard deviation of the measurements for a particular day. For instance, the standard deviation plotted for the first day is

$$\sqrt{\frac{(3196 - 3487.4)^2 + (3507 - 3487.4)^2 + \dots + (3721 - 3487.4)^2}{19}} = 220.26$$

Since all the points lie within the control limits, it can be concluded that the process is in statistical control.

By default, the control limits shown are 3σ limits estimated from the data; the formulas for the limits are given in Table 44.22 on page 1574. You can also read control limits from an input data set; see “Reading Prestablished Control Limits” on page 1561.

For computational details, see “Constructing Charts for Means and Standard Deviations” on page 1573. For more details on reading raw data, see “DATA= Data Set” on page 1578.

Creating Charts for Means and Standard Deviations from Summary Data

The previous example illustrates how you can create \bar{X} and s charts using raw data (process measurements). However, in many applications the data are provided as subgroup summary statistics. This example illustrates how you can use the XSCART statement with data of this type.

See SHWXS1
in the SAS/QC
Sample Library

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

```
data oilsum;
  input day kwattsx kwattss kwattsn;
  informat day date7. ;
  format day date5. ;
  label day='Date of Measurement';
  datalines;
04JUL94 3487.40 220.260 20
05JUL94 3471.65 210.427 20
06JUL94 3488.30 147.025 20
07JUL94 3434.20 157.637 20
08JUL94 3475.80 258.949 20
09JUL94 3518.10 211.566 20
10JUL94 3492.65 193.779 20
11JUL94 3496.40 212.024 20
12JUL94 3398.50 199.201 20
13JUL94 3456.05 173.455 20
14JUL94 3493.60 187.465 20
15JUL94 3563.30 205.472 20
16JUL94 3519.05 173.676 20
17JUL94 3474.20 200.576 20
18JUL94 3443.60 222.084 20
19JUL94 3586.35 185.724 20
20JUL94 3486.45 223.474 20
21JUL94 3492.90 145.267 20
22JUL94 3432.80 190.994 20
23JUL94 3496.90 208.858 20
;
```

A partial listing of OILSUM is shown in Figure 44.3.

Summary Data Set for Power Output				
day	kwattsx	kwattss	kwattsn	
04JUL	3487.40	220.260	20	
05JUL	3471.65	210.427	20	
06JUL	3488.30	147.025	20	
07JUL	3434.20	157.637	20	
08JUL	3475.80	258.949	20	
.	.	.	.	
.	.	.	.	
.	.	.	.	

Figure 44.3. The Summary Data Set OILSUM

Part 9. The CAPABILITY Procedure

There is exactly one observation for each subgroup (note that the subgroups are still indexed by DAY). The variable KWATTSSX contains the subgroup means, the variable KWATTSS contains the subgroup standard deviations, and the variable KWATTSN contains the subgroup sample sizes (which are all 20). You can read this data set by specifying it as a HISTORY= data set in the PROC SHEWHART statement, as follows:

```

title 'Mean and Standard Deviation Charts for Power Output';
proc shewhart history=oilsum lineprinter;
  xschart kwatts*day='*';
run;

```

The resulting \bar{X} and s charts are shown in Figure 44.4. 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 KWATTS is *not* the name of a SAS variable in the data set OILSUM but is, instead, the common prefix for the names of the three SAS variables KWATTSSX, KWATTSS, and KWATTSN. The suffix characters X, S, and N indicate *mean*, *standard deviation*, and *sample size*, respectively. Thus, you can specify three subgroup summary variables in the HISTORY= data set with a single name (KWATTS), which is referred to as the *process*. The name DAY specified after the asterisk is the name of the *subgroup-variable*.

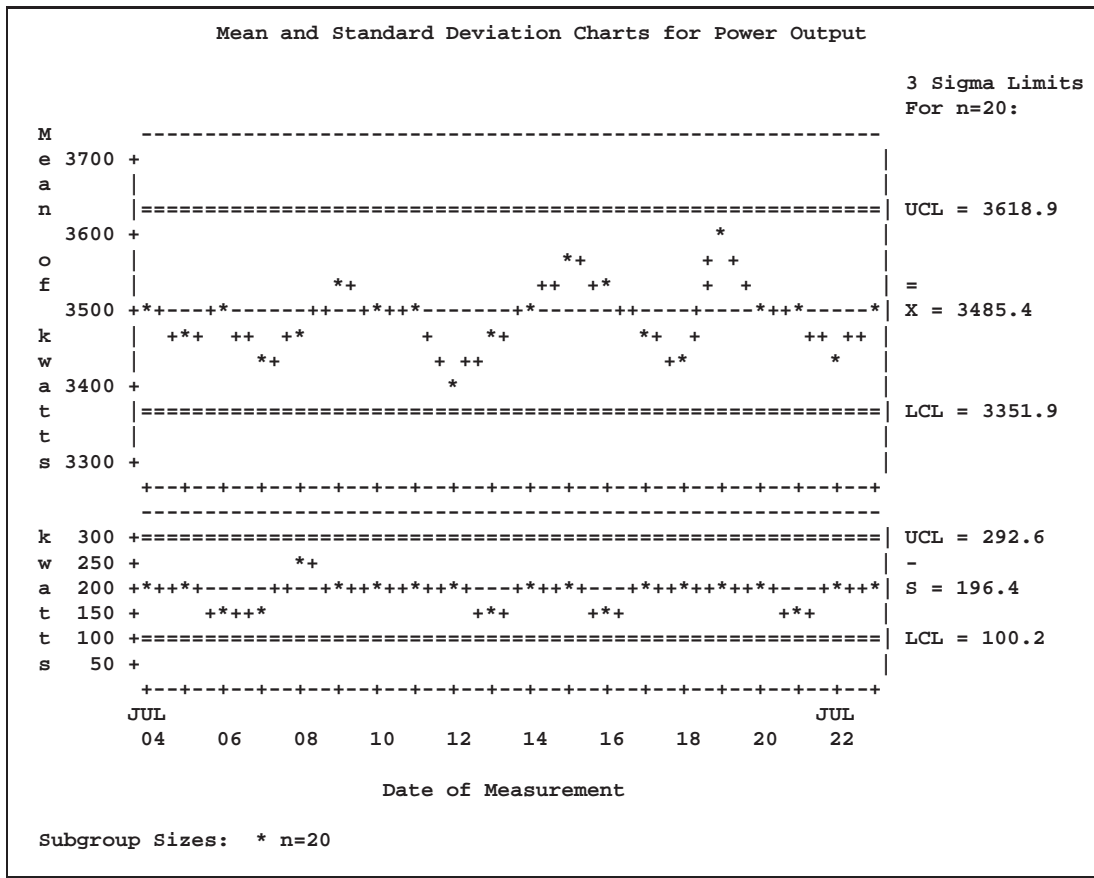


Figure 44.4. \bar{X} and s Charts for Power Output Data

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

- subgroup variable
- subgroup mean variable
- subgroup standard deviation variable
- subgroup sample size variable

Furthermore, the names of the subgroup mean, standard deviation, and sample size variables must begin with the *process* name specified in the XSCHART statement and end with the special suffix characters *X*, *S*, 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. For an illustration, see Example 44.2 on page 1586.

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 1579.

Saving Summary Statistics

In this example, the XSCHART 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 TURBINE (see page 1552) and create a summary data set named TURBHIST:

See SHWXS1 in the SAS/QC Sample Library

```

title 'Summary Data Set for Power Output';
proc shewhart data=turbine;
    xschart kwatts*day / outhistory = turbhist
                        nochart;
run;

```

The OUTHISTORY= option names the output data set, and the NOCHART option suppresses the display of the charts, which would be identical to those in Figure 44.2. Options such as OUTHISTORY= and NOCHART are specified after the slash (/) in the XSCHART statement. A complete list of options is presented in the “Syntax” section on page 1562.

Figure 44.5 contains a partial listing of TURBHIST.

Summary Data Set for Power Output			
day	kwattsX	kwattsS	kwattsN
04JUL	3487.40	220.260	20
05JUL	3471.65	210.427	20
06JUL	3488.30	147.025	20
07JUL	3434.20	157.637	20
08JUL	3475.80	258.949	20
.	.	.	.
.	.	.	.
.	.	.	.

Figure 44.5. The Summary Data Set TURBHIST

There are four variables in the data set TURBHIST.

- DAY contains the subgroup index.
- KWATTSSX contains the subgroup means.
- KWATTSSS contains the subgroup standard deviations.
- KWATTSSN contains the subgroup sample sizes.

Note that the summary statistic variables are named by adding the suffix characters *X*, *S*, and *N* to the *process* KWATTS specified in the XSCHART 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 1576.

Saving Control Limits

See SHWXS1
in the SAS/QC
Sample Library

You can save the control limits for \bar{X} and *s* charts in a SAS data set; this enables you to apply the control limits to future data (see “Reading Preestablished Control Limits” on page 1561) or modify the limits with a DATA step program.

The following statements read measurements from the data set TURBINE (see page 1552) and save the control limits displayed in Figure 44.2 in a data set named TURBLIM:

```

title1 'Control Limits for Power Output Data';
proc shewhart data=turbine;
  xschart kwatts*day / outlimits=turblim
  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 TURBLIM is listed in Figure 44.6.

Control Limits for Power Output Data						
<u>_VAR_</u>	<u>_SUBGRP_</u>	<u>_TYPE_</u>	<u>_LIMITN_</u>	<u>_ALPHA_</u>	<u>_SIGMAS_</u>	<u>_LCLX_</u>
kwatts	day	ESTIMATE	20	.002699796	3	3351.92
<u>_MEAN_</u>	<u>_UCLX_</u>	<u>_LCLS_</u>	<u>_S_</u>	<u>_UCLS_</u>	<u>_STDDEV_</u>	
3485.41	3618.90	100.207	196.396	292.584	198.996	

Figure 44.6. The Data Set TURBLIM Containing Control Limit Information

The data set TURBLIM contains one observation with the limits for *process* KWATTS. The variables _LCLX_ and _UCLX_ contain the lower and upper control limits for the \bar{X} chart, and the variables _LCLS_ and _UCLS_ contain the lower and upper control limits for the *s* chart. The variable _MEAN_ contains the central line for the \bar{X} chart, and the variable _S_ contains the central line for the *s* 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. For more information, see “OUTLIMITS= Data Set” on page 1575.

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=turbine;
  xschart kwatts*day / outtable=turbtabs
                    nochart;
run;

```

The data set TURBTABS contains one observation for each subgroup sample. The variables _SUBX_, _SUBS_, and _SUBN_ contain the subgroup means, subgroup standard deviations, and subgroup sample sizes. The variables _LCLX_ and _UCLX_ contain the lower and upper control limits for the \bar{X} chart. The variables _LCLS_ and _UCLS_ contain the lower and upper control limits for the *s* chart. The variable _MEAN_ contains the central line for the \bar{X} chart. The variable _S_ contains the central line for the *s* 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 1577.

The data set TURBTABS is listed in Figure 44.7.

Summary Statistics and Control Limit Information							
VAR	day	_SIGMAS_	_LIMITN_	_SUBN_	_LCLX_	_SUBX_	_MEAN_
kwatts	04JUL	3	20	20	3351.92	3487.40	3485.41
kwatts	05JUL	3	20	20	3351.92	3471.65	3485.41
kwatts	06JUL	3	20	20	3351.92	3488.30	3485.41
kwatts	07JUL	3	20	20	3351.92	3434.20	3485.41
kwatts	08JUL	3	20	20	3351.92	3475.80	3485.41
kwatts	09JUL	3	20	20	3351.92	3518.10	3485.41
kwatts	10JUL	3	20	20	3351.92	3492.65	3485.41
kwatts	11JUL	3	20	20	3351.92	3496.40	3485.41
kwatts	12JUL	3	20	20	3351.92	3398.50	3485.41
kwatts	13JUL	3	20	20	3351.92	3456.05	3485.41
kwatts	14JUL	3	20	20	3351.92	3493.60	3485.41
kwatts	15JUL	3	20	20	3351.92	3563.30	3485.41
kwatts	16JUL	3	20	20	3351.92	3519.05	3485.41
kwatts	17JUL	3	20	20	3351.92	3474.20	3485.41
kwatts	18JUL	3	20	20	3351.92	3443.60	3485.41
kwatts	19JUL	3	20	20	3351.92	3586.35	3485.41
kwatts	20JUL	3	20	20	3351.92	3486.45	3485.41
kwatts	21JUL	3	20	20	3351.92	3492.90	3485.41
kwatts	22JUL	3	20	20	3351.92	3432.80	3485.41
kwatts	23JUL	3	20	20	3351.92	3496.90	3485.41
UCLX	_EXLIM_	_LCLS_	_SUBS_	_S_	_UCLS_	_EXLIMS_	
3618.90		100.207	220.260	196.396	292.584		
3618.90		100.207	210.427	196.396	292.584		
3618.90		100.207	147.025	196.396	292.584		
3618.90		100.207	157.637	196.396	292.584		
3618.90		100.207	258.949	196.396	292.584		
3618.90		100.207	211.566	196.396	292.584		
3618.90		100.207	193.779	196.396	292.584		
3618.90		100.207	212.024	196.396	292.584		
3618.90		100.207	199.201	196.396	292.584		
3618.90		100.207	173.455	196.396	292.584		
3618.90		100.207	187.465	196.396	292.584		
3618.90		100.207	205.472	196.396	292.584		
3618.90		100.207	173.676	196.396	292.584		
3618.90		100.207	200.576	196.396	292.584		
3618.90		100.207	222.084	196.396	292.584		
3618.90		100.207	185.724	196.396	292.584		
3618.90		100.207	223.474	196.396	292.584		
3618.90		100.207	145.267	196.396	292.584		
3618.90		100.207	190.994	196.396	292.584		
3618.90		100.207	208.858	196.396	292.584		

Figure 44.7. The OUTTABLE= Data Set TURBTAB

A data set created with the OUTTABLE= option can be read later as a TABLE= data set. For example, the following statements read TURBTAB and display charts (not shown here) identical to those in Figure 44.2:

```

title 'Mean and Standard Deviation Charts for Power Output';
proc shewhart table=turbtab;
    xschart kwatts*day;
run;

```

Because the SHEWHART procedure simply displays the information in 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 1580.

Reading Preestablished Control Limits

In the previous example, the OUTLIMITS= data set TURBLIM saved control limits computed from the measurements in TURBINE. This example shows how these limits can be applied to new data. The following statements create \bar{X} and s charts for new measurements in a data set named TURBINE2 (not listed here) using the control limits in TURBLIM:

See SHWXS1
in the SAS/QC
Sample Library

```
title 'Mean and Standard Deviation Charts for Power Output';
proc shewhart data=turbine2 limits=turblim;
  xschart kwatts*day;
run;
```

The charts are shown in Figure 44.8. The LIMITS= option in the PROC SHEWHART statement specifies the data set containing preestablished control limit information. 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 KWATTS
- the value of `_SUBGRP_` matches the *subgroup-variable* name DAY

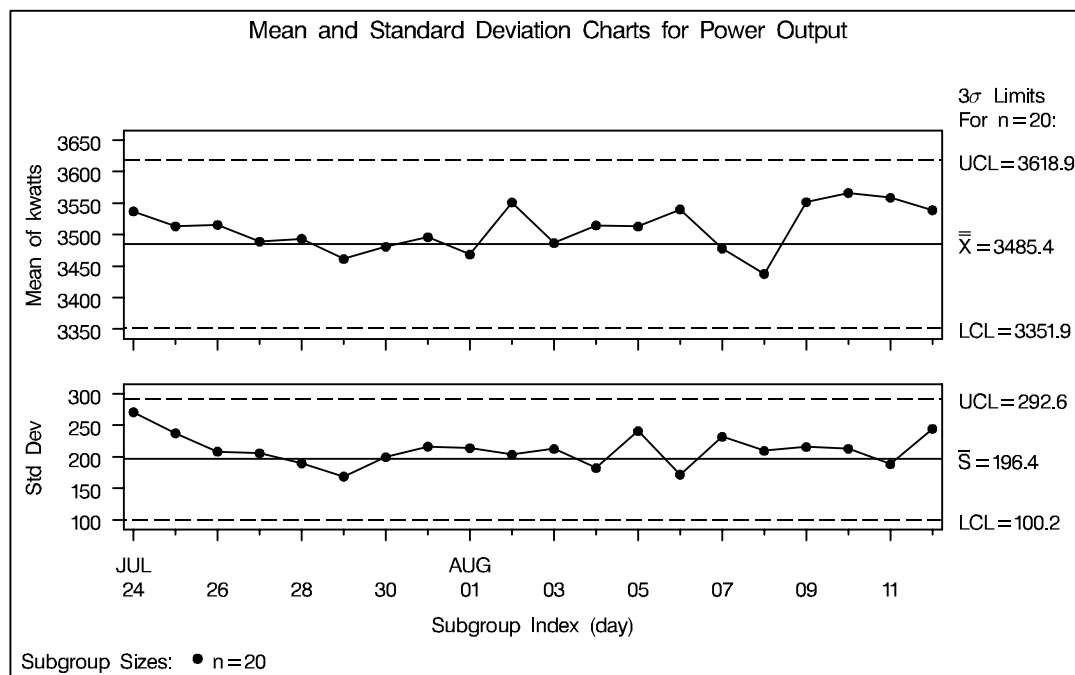


Figure 44.8. \bar{X} and s Charts for Second Set of Power Outputs

The means and standard deviations lie within the control limits, indicating that the heating process is still in statistical control.

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 1579 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 XSCHART statement is as follows:

```
XSCHART process*subgroup-variable ;
```

The general form of this syntax is as follows:

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

You can use any number of XSCHART statements in the SHEWHART procedure. The components of the XSCHART 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 the raw data are read using 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 Standard Deviations from Raw Data” on page 1552.
- 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 Standard Deviations from Summary Data” on page 1555.
- 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 1558.

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

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

subgroup-variable

is the variable that identifies subgroups in the data. The *subgroup-variable* is required. In the preceding XSCHART 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 standard deviations.

- 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 s charts using an asterisk (*) to plot the points:

```
proc shewhart data=values;
  xschart 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 XSCHART statement options by function. For complete descriptions, see Chapter 46, “Dictionary of Options.”

Table 44.1. Tabulation Options

TABLE	creates a basic table of subgroup means, subgroup sample sizes, subgroup standard deviations, 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 44.2. Options for Specifying Tests for Special Causes

NO3SIGMACHECK	allows tests to be applied with control limits other than 3σ limits
TESTS= <i>value-list</i> <i>customized-pattern-list</i>	specifies tests for special causes for the \bar{X} chart
TESTS2= <i>value-list</i> <i>customized-pattern-list</i>	specifies tests for special causes for the s chart
TEST2RUN= <i>n</i>	specifies length of pattern for Test 2
TEST3RUN= <i>n</i>	specifies length of pattern for Test 3
TESTACROSS	applies tests across <i>phase</i> boundaries
TESTLABEL= <i>'label'</i> <i>(variable)</i> <i>keyword</i>	provides labels for points where test is positive
TESTLABEL <i>n</i> = <i>'label'</i>	specifies label for n^{th} test for special causes
TESTMETHOD= 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 s chart
ZONES	adds lines to \bar{X} chart delineating zones A, B, and C
ZONES2	adds lines to s chart delineating zones A, B, and C
ZONEVALPOS= <i>n</i>	specifies position of ZONEVALUES and ZONE2VALUES labels
ZONEVALUES	labels \bar{X} chart zone lines with their values
ZONE2VALUES	labels s chart zone lines with their values

Table 44.3. Graphical Options for Displaying Tests for Special Causes

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

Table 44.4. Line Printer Options for Displaying Tests for Special Causes

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

Table 44.5. Reference Line Options

CHREF= <i>color</i>	specifies color for lines requested by the HREF= and HREF2= options
CVREF= <i>color</i>	specifies color for lines requested by the VREF= and VREF2= options
HREF= <i>values</i> <i>SAS-data-set</i>	specifies position of reference lines perpendicular to horizontal axis on \bar{X} chart
HREF2= <i>values</i> <i>SAS-data-set</i>	specifies position of reference lines perpendicular to horizontal axis on <i>s</i> 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 \bar{X} chart
HREF2DATA= <i>SAS-data-set</i>	specifies position of reference lines perpendicular to horizontal axis on <i>s</i> 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 applies uniformly to charts created for all BY groups
VREF= <i>values</i> <i>SAS-data-set</i>	specifies position of reference lines perpendicular to vertical axis on \bar{X} chart
VREF2= <i>values</i> <i>SAS-data-set</i>	specifies position of reference lines perpendicular to vertical axis on <i>s</i> 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

Table 44.6. 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 colors 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 major 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
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 <i>s</i> chart
NOVANGLE	requests vertical axis labels that are strung out vertically
SKIPHLABELS= <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>values</i> AXIS <i>n</i>	specifies major tick mark values for vertical axis of \bar{X} chart
VAXIS2= <i>values</i> AXIS <i>n</i>	specifies major tick mark values for vertical axis of <i>s</i> 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
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= <i>n</i>	specifies width of axis lines

Table 44.7. Specification Limit Options

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

Table 44.8. Clipping Options

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

Table 44.9. Block Variable Legend Options

BLOCKLABELPOS= <i>keyword</i>	specifies position of label for <i>block-variable</i> legend
BLOCKLABTYPE= <i>value 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

Table 44.10. Options for Specifying Control Limits

ALPHA= <i>value</i>	requests probability limits for control charts
LIMITN= <i>n </i> VARYING	specifies either nominal sample size for fixed control limits or varying limits
NOREADLIMITS	computes control limits for each <i>process</i> 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 <i>process</i> from a LIMITS= data set
READLIMITS	reads single set of control limits for each <i>process</i> from a LIMITS= data set (Release 6.09 and earlier releases)
SIGMAS= <i>k</i>	specifies width of control limits in terms of multiple <i>k</i> of standard error of plotted means and standard deviations

Table 44.11. Options for Displaying Control Limits

CINFILL= <i>color</i>	specifies color for area inside control limits
CLIMITS= <i>color</i>	specifies color of control limits, central line, and related labels
LCLLABEL= <i>'label'</i>	specifies label for lower control limit on \bar{X} chart
LCLLABEL2= <i>'label'</i>	specifies label for lower control limit on <i>s</i> chart
LIMLABSUBCHAR= <i>'character'</i>	specifies a substitution character for labels provided as quoted strings; the character is replaced with the value of the control limit
LLIMITS= <i>linetype</i>	specifies line type for control limits
NDECIMAL= <i>n</i>	specifies number of digits to right of decimal place in default labels for control limits and central line on \bar{X} chart
NDECIMAL2= <i>n</i>	specifies number of digits to right of decimal place in default labels for control limits and central line on <i>s</i> chart
NOCTL	suppresses display of central line on \bar{X} chart
NOCTL2	suppresses display of central line on <i>s</i> chart
NOLCL	suppresses display of lower control limit on \bar{X} chart
NOLCL2	suppresses display of lower control limit on <i>s</i> 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 <i>s</i> chart
NOUCL	suppresses display of upper control limit on \bar{X} chart
NOUCL2	suppresses display of upper control limit on <i>s</i> chart
SSYMBOL= <i>'string'</i> <i>keyword</i>	specifies label for central line on <i>s</i> chart
UCLLABEL= <i>'string'</i>	specifies label for upper control limit on \bar{X} chart
UCLLABEL2= <i>'string'</i>	specifies label for upper control limit on <i>s</i> chart
WLIMITS= <i>n</i>	specifies width for control limits and central line
XSYMBOL= <i>'string'</i> <i>keyword</i>	specifies label for central line on \bar{X} chart

Table 44.12. Grid Options

ENDGRID	adds grid after last plotted point
GRID	adds grid to control 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 44.13. Options for Plotting and Labeling Points

ALLLABEL=VALUE (variable)	labels every point on \bar{X} chart
ALLLABEL2=VALUE (variable)	labels every point on s 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 s chart
SYMBOLCHARS= 'characters'	specifies characters indicating <i>symbol-variable</i>
SYMBOLLEGEND= NONE name	specifies LEGEND statement for levels of <i>symbol-variable</i>
SYMBOLORDER= keyword	specifies order in which symbols are assigned for levels of <i>symbol-variable</i>
TURNALL TURNOUT	turns point labels so that they are strung out vertically

Table 44.14. Input Data Set Options

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

Table 44.15. Output Data Set Options

OUTHISTORY= SAS-data-set	creates output data set containing subgroup summary statistics
OUTINDEX='string'	specifies value of <code>_INDEX_</code> 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 44.16. Phase Options

CPHASEBOX= <i>color</i>	specifies color for box enclosing all plotted points for a phase
CPHASEBOX- CONNECT= <i>color</i>	specifies color for line segments connecting adjacent enclosing boxes
CPHASEBOXFILL= <i>color</i>	specifies fill color for box enclosing all plotted points for a phase
CPHASELEG= <i>color</i>	specifies text color for <i>phase</i> legend
CPHASEMEAN- CONNECT= <i>color</i>	specifies color for line segments connecting average value points within a phase
NOPHASEFRAME	suppresses default frame for <i>phase</i> legend
OUTPHASE= <i>'string'</i>	specifies value of <code>_PHASE_</code> in the 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 charts
PHASELIMITS	labels control limits for each phase, provided they are constant within that phase
PHASEMEANSYMBOL= <i>symbol</i>	specifies symbol marker for average of values within a phase
PHASEREF	delineates <i>phases</i> with vertical reference lines
READPHASES=ALL <i>'label1' ...'labeln'</i>	specifies <i>phases</i> to be read from an input data set

Table 44.17. Process Mean and Standard Deviation Options

MU0= <i>value</i>	specifies known (standard) value of μ_0 for process mean μ
SIGMA0= <i>value</i>	specifies known (standard) value σ_0 for process standard deviation σ
SMETHOD= <i>keyword</i>	specifies method for estimating the process standard deviation σ
TYPE= <i>keyword</i>	identifies whether parameters are estimates or standard values and specifies value of <code>_TYPE_</code> in the OUTLIMITS= data set

Table 44.18. 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
TESTURLS= <i>SAS-data-set</i>	associates URLs with tests for special causes
WEBOUT= <i>SAS-data-set</i>	creates an OUTTABLE= data set with additional graphics coordinate data

Table 44.19. Plot Layout Options

ALLN	plots means and standard deviations for all subgroups
BILEVEL	creates control charts using half-screens and half-pages
EXCHART	creates control charts for a process variable 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 charts
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 <i>s</i> chart
NOFRAME	suppresses frame for plot area
NOLEGEND	suppresses legend for subgroup sample sizes
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
SEPARATE	displays \bar{X} and <i>s</i> charts on separate screens or pages
TOTPANELS= <i>n</i>	specifies number of pages or screens to be used to display chart
YPCT1= <i>value</i>	specifies length of vertical axis on \bar{X} chart as a percentage of the sum of lengths of vertical axes for \bar{X} and <i>s</i> charts
ZEROSTD	displays \bar{X} and <i>s</i> charts regardless of whether $\hat{\sigma} = 0$

Table 44.20. Graphical Enhancement Options

ANNOTATE= <i>SAS-data-set</i>	specifies annotate data set that adds features to \bar{X} chart
ANNOTATE2= <i>SAS-data-set</i>	specifies annotate data set that adds features to <i>s</i> chart
DESCRIPTION= <i>'string'</i>	specifies string that appears in the description field of the PROC GREPLAY master menu for \bar{X} chart
DESCRIPTION2= <i>'string'</i>	specifies string that appears in the description field of the PROC GREPLAY master menu for <i>s</i> chart
FONT= <i>font</i>	specifies software font for labels and legends on charts
NAME= <i>'string'</i>	specifies name that appears in the name field of the PROC GREPLAY master menu for \bar{X} chart
NAME2= <i>'string'</i>	specifies name that appears in the name field of the PROC GREPLAY master menu for <i>s</i> chart
PAGENUM= <i>'string'</i>	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

Table 44.21. Star Options

CSTARCIRCLES= <i>color</i>	specifies color for STARCIRCLES= circles
CSTARFILL= <i>color</i> (<i>variable</i>)	specifies color for filling stars
CSTAROUT= <i>color</i>	specifies outline color for stars exceeding inner or outer circles
CSTARS= <i>color</i> (<i>variable</i>)	specifies color for outlines of stars
LSTARCIRCLES= <i>linetypes</i>	specifies line types for STARCIRCLES= circles
LSTARS= <i>linetype</i> (<i>variable</i>)	specifies line types for outlines of STARVERTICES= stars
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 \bar{X} chart
WSTARCIRCLES= <i>n</i>	specifies width of STARCIRCLES= circles
WSTARS= <i>n</i>	specifies width of STARVERTICES= stars

Details

Constructing Charts for Means and Standard Deviations

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
s_i	standard deviation of the measurements x_{i1}, \dots, x_{in_i} in the i^{th} subgroup
	$s_i = \sqrt{((x_{i1} - \bar{X}_i)^2 + \dots + (x_{in_i} - \bar{X}_i)^2)/(n_i - 1)}$
n_i	sample size of i^{th} subgroup
N	number of subgroups
$\bar{\bar{X}}$	weighted average of subgroup means
z_p	100 p^{th} percentile of the standard normal distribution
$c_4(n)$	expected value of the standard deviation of n independent normally distributed variables with unit standard deviation
$c_5(n)$	standard error of the standard deviation of n independent observations from a normal population with unit standard deviation
$\chi_p^2(n)$	100 p^{th} percentile ($0 < p < 1$) of the χ^2 distribution with n degrees of freedom

Plotted Points

Each point on an \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 13, the mean plotted for this subgroup is

$$\bar{X}_{10} = \frac{12 + 15 + 19 + 16 + 13}{5} = 15$$

Each point on an s chart indicates the value of a subgroup standard deviation (s_i). For example, the standard deviation plotted for the tenth subgroup is

$$s_{10} = \sqrt{((12 - 15)^2 + (15 - 15)^2 + (19 - 15)^2 + (16 - 15)^2 + (13 - 15)^2)/4} = 2.739$$

Central Lines

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

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

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

On the s chart, by default, the central line for the i^{th} subgroup indicates an estimate for the expected value of s_i , which is computed as $c_4(n_i)\hat{\sigma}$, 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 $c_4(n_i)\sigma_0$. 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 s_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 s_i exceeds the limits

The following table provides the formulas for the limits:

Table 44.22. Limits for \bar{X} and s Charts

Control Limits	
\bar{X} Chart	LCL = lower limit = $\bar{\bar{X}} - k\hat{\sigma}/\sqrt{n_i}$ UCL = upper limit = $\bar{\bar{X}} + k\hat{\sigma}/\sqrt{n_i}$
s Chart	LCL = lower limit = $\max(c_4(n_i)\hat{\sigma} - kc_5(n_i)\hat{\sigma}, 0)$ UCL = upper limit = $c_4(n_i)\hat{\sigma} + kc_5(n_i)\hat{\sigma}$
Probability Limits	
\bar{X} Chart	LCL = lower limit = $\bar{\bar{X}} - z_{\alpha/2}(\hat{\sigma}/\sqrt{n_i})$ UCL = upper limit = $\bar{\bar{X}} + z_{\alpha/2}(\hat{\sigma}/\sqrt{n_i})$
s Chart	LCL = lower limit = $\hat{\sigma}\sqrt{\chi_{\alpha/2}^2(n_i - 1)/(n_i - 1)}$ UCL = upper limit = $\hat{\sigma}\sqrt{\chi_{1-\alpha/2}^2(n_i - 1)/(n_i - 1)}$

The formulas for s charts assume that the data are normally distributed. If standard values μ_0 and σ_0 are available for μ and σ , respectively, replace $\bar{\bar{X}}$ with μ_0 and $\hat{\sigma}$ with σ_0 in Table 44.22. Note that the limits vary with n_i and that the probability limits for s_i are asymmetric about 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 44.23. 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
LCLS	lower control limit for subgroup standard deviation
LCLX	lower control limit for subgroup mean
LIMITN	nominal sample size associated with the control limits
LSL	lower specification limit
MEAN	process mean (\bar{X} or μ_0)
S	value of central line on s chart
SIGMAS	multiple (k) of standard error of \bar{X}_i or s_i
STDDEV	process standard deviation ($\hat{\sigma}$ or σ_0)
SUBGRP	<i>subgroup-variable</i> specified in the XSCHART statement
TARGET	target value
TYPE	type (estimate or standard value) of _MEAN_ and _STDDEV_
UCLS	upper control limit for subgroup standard deviation
UCLX	upper control limit for subgroup mean
USL	upper specification limit
VAR	<i>process</i> specified in the XSCHART 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_, _LCLS_, _S_, and _UCLS_.
2. If the limits are defined in terms of a multiple k of the standard errors of \bar{X}_i and s_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 `XSCHART` statement. For an example, see “Saving Control Limits” on page 1558.

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 standard deviation variable named by *process* suffixed with *S*
- 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.

Subgroup summary variables are created for each *process* specified in the `XSCHART` statement. For example, consider the following statements:

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

The data set `SUMMARY` contains variables named `LOT`, `WIDTHX`, `WIDTHS`, `WIDTHN`, `DIAMTERX`, `DIAMTERS`, 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 1557.

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
EXLIMS	control limit exceeded on s chart
LCLS	lower control limit for standard deviation
LCLX	lower control limit for mean
LIMITN	nominal sample size associated with the control limits
MEAN	process mean
S	average standard deviation
SIGMAS	multiple (k) of the standard error associated with control limits
<i>subgroup</i>	values of the subgroup variable
SUBN	subgroup sample size
SUBS	subgroup standard deviation
SUBX	subgroup mean
TESTS	tests for special causes signaled on \bar{X} chart
TESTS2	tests for special causes signaled on s chart
UCLS	upper control limit for standard deviation
UCLX	upper control limit for mean
VAR	<i>process</i> specified in the XSCHART 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_, _EXLIMS_, _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 1558.

ODS Tables

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

Table 44.24. ODS Tables Produced with the XSCHEM Statement

Table Name	Description	Options
XSCHEM	\bar{X} and s 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 XSCHEM statement must be a SAS variable in the DATA= data set. This variable provides measurements that must be grouped into subgroup samples indexed by the *subgroup-variable*. The *subgroup-variable*, which is specified in the XSCHEM 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 Standard Deviations from Raw Data” on page 1552.

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;
    xschart 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_`, `_LCLS_`, `_S_`, and `_UCLS_`, 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 44.22 on page 1574

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 1561.

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 XSCHART statement must contain the following variables:

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

Part 9. The CAPABILITY Procedure

- the *subgroup-variable*
- a subgroup mean variable for each *process*
- a subgroup standard deviation variable for each *process*
- a 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* name concatenated with the special suffix characters *X*, *S*, and *N*, respectively. For example, consider the following statements:

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

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

Note that if you specify a *process* name that contains eight characters, the names of 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 of 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 Standard Deviations from Summary Data” on page 1555.

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. 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. Examples are provided in Chapter 49, “Specialized Control Charts.”

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

Table 44.25. Variables Required in a TABLE= Data Set

Variable	Description
LCLS	lower control limit for standard deviation
LCLX	lower control limit for mean
LIMITN	nominal sample size associated with the control limits
MEAN	process mean
S	average standard deviation
<i>subgroup-variable</i>	values of the <i>subgroup-variable</i>
SUBN	subgroup sample size
SUBS	subgroup standard deviation
SUBX	subgroup mean
UCLS	upper control limit for standard deviation
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 standard deviations 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 1558.

Methods for Estimating the Standard Deviation

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

Default Method

The default estimate for σ is

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

where N is the number of subgroups for which $n_i \geq 2$, s_i is the sample standard deviation of the i^{th} subgroup

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

and

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

Here, $\Gamma(\cdot)$ 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 $c_4(n_i)\sigma$. 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* (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). This estimate is a weighted average of N unbiased estimates of σ of the form $s_i/c_4(n_i)$, and it is computed as

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

where

$$h_i = \frac{[c_4(n_i)]^2}{1 - [c_4(n_i)]^2}$$

A subgroup standard deviation s_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 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.

RMSDF Method

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

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

The weights are the degrees of freedom $n_i - 1$. A subgroup standard deviation s_i is included in the calculation only if $n_i \geq 2$, and N is the number of subgroups for which $n_i \geq 2$.

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, in process control applications it is generally not assumed that σ is constant, and if σ varies across subgroups, the root-mean-square estimate tends to be more inflated than the MVLUE.

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	<i>subgroup-variable</i>
Vertical (\bar{X} chart)	DATA=	<i>process</i>
Vertical (\bar{X} chart)	HISTORY=	subgroup mean variable
Vertical (\bar{X} chart)	TABLE=	<code>_SUBX_</code>

You can specify distinct labels for the vertical axes of the \bar{X} and s 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 s chart.

For example, the following sets of statements specify the label *Avg Power Output* for the vertical axis of the \bar{X} chart and the label *Std Deviation* for the vertical axis of the s chart:

```
proc shewhart data=turbine;
  xschart kwatts*day / split = '/' ;
  label kwatts = 'Avg Power Output/Std Deviation';
run;

proc shewhart history=turbhist;
  xschart kwatts*day / split = '/' ;
  label kwattsx = 'Avg Power Output/Std Deviation';
run;

proc shewhart table=turbtabs;
  xschart kwatts*day / split = '/' ;
  label _SUBX_ = 'Avg Power Output/Std Deviation';
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

This section provides advanced examples of the XSCHEM statement.

Example 44.1. Specifying Probability Limits

This example illustrates how to create \bar{X} and s charts with probability limits. The following statements read the kilowatt power output measurements from the data set TURBINE (see page 1552) and create the \bar{X} and s charts shown in Output 44.1.1:

See SHWXS2
in the SAS/QC
Sample Library

```

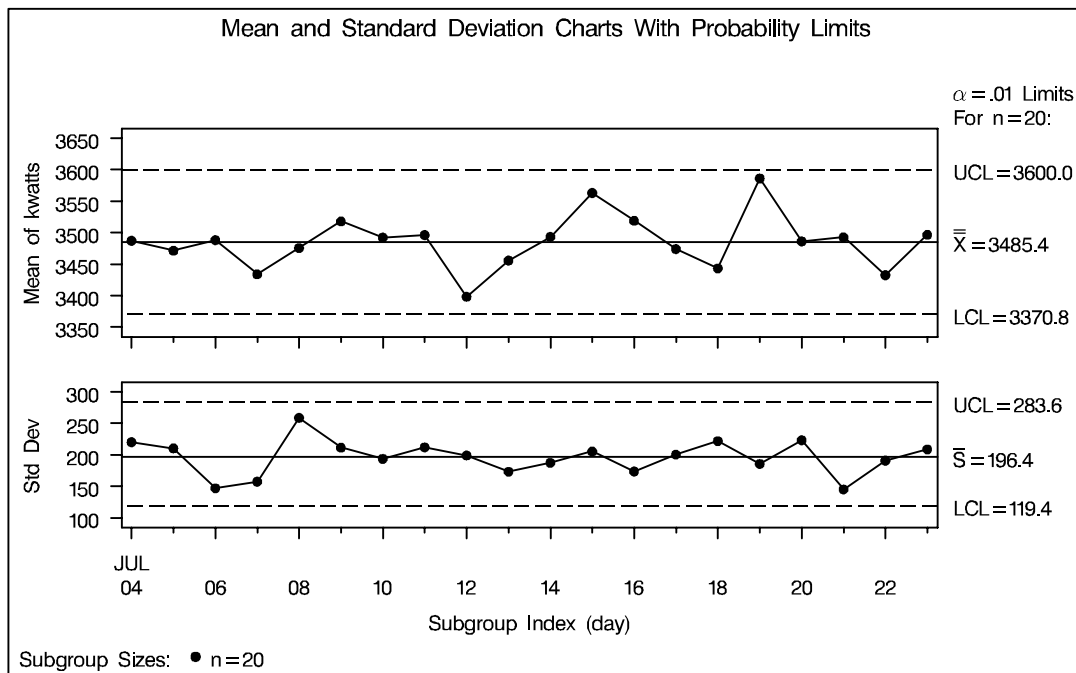
title1 'Mean and Standard Deviation Charts With Probability Limits';
symbol v=dot;
proc shewhart data=turbine;
    xschart kwatts*day / alpha      = 0.01
                        outlimits = oillim;
run;

```

The ALPHA= option specifies the probability (α) that a subgroup summary statistic is outside the limits. Here, the limits are computed so that the probability that a subgroup mean or standard deviation is less than its lower limit is $\alpha/2 = 0.005$, and the probability that a subgroup mean or standard deviation is greater than its upper limit is $\alpha/2 = 0.005$. This assumes that the measurements are normally distributed.

The OUTLIMITS= option names an output data set (OILSUM) that saves the probability limits. The data set OILLIM is shown in Output 44.1.2.

Output 44.1.1. Probability Limits on \bar{X} and s Charts



Output 44.1.2. Probability Limit Information

Mean and Standard Deviation Charts with Probability Limits						
<u>_VAR_</u>	<u>_SUBGRP_</u>	<u>_TYPE_</u>	<u>_LIMITN_</u>	<u>_ALPHA_</u>	<u>_SIGMAS_</u>	<u>_LCLX_</u>
kwatts	day	ESTIMATE	20	0.01	2.57583	3370.79
<u>_MEAN_</u>	<u>_UCLX_</u>	<u>_LCLS_</u>	<u>_S_</u>	<u>_UCLS_</u>	<u>_STDDEV_</u>	
3485.41	3600.03	119.432	196.396	283.570	198.996	

The variable _ALPHA_ saves the value of α . The value of the variable _SIGMAS_ is computed as $k = \Phi^{-1}(1 - \alpha/2)$, where Φ^{-1} is the inverse standard normal distribution function. Note that, in this case, the probability limits for the mean are equivalent to 2.58σ limits.

Since all the points fall within the probability limits, it can be concluded that the process is in statistical control.

Example 44.2. Computing Subgroup Summary Statistics

See SHWXS3
in the SAS/QC
Sample Library

You can use output data sets from a number of SAS procedures as input data sets for the SHEWHART procedure. In this example, the MEANS procedure is used to create a data set containing subgroup summary statistics, which can be read by the SHEWHART procedure as a HISTORY= data set. The following statements create an output data set named OILSUMM, which contains subgroup means, standard deviations, and sample sizes for the variable KWATTS in the data set TURBINE (see page 1552):

```
proc means data=turbine noprint;
  var kwatts;
  by day;
  output out=oilsumm mean=means std=stds n=sizes;
run;
```

A listing of OILSUMM is shown in Output 44.2.1.

Output 44.2.1. The Data Set OILSUMM

Summary Statistics for Power Output Data				
day	kwattsx	kwattss	kwattsn	
04JUL	3487.40	220.260	20	
05JUL	3471.65	210.427	20	
06JUL	3488.30	147.025	20	
07JUL	3434.20	157.637	20	
08JUL	3475.80	258.949	20	
09JUL	3518.10	211.566	20	
10JUL	3492.65	193.779	20	
11JUL	3496.40	212.024	20	
12JUL	3398.50	199.201	20	
13JUL	3456.05	173.455	20	
14JUL	3493.60	187.465	20	
15JUL	3563.30	205.472	20	
16JUL	3519.05	173.676	20	
17JUL	3474.20	200.576	20	
18JUL	3443.60	222.084	20	
19JUL	3586.35	185.724	20	
20JUL	3486.45	223.474	20	
21JUL	3492.90	145.267	20	
22JUL	3432.80	190.994	20	
23JUL	3496.90	208.858	20	

The variables MEANS, STDS, and SIZES do not follow the naming convention required for HISTORY= data sets (see “HISTORY= Data Set” on page 1579). The following statements temporarily rename these variables to KWATTSSX, KWATTSS, and KWATTSN, respectively (the names required when the *process* KWATTS is specified in the XSCHART statement):

```

title 'Mean and Standard Deviation Charts for Power Output';
symbol v=dot;
proc shewhart
    history=oilsumm (rename=(means = kwattsx
                             stds   = kwattss
                             sizes  = kwattsn ));
    xschart kwatts*day;
run;

```

The resulting charts are identical to the charts in Figure 44.2 on page 1554.

Example 44.3. Analyzing Nonnormal Process Data

See SHWXS4 in the SAS/QC Sample Library

The standard control limits for s charts (see Table 44.22 on page 1574) are calculated under the assumption that the data are normally distributed. This example illustrates how a transformation to normality can be used in conjunction with \bar{X} and s charts.

The length of a metal brace is measured in centimeters for each of 20 braces sampled daily. Subgroup samples are collected for nineteen days, and the data are analyzed to determine if the manufacturing process is in statistical control.

```

data lengdata;
  informat day date7.;
  format day date5.;
  label length='Brace Length (cm)';
  input day @;
  do i=1 to 5;
    input length @;
    output;
  end;
  drop i;
  datalines;
02JAN86  113.64  119.60  111.66  111.88  125.29
02JAN86  114.08  115.28  127.84  109.97  109.34
02JAN86  109.65  121.76  112.17  116.01  111.64
02JAN86  112.70  114.43  110.27  114.76  125.89
03JAN86  115.92  113.62  117.52  114.44  118.08
03JAN86  111.13  118.42  112.16  112.25  107.71
03JAN86  110.46  113.78  109.89  114.59  116.98
03JAN86  111.06  113.76  115.53  110.88  115.47
.         .         .         .         .         .
.         .         .         .         .         .
.         .         .         .         .         .
20JAN86  112.84  109.01  110.03  114.44  115.21
20JAN86  115.15  112.34  114.99  109.70  111.20
20JAN86  117.81  119.51  109.03  111.61  118.01
20JAN86  113.55  114.78  112.91  111.87  118.54
;

```

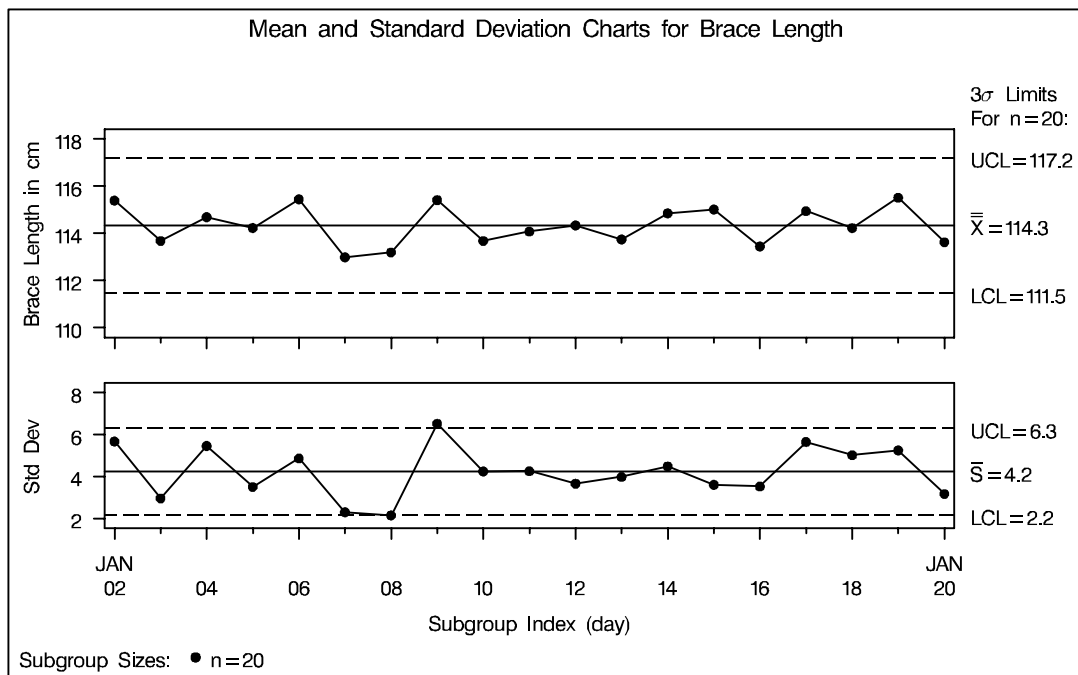
The following statements create preliminary \bar{X} and s charts for the lengths:

```

title 'Mean and Standard Deviation Charts for Brace Length';
symbol v=dot;
proc shewhart data=lengdata;
  xschart length*day;
run;

```

The charts are shown in Output 44.3.1.

Output 44.3.1. \bar{X} and s Charts

The s chart suggests that the process is not in control, since the standard deviation of the measurements recorded on January 9 exceeds its upper control limit. In addition, a number of other points on the s chart are close to the control limits.

The following statements create a box chart for the lengths (for more information on box charts, see Chapter 32, “BOXCHART Statement”).

```

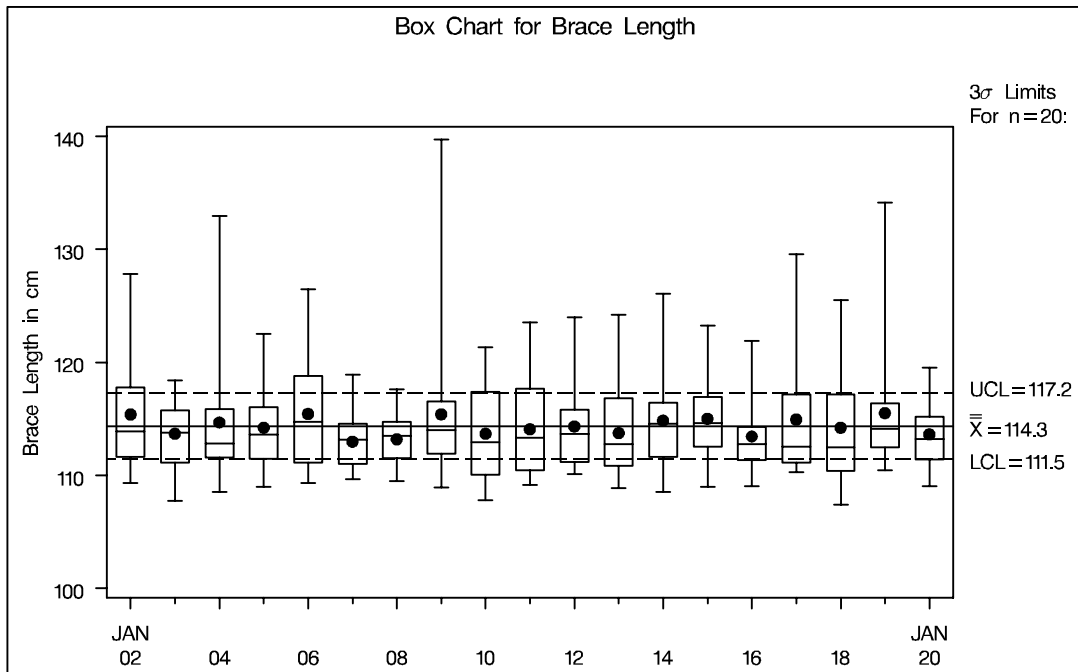
title 'Box Chart for Brace Length';
proc shewhart data=lengdata;
    boxchart length*day / serifs
                    nohlabel
                    nolegend;
run;

```

The chart, shown in Output 44.3.2, reveals that most of the subgroup distributions are skewed to the right. Consequently, the s chart shown in Output 44.3.1 should be interpreted with caution, since control limits for s charts are based on the assumption that the data are normally distributed.

No special cause for the skewness of the subgroup distributions is discovered. This indicates that the process is in statistical control and that the length distribution is naturally skewed.

Output 44.3.2. Box Chart



The following statements apply a lognormal transformation to the length measurements and display a box chart for the transformed data:

```
data lengdata;
  set lengdata;
  logleng=log(length-105);
  label logleng='log(Length-105)';
run;

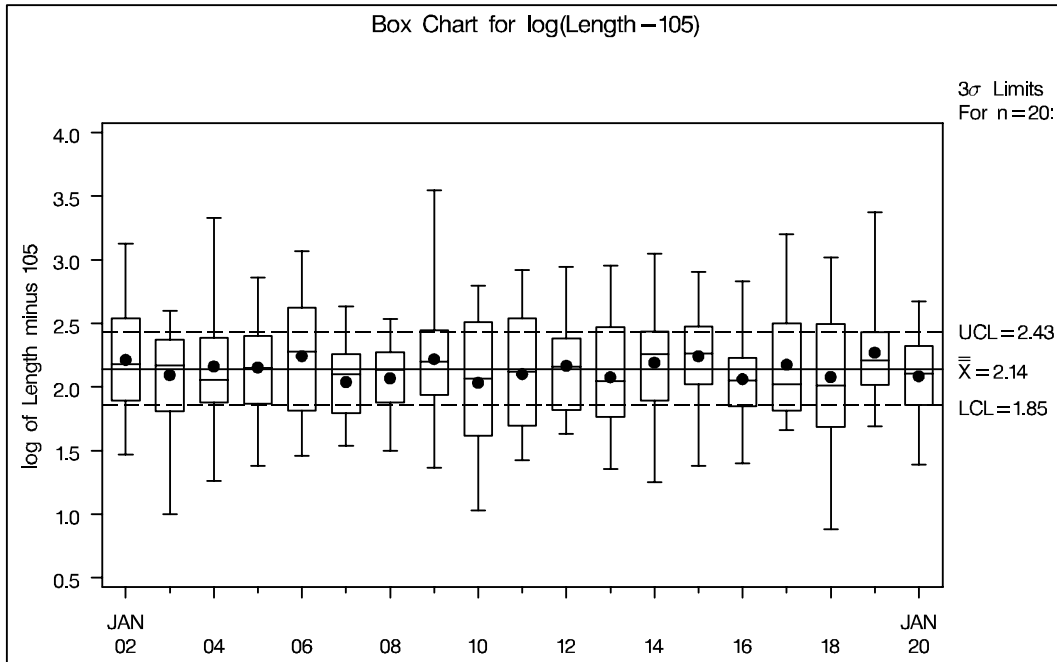
title 'Box Chart for log(Length-105)';
proc shewhart data=lengdata;
  boxchart logleng*day / serifs
                    nohlabel
                    nolegend;
run;
```

The chart, shown in Output 44.3.3, indicates that the subgroup distributions of LOGLENG are approximately normal (this can be verified with goodness-of-fit tests by using the CAPABILITY procedure).

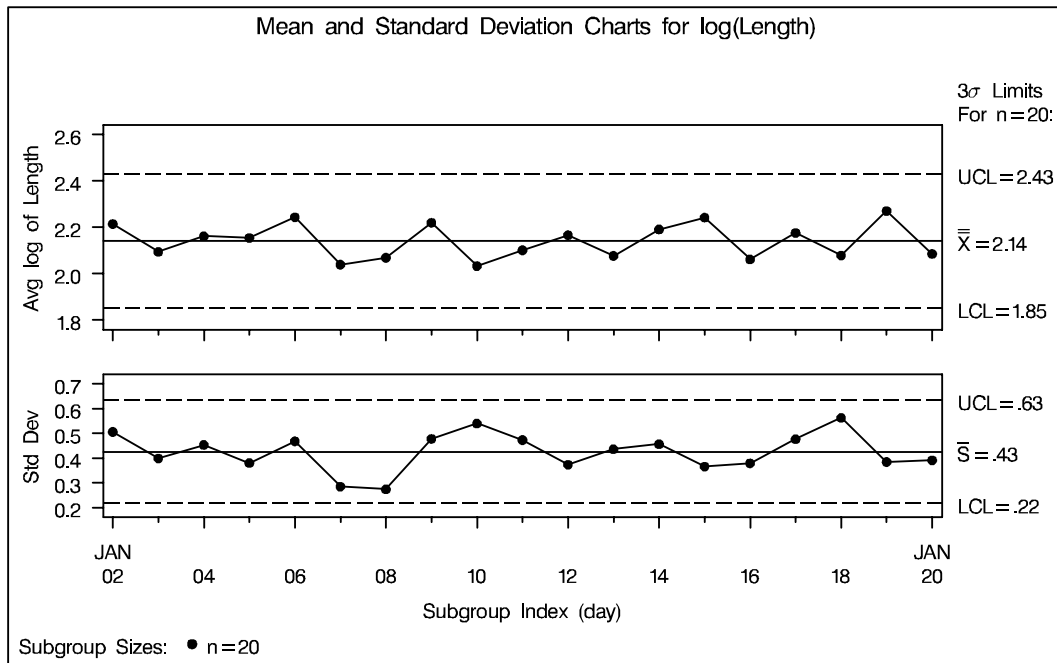
Finally, \bar{X} and s charts, shown in Output 44.3.4, are created for LOGLENG. They indicate that the variability and mean level of the transformed lengths are in control.

```
title 'Mean and Standard Deviation Charts for log(Length)';
proc shewhart data=lengdata;
  xschart logleng*day / split='//';
  label logleng='Avg log(Length)/Std Dev';
run;
```

Output 44.3.3. Box Chart for Transformed Data



Output 44.3.4. \bar{X} and s Charts for Transformed Length



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.