Chapter 21
MACHART Statement

Chapter Table of Contents

Chapter 21
MACHART Statement

Overview

The MACHART statement creates a uniformly weighted moving average control chart (commonly referred to as a moving average control chart), which is used to decide whether a process is in a state of statistical control and to detect shifts in the process average.

You can use options in the MACHART statement to

- specify the span of the moving averages (the number of terms in the moving average)

- compute control limits from the data based on a multiple of the standard error of the plotted moving averages or as probability limits

- tabulate the moving averages, subgroup sample sizes, subgroup means, subgroup standard deviations, control limits, and other information

- save control limit parameters in an output data set

- save the moving averages, subgroup sample sizes, subgroup means, and subgroup standard deviations in an output data set

- read control limit parameters from an input data set

- specify one of several methods for estimating the process standard deviation

- specify a known (standard) process mean and standard deviation for computing control limits

- display a secondary chart that plots a time trend that has been removed from the data

- 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 chart 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 MACHART statement with simple examples that illustrate the most commonly used options. Complete syntax for the MACHART statement is presented in the "Syntax" section on page 674, and advanced examples are given in the "Examples" section on page 702.

Creating Moving Average Charts from Raw Data

See MACMA1
in the SAS/QC
Sample Library

In the manufacture of a metal clip, the gap between the ends of the clip is a critical dimension. To monitor the process for a change in the average gap, subgroup samples of five clips are selected daily. The data are analyzed with a uniformly weighted moving average chart. The gaps recorded during the first twenty days are saved in a SAS data set named CLIPS1.

```
data clips1;
   input day @ ;
   do i=1 to 5;
      input gap @ ;
      output;
      end;
   drop i;
   datalines;
 1  14.76  14.82  14.88  14.83  15.23
 2  14.95  14.91  15.09  14.99  15.13
 3  14.50  15.05  15.09  14.72  14.97
 4  14.91  14.87  15.46  15.01  14.99
 5  14.73  15.36  14.87  14.91  15.25
 6  15.09  15.19  15.07  15.30  14.98
 7  15.34  15.39  14.82  15.32  15.23
 8  14.80  14.94  15.15  14.69  14.93
 9  14.67  15.08  14.88  15.14  14.78
10  15.27  14.61  15.00  14.84  14.94
11  15.34  14.84  15.32  14.81  15.17
12  14.84  15.00  15.13  14.68  14.91
13  15.40  15.03  15.05  15.03  15.18
14  14.50  14.77  15.22  14.70  14.80
15  14.81  15.01  14.65  15.13  15.12
16  14.82  15.01  14.82  14.83  15.00
17  14.89  14.90  14.60  14.40  14.88
18  14.90  15.29  15.14  15.20  14.70
19  14.77  14.60  14.45  14.78  14.91
20  14.80  14.58  14.69  15.02  14.85
;
```

The following statements produce the listing of the data set CLIPS1 shown in Figure 21.1:

```
title 'The Data Set CLIPS1';
proc print data=clips1 noobs;
run;
```

```
                    The Data Set CLIPS1
                 day      gap

                   1     14.76
                   1     14.82
                   1     14.88
                   1     14.83
                   1     15.23
                   2     14.95
                   2     14.91
                   2     15.09
                   2     14.99
                   2     15.13
                   .      .
                   .      .
                   .      .
                  20     14.80
                  20     14.58
                  20     14.69
                  20     15.02
                  20     14.85
```

Figure 21.1.   Partial Listing of the Data Set CLIPS1

The data set CLIPS1 is said to be in "strung-out" form, since each observation contains the day and gap measurement of a single clip. The first five observations contain the gap measurements for the first day, the second five observations contain the gap measurements 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 GAP contains the gap measurements and is referred to as the *process variable* (or *process* for short).

The within-subgroup variability of the gap measurements is known to be stable. You can use a uniformly weighted moving average chart to determine whether the mean level is in control. The following statements create the chart shown in Figure 21.2:

```
title 'Moving Average Chart for Gap Measurements';
symbol v=dot;
proc macontrol data=clips1;
   machart gap*day / span=3;
run;
```

This example illustrates the basic form of the MACHART statement. After the keyword MACHART, you specify the *process* to analyze (in this case, GAP) followed by an asterisk and the *subgroup-variable* (DAY). The SPAN= option specifies the number of terms to include in the moving average. Options such as SPAN= are specified after the slash (/) in the MACHART statement. A complete list of options is presented in the "Syntax" section on page 674. You must provide the span of the moving average. As an alternative to specifying the SPAN= option, you can read the span from an input data set; see "Reading Preestablished Control Limit Parameters" on page 672.

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

665
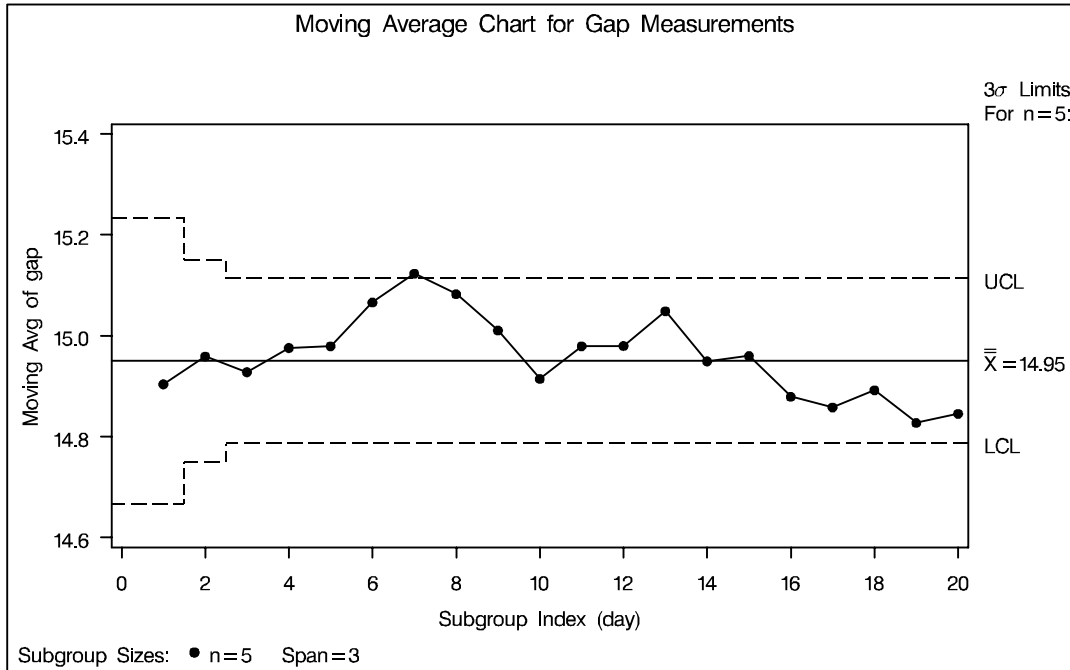
Figure 21.2. Uniformly Weighted Moving Average Chart for Gap Data

Each point on the chart represents the uniformly weighted moving average for a particular day. The moving average A_1 plotted at DAY=1 is simply the subgroup mean for DAY=1. The moving average A_2 plotted at DAY=2 is the average of the subgroup means for DAY=1 and DAY=2. The moving average A_3 plotted at DAY=3 is the average of the subgroup means for DAY=1, DAY=2, and DAY=3.

$$A_1 = \frac{14.76 + 14.82 + 14.88 + 14.83 + 15.23}{5} = 14.904 \text{ mm}$$

$$A_2 = \frac{14.904 + 15.014}{2} = 14.959 \text{ mm}$$

$$A_3 = \frac{14.904 + 15.014 + 14.866}{3} = 14.928 \text{ mm}$$

For succeeding days, the moving average is similarly calculated as the average of the present and the two previous subgroup means (since a span of three is specified with the SPAN= option).

Note that the moving average for the seventh day lies above the upper control limit, signaling an out-of-control process.

By default, the control limits shown are 3σ limits estimated from the data; the formulas for the limits are given in Table 21.19 on page 688.

For computational details, see "Constructing Uniformly Weighted Moving Average Charts" on page 687. For more details on reading from a DATA= data set, see "DATA= Data Set" on page 696.

Creating Moving Average Charts from Subgroup Summary Data

The previous example illustrates how you can create moving average 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 MACHART statement with data of this type. The following data set (CLIPSUM) provides the data from the preceding example in summarized form:

```
data clipsum;
   input day gapx gaps;
   gapn=5;
datalines;
 1  14.904  0.18716
 2  15.014  0.09317
 3  14.866  0.25006
 4  15.048  0.23732
 5  15.024  0.26792
 6  15.126  0.12260
 7  15.220  0.23098
 8  14.902  0.17254
 9  14.910  0.19824
10  14.932  0.24035
11  15.096  0.25618
12  14.912  0.16903
13  15.138  0.15928
14  14.798  0.26329
15  14.944  0.20876
16  14.896  0.09965
17  14.734  0.22512
18  15.046  0.24141
19  14.702  0.17880
20  14.788  0.16634
;
```

A partial listing of CLIPSUM is shown in Figure 21.3. There is exactly one observation for each subgroup (note that the subgroups are still indexed by DAY). The variable GAPX contains the subgroup means, the variable GAPS contains the subgroup standard deviations, and the variable GAPN contains the subgroup sample sizes (these are all five).

```
               The Data Set CLIPSUM


        day      gapx        gaps      gapn

         1      14.904     0.18716       5
         2      15.014     0.09317       5
         3      14.866     0.25006       5
         .        .          .          .
         .        .          .          .
         .        .          .          .
        20      14.788     0.16634       5
```

Figure 21.3. The Summary Data Set CLIPSUM

You can read this data set by specifying it as a HISTORY= data set in the PROC MACONTROL statement, as follows:
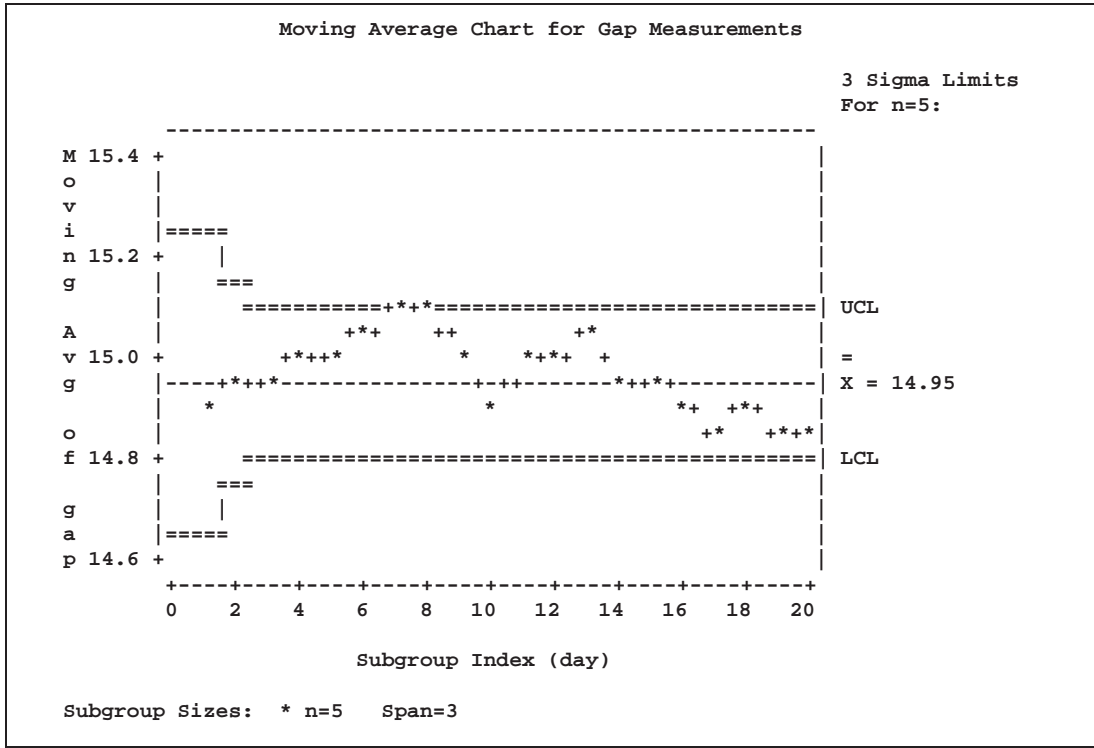
```
title 'Moving Average Chart for Gap Measurements';
proc macontrol history=clipsum lineprinter;
   machart gap*day='*' / span=3;
run;
```

The resulting moving average chart is shown in Figure 21.4. Since the LINEPRINTER option is specified in the PROC MACONTROL statement, line printer output is produced. The asterisk (*) specified in single quotes after the *subgroup-variable* indicates the character used to plot points. This character must follow an equal sign.

Note that GAP is *not* the name of a SAS variable in the data set but is, instead, the common prefix for the names of the three SAS variables GAPX, GAPS, and GAPN. The suffix characters *X*, *S*, and *N* indicate *mean*, *standard deviation*, and *sample size*, respectively. Thus, you can specify three subgroup summary variables in a HISTORY= data set with a single name (GAP), which is referred to as the *process*. The variables GAPX, GAPS, and GAPN are all required. The name DAY specified after the asterisk is the name of the *subgroup-variable*.

```
                 Moving Average Chart for Gap Measurements

                                                    3 Sigma Limits
                                                    For n=5:

             -------------------------------------------------
 M 15.4 +                                                     |
 o      |                                                     |
 v      |                                                     |
 i      |=====                                                |
 n 15.2 +    |                                                |
 g      |    ===                                              |
        |      ===========+*+*=========================|  UCL
 A      |             +*+    ++        +*                     |
 v 15.0 +         +*++*         *    *+*+  +         | =
 g      |----+*++*---------------+-++-------*++*+-----------|  X = 14.95
        |    *                 *              *+  +*+        |
 o      |                                     +*    +*+*|
 f 14.8 +      =======================================|  LCL
        |    ===                                              |
 g      |    |    |                                           |
 a      |=====                                                |
 p 14.6 +                                                     |
         +----+----+----+----+----+----+----+----+----+----+
         0    2    4    6    8   10   12   14   16   18   20

                         Subgroup Index (day)

 Subgroup Sizes:  * n=5   Span=3
```

Figure 21.4. Uniformly Weighted Moving Average Chart from Summary Data

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

668

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

Furthermore, the names of subgroup mean, standard deviation, and sample size variables must begin with the *process* name specified in the MACHART 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 in the PROC MACONTROL statement to rename the variables for the duration of the MACONTROL procedure step (see page 1507 for an example).

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 697.

Saving Summary Statistics

In this example, the MACHART statement is used to create a summary data set that can be read later by the MACONTROL procedure (as in the preceding example). The following statements read measurements from the data set CLIPS1 and create a summary data set named CLIPHIST:

See MACMA1 in the SAS/QC Sample Library

```
title 'Summary Data Set for Gap Measurements';
proc macontrol data=clips1;
   machart gap*day / span      = 3
                     outhistory = cliphist
                     nochart;
run;
```

The OUTHISTORY= option names the output data set, and the NOCHART option suppresses the display of the chart, which would be identical to the chart in Figure 21.2.

Figure 21.5 contains a partial listing of CLIPHIST.

```
       Summary Data Set for Gap Measurements


   day     gapX       gapS       gapA      gapN

     1    14.904    0.18716    14.9040      5
     2    15.014    0.09317    14.9590      5
     3    14.866    0.25006    14.9280      5
     .       .         .          .         .
     .       .         .          .         .
     .       .         .          .         .
    20    14.788    0.16634    14.8453      5
```

Figure 21.5. The Summary Data Set CLIPHIST

There are five variables in the data set CLIPHIST.

- DAY contains the subgroup index.
- GAPX contains the subgroup means.
- GAPS contains the subgroup standard deviations.
- GAPA contains the subgroup moving averages.
- GAPN contains the subgroup sample sizes.

Note that the summary statistic variables are named by adding the suffix characters *X*, *S*, *A*, and *N* to the *process* GAP specified in the MACHART 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 693.

Saving Control Limit Parameters

See MACMA1
in the SAS/QC
Sample Library

You can save the control limit parameters used for a moving average chart in a SAS data set; this enables you to use these parameters with future data (see "Reading Preestablished Control Limit Parameters" on page 672) or modify the parameters with a DATA step program.

The following statements read measurements from the data set CLIPS1 (see page 664) and save the control limit parameters in a data set named CLIPLIM:

```
title 'Control Limit Parameters';
proc macontrol data=clips1;
   machart gap*day / span      = 3
                     outlimits = cliplim
                     nochart;
run;
```

The OUTLIMITS= option names the data set containing the control limits, and the NOCHART option suppresses the display of the chart. The data set CLIPLIM is listed in Figure 21.6.

```
                        Control Limit Parameters

_VAR_  _SUBGRP_   _TYPE_   _LIMITN_    _ALPHA_  _SIGMAS_  _MEAN_  _STDDEV_  _SPAN_

 gap     day     ESTIMATE     5      .002699796     3     14.95   0.21108     3
```

Figure 21.6. The Data Set CLIPLIM Containing Control Limit Information

Note that the data set CLIPLIM does not contain the actual control limits, but rather the parameters required to compute the limits.

The data set contains one observation with the parameters for *process* GAP. The variable _SPAN_ contains the number of terms used to calculate the moving average. 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 that the values of _MEAN_ and _STDDEV_ are estimates rather than standard values. For more information, see "OUTLIMITS= Data Set" on page 693.

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

```
title 'Summary Statistics and Control Limits';
proc macontrol data=clips1;
   machart gap*day / span     = 3
                     outtable = cliptab
                     nochart;
run;
```

The data set CLIPTAB is listed in Figure 21.7.

This data set contains one observation for each subgroup sample. The variable _UWMA_ contains the uniformly weighted moving average. The variables _SUBX_, _SUBS_, and _SUBN_ contain the subgroup means, subgroup standard deviations, and subgroup sample sizes, respectively. The variables _LCLA_ and _UCLA_ contain the lower and upper control limits, and the variable _MEAN_ contains the central line. The variables _VAR_ and DAY contain the *process* name and values of the *subgroup-variable*, respectively. For more information, see "OUTTABLE= Data Set" on page 694.

```
            Summary Statistics and Control Limits

      _ _
      S L
      I I _ _    _      _      _      _      _      _      _
      G M S S    S      S      L      U      M      U      E
 _    M I P U    U      U      C      W      E      C      X
 V  d A T A B    B      B      L      M      A      C      L
 A  a S N N N    X      S      L      M      A      L      I
 R  Y _ _ _ _    _      _      _      _      _      _      M
 _  _                                                     _

gap  1 3 5 3 5 14.904 0.18716 14.6668 14.9040 14.95 15.2332
gap  2 3 5 3 5 15.014 0.09317 14.7498 14.9590 14.95 15.1502
gap  3 3 5 3 5 14.866 0.25006 14.7865 14.9280 14.95 15.1135
gap  4 3 5 3 5 15.048 0.23732 14.7865 14.9760 14.95 15.1135
gap  5 3 5 3 5 15.024 0.26792 14.7865 14.9793 14.95 15.1135
gap  6 3 5 3 5 15.126 0.12260 14.7865 15.0660 14.95 15.1135
gap  7 3 5 3 5 15.220 0.23098 14.7865 15.1233 14.95 15.1135 UPPER
gap  8 3 5 3 5 14.902 0.17254 14.7865 15.0827 14.95 15.1135
gap  9 3 5 3 5 14.910 0.19824 14.7865 15.0107 14.95 15.1135
gap 10 3 5 3 5 14.932 0.24035 14.7865 14.9147 14.95 15.1135
gap 11 3 5 3 5 15.096 0.25618 14.7865 14.9793 14.95 15.1135
gap 12 3 5 3 5 14.912 0.16903 14.7865 14.9800 14.95 15.1135
gap 13 3 5 3 5 15.138 0.15928 14.7865 15.0487 14.95 15.1135
gap 14 3 5 3 5 14.798 0.26329 14.7865 14.9493 14.95 15.1135
gap 15 3 5 3 5 14.944 0.20876 14.7865 14.9600 14.95 15.1135
gap 16 3 5 3 5 14.896 0.09965 14.7865 14.8793 14.95 15.1135
gap 17 3 5 3 5 14.734 0.22512 14.7865 14.8580 14.95 15.1135
gap 18 3 5 3 5 15.046 0.24141 14.7865 14.8920 14.95 15.1135
gap 19 3 5 3 5 14.702 0.17880 14.7865 14.8273 14.95 15.1135
gap 20 3 5 3 5 14.788 0.16634 14.7865 14.8453 14.95 15.1135
```

Figure 21.7. The OUTTABLE= Data Set CLIPTAB

An OUTTABLE= data set can be read later as a TABLE= data set. For example, the following statements read CLIPTAB and display a moving average chart (not shown here) identical to Figure 21.2:

```
title 'Moving Average Chart for Gap Measurements';
proc macontrol table=cliptab;
   machart gap*day;
run;
```

For more information, see "TABLE= Data Set" on page 698.

Reading Preestablished Control Limit Parameters

See MACMA1
in the SAS/QC
Sample Library

In the previous example, the OUTLIMITS= data set saved the control limit parameters in the data set CLIPLIM. This example shows how to apply these parameters to new data provided in the following data set:

```
data clips1a;
   label gap='Gap Measurement (mm)';
   input day @;
   do i=1 to 5;
      input gap @;
      output;
      end;
   drop i;
datalines;
21  14.86 15.01 14.67 14.67 15.07
22  14.93 14.53 15.07 15.10 14.98
23  15.27 14.90 15.12 15.10 14.80
24  15.02 15.21 14.93 15.11 15.20
25  14.90 14.81 15.26 14.57 14.94
26  14.78 15.29 15.13 14.62 14.54
27  14.78 15.15 14.61 14.92 15.07
28  14.92 15.31 14.82 14.74 15.26
29  15.11 15.04 14.61 15.09 14.68
30  15.00 15.04 14.36 15.20 14.65
31  14.99 14.76 15.18 15.04 14.82
32  14.90 14.78 15.19 15.06 15.06
33  14.95 15.10 14.86 15.27 15.22
34  15.03 14.71 14.75 14.99 15.02
35  15.38 14.94 14.68 14.77 14.83
36  14.95 15.43 14.87 14.90 15.34
37  15.18 14.94 15.32 14.74 15.29
38  14.91 15.15 15.06 14.78 15.42
39  15.34 15.34 15.41 15.36 14.96
40  15.12 14.75 15.05 14.70 14.74
   ;
```

672

The following statements create a moving average chart for the data in CLIPS1A using the control limit parameters in CLIPLIM:

```
title 'Moving Average Chart for Second Set of Gap Measurements';
symbol v=dot;
proc macontrol data=clips1a limits=cliplim;
    machart gap*day;
run;
```

The chart is shown in Figure 21.8.

The LIMITS= option in the PROC MACONTROL statement specifies the data set containing the control limits parameters. 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 GAP
- the value of _SUBGRP_ matches the *subgroup-variable* name DAY
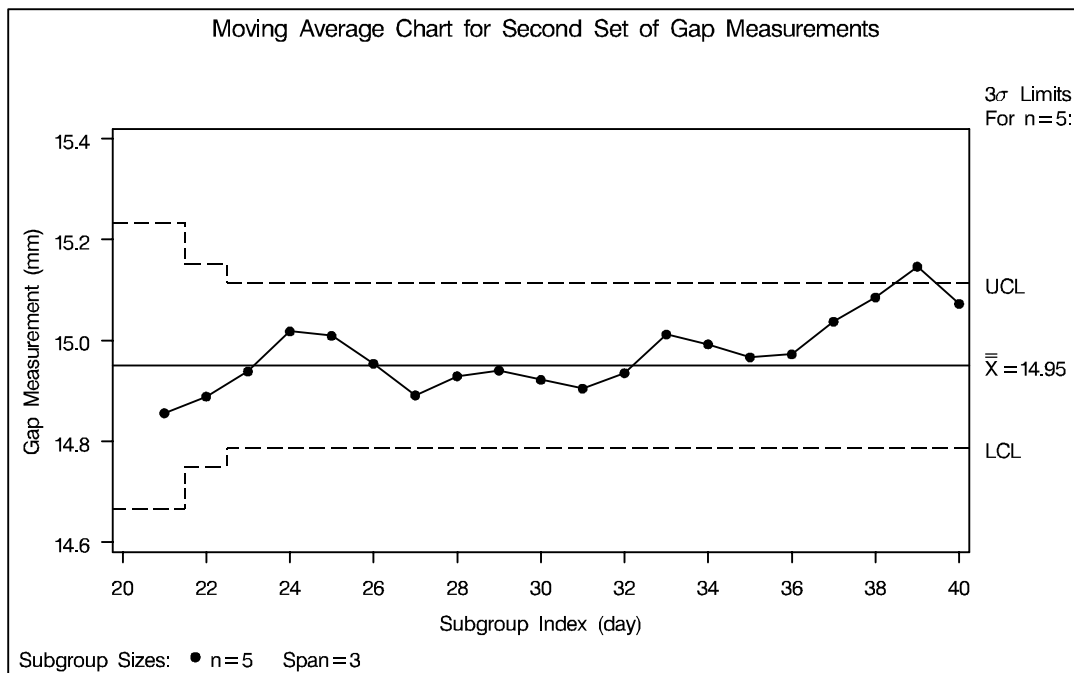
Figure 21.8. Using Control Limit Parameters from a LIMITS= Data Set

Note that the moving average plotted for the 39th day lies above the upper control limit, signalling an out-of-control process.

In this example, the LIMITS= data set was created in a previous run of the MACONTROL procedure. You can also create a LIMITS= data set with the DATA step. See "LIMITS= Data Set" on page 697 for details concerning the variables that you must provide, and see Example 21.1 on page 702 for an illustration.

---

[*]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 MACHART statement is as follows:

MACHART *process*subgroup-variable* **/ SPAN=***value* < *options* > ;

The general form of this syntax is as follows:

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

Note that the SPAN= option is required unless its *value* is read from a LIMITS= data set. You can use any number of MACHART statements in the MACONTROL procedure. The components of the MACHART 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 MACONTROL 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 Moving Average Charts from Raw Data" on page 664.

> - 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 Moving Average Charts from Subgroup Summary Data" on page 667.

> - 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 Limit Parameters" on page 670.

> A *process* is required. If more than one *process* is specified, enclose the list in parentheses. For example, the following statements request distinct moving average charts (each with a span of 3) for WEIGHT, LENGTH, and WIDTH:

```
proc macontrol data=measures;
   machart (weight length width)*day / span=3;
run;
```

subgroup-variable
> is the variable that classifies the data into subgroups. The *subgroup-variable* is required. In the preceding MACHART 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 plotting character used to plot the the moving averages.

>   - If you produce a chart on a line printer, an 'A' is displayed for points corresponding to the first level of the *symbol-variable*, a 'B' is displayed for points corresponding to the second level, and so on.

>   - If you produce a chart on a graphics device, distinct symbol markers are displayed for points corresponding to the various levels of the *symbol-variable*. You can specify the symbol markers with SYMBOL*n* statements. 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 a moving average chart using an asterisk (*) to plot the points:

>   ```
>   proc macontrol data=values;
>       machart weight*hour='*' / span=3;
>   run;
>   ```

options

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

Summary of Options

The following tables list the MACHART statement options by function. Options unique to the MACONTROL procedure are listed in Table 21.1 and Table 21.2, and they are described in detail in "Dictionary of Special Options" on page 684. Options that are common to both the MACONTROL and SHEWHART procedures are listed in Table 21.3 to Table 21.18. They are described in detail beginning on page 1613 of Part 9, "The SHEWHART Procedure."

Table 21.1. Options for Specifying Uniformly Weighted Moving Average Charts[*]

ALPHA=*value*	requests probability limits for control charts
ASYMPTOTIC	requests constant control limits
LIMITN=n│VARYING	specifies either a fixed nominal sample size (n) for control limits or allows the control limits to vary with subgroup sample size
MU0=*value*	specifies a standard (known) value μ_0 for the process mean
NOREADLIMITS	specifies that control limit parameters are not to be read from LIMITS= data set (Release 6.10 and later releases)
READALPHA	reads _ALPHA_ instead of _SIGMAS_ from LIMITS= data set when both variables are available
READINDEX='*value*'	reads control limit parameters from the first observation in the LIMITS= data set where the variable _INDEX_ equals *value*
READLIMITS	reads control limit parameters from a LIMITS= data set (Release 6.09 and earlier releases)
SIGMA0=*value*	specifies standard (known) value σ_0 for process standard deviation
SIGMAS=k	specifies width of control limits in terms of multiple k of standard error of plotted moving averages
SPAN=*value*	specifies the number of terms in the moving average

Table 21.2. Options for Plotting Subgroup Means[*]

CMEANSYMBOL=*color*	specifies color for MEANSYMBOL= symbol
MEANCHAR='*character*'	specifies *character* to plot subgroup means on line printer
MEANSYMBOL= *keyword*	specifies symbol to plot subgroup means on graphics device

[*]The options in these tables are described in "Dictionary of Special Options" on page 684.

Table 21.3. Tabulation Options[†]

TABLE	creates a basic table of subgroup variable values, subgroup sample sizes, subgroup means, subgroup moving averages, and control limits
TABLEALL	is equivalent to the options TABLE, TABLECENTRAL, TABLEID, and TABLEOUTLIM
TABLECENTRAL	augments basic table with the value of the central line
TABLEID	augments basic table with columns for ID variables
TABLEOUTLIM	augments basic table with columns indicating control limits exceeded

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

Table 21.4. 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*|AXIS*n*	specifies major tick mark values for horizontal axis
HEIGHT=*value*	specifies height of axis label and axis legend text
HMINOR=*n*	specifies minor tick marks between major horizontal tick marks
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
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*|AXIS*n*	specifies major tick mark values for vertical axis on moving average chart
VAXIS2=*values*|AXIS*n*	specifies major tick mark values for vertical axis on trend chart
VMINOR=*n*	specifies minor tick marks between major vertical tick marks
VOFFSET=*value*	specifies length of offset at both ends of vertical axis
WAXIS=*n*	specifies width of axis lines

Table 21.5. Process Mean and Standard Deviation Options[†]

SMETHOD=*keyword*	specifies method for estimating process standard deviation σ
TYPE=*keyword*	identifies whether parameters are estimates or standard values and specifies value of _TYPE_ in OUTLIMITS= data set

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

677

Table 21.6. Grid Options[†]

ENDGRID	adds grid after last plotted point
GRID	adds grid to 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 21.7. 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 reference lines perpendicular to horizontal axis on moving average chart
HREF2=*values*| *SAS-data-set*	specifies reference lines perpendicular to horizontal axis on trend 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 moving average chart
HREF2DATA= *SAS-data-set*	specifies position of reference lines perpendicular to horizontal axis on trend 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 is to be applied uniformly to charts created for all BY groups
VREF=*values*| *SAS-data-set*	specifies reference lines perpendicular to vertical axis on moving average chart
VREF2=*values*| *SAS-data-set*	specifies reference lines perpendicular to vertical axis on trend 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

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

678

Table 21.8. 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 21.9. 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
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*	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
NOCTL	suppresses display of central line
NOLCL	suppresses display of lower control limit
NOLIMITLABEL	suppresses labels for control limits and center line
NOLIMITS	suppresses display of control limits
NOLIMITSLEGEND	suppresses legend for control limits
NOUCL	suppresses display of upper control limit
UCLLABEL='*string*'	specifies label for upper control limit
WLIMITS=*n*	width for control limits and central line
XSYMBOL='*string*'| *keyword*	specifies label for central line

Table 21.10. 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
WEBOUT=*SAS-data-set*	creates an OUTTABLE= data set with additional graphics co-ordinate data

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

Table 21.11. Options for Plotting and Labeling Points[†]

ALLLABEL=VALUE| (*variable*)	labels every point on moving average chart
ALLLABEL2=VALUE| (*variable*)	labels every point on trend 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 moving average chart
COUT=*color*	specifies color for line segments that connect points exceeding control limits
COUTFILL=*color*	specifies color for areas between connected points and control limits
LABELFONT=*font*	specifies a software font for labels requested by the ALLLABEL=, ALLLABEL2=, OUTLABEL=, and STARLABEL= options
LABELHEIGHT=*font*	specifies the height (in vertical percent screen units) for labels requested by the ALLLABEL=, ALLLABEL2=, OUTLABEL=, and STARLABEL= options
NEEDLES	connects points to central line with vertical needles
NOCONNECT	suppresses line segments that connect points on moving average chart
NOTRENDCONNECT	suppresses line segments that connect points on trend chart
OUTLABEL=VALUE| (*variable*)	labels points exceeding control limits
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	turns point labels so that they are strung out vertically

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

Table 21.12. Input Data Set Options[†]

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

Table 21.13. Output Data Set Options[†]

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

Table 21.14. Plot Layout Options[†]

ALLN	plots moving averages for all subgroups
BILEVEL	creates control charts using half-screens and half-pages
EXCHART	creates control charts 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 moving average 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
TOTPANELS=*n*	specifies number of pages or screens to be used to display chart
TRENDVAR=*variable*| (*variable-list*)	specifies list of trend variables
YPCT1=*value*	specifies length of vertical axis on EWMA chart as a percentage of sum of lengths of vertical axes for EWMA and trend charts
ZEROSTD	displays $\bar{X}$ chart regardless of whether $\hat{\sigma} = 0$

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

SAS OnlineDoc[TM]*: Version 8*

Table 21.15. Phase Options[†]

CPHASELEG=*color*	specifies text color for *phase* legend
OUTPHASE='*string*'	specifies value of _PHASE_ in 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 chart
PHASEREF	delineates *phases* with vertical reference lines
READPHASES= ALL| '*label1*'...'*labeln*'	specifies *phases* to be read from input data set

Table 21.16. Graphical Enhancement Options[†]

ANNOTATE= *SAS-data-set*	specifies annotate data set that adds features to moving average chart
ANNOTATE2= *SAS-data-set*	specifies annotate data set that adds features to trend chart
DESCRIPTION='*string*'	specifies string that appears in the description field of PROC GREPLAY master menu for moving average chart
FONT=*font*	specifies software font for labels and legends on chart
NAME='*string*'	specifies name that appears in the name field of the PROC GRE-PLAY master menu for moving average chart
PAGENUM='*string*'	specifies the form of the label used in pagination
PAGENUMPOS= *keyword*	specifies the position of the page number requested with the PA-GENUM= option

Table 21.17. Clipping Options[†]

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

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

Table 21.18. 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 stars requested with the STARVERTICES= option
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 moving average chart
WSTARCIRCLES=*n*	specifies width of STARCIRCLES= circles
WSTARS=*n*	specifies width of STARVERTICES= stars

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

SAS OnlineDoc[TM]: Version 8

Dictionary of Special Options

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

ALPHA=*value*

requests *probability limits*. If you specify ALPHA=α, the control limits are computed so that the probability is α that a single moving average exceeds its control limits. The value of α can range between 0 and 1. This assumes that the process is in statistical control and that the data follow a normal distribution. For the equations used to compute probability limits, see "Control Limits" on page 687.

Note the following:

- As an alternative to specifying ALPHA=α, you can read α from the variable _ALPHA_ in a LIMITS= data set by specifying the READALPHA option.

- As an alternative to specifying ALPHA=α (or reading _ALPHA_ from a LIMITS= data set), you can request "$k\sigma$ control limits" by specifying SIGMAS=k (or reading _SIGMAS_ from a LIMITS= data set).

If you specify neither the ALPHA= option nor the SIGMAS= option, the procedure computes 3σ control limits by default.

ASYMPTOTIC

requests constant upper and lower control limits for all subgroups having the following values:

$$\text{LCL} = \overline{\overline{X}} - \frac{k\hat{\sigma}}{\sqrt{nw}}$$

$$\text{UCL} = \overline{\overline{X}} + \frac{k\hat{\sigma}}{\sqrt{nw}}$$

Here w is the span of the moving average, and n is the nominal sample size associated with the control limits. Substitute $\Phi^{-1}(1-\alpha/2)$ for k if you specify probability limits with the ALPHA= option. When you do not specify the ASYMPTOTIC option, the control limits are computed using the exact formulas in Table 21.19 on page 688. Use the ASYMPTOTIC option only if all the subgroup sample sizes are the same or if you specify LIMITN=n.

CMEANSYMBOL=*color*

specifies the *color* for the symbol requested with the MEANSYMBOL= option. The default *color* is the first color in the device color list.

LIMITN=*n*
LIMITN=VARYING

specifies either a fixed or varying nominal sample size for the control limits.

If you specify LIMITN=n, moving averages are calculated and displayed only for those subgroups with a sample size equal to n, unless you also specify the ALLN option, which causes all the moving averages to be calculated and displayed. By default (or if you specify LIMITN=VARYING), moving averages are calculated and displayed for all subgroups, regardless of sample size.

MEANCHAR='*character***'**

specifies a *character* used to plot the subgroup mean for each subgroup. By default, subgroup means are not plotted.

MEANSYMBOL=*keyword*

specifies a symbol used to plot the subgroup mean for each subgroup. By default, subgroup means are not plotted.

MU0=*value*

specifies a known (standard) value μ_0 for the process mean μ. By default, μ is estimated from the data.

Note: As an alternative to specifying MU0=μ_0, you can read a predetermined value for μ_0 from the variable _MEAN_ in a LIMITS= data set.

See Example 21.1 on page 702.

NOREADLIMITS

specifies that control limit parameters for each *process* listed in the MACHART statement are *not* to be read from the LIMITS= data set specified in the PROC MACONTROL statement. The NOREADLIMITS option is available only in Release 6.10 and later releases.

The following example illustrates the NOREADLIMITS option:

```
proc macontrol data=pistons limits=diamlim;
   machart diameter*hour;
   machart diameter*hour / noreadlimits span=3;
run;
```

The first MACHART statement reads the control limits from the first observation in the data set DIAMLIM for which the variable _VAR_ is equal to `diameter` and the variable _SUBGRP_ is equal to `hour`. The second MACHART statement computes estimates of the process mean and standard deviation for the control limits from the measurements in the data set PISTONS. Note that the second MACHART statement is equivalent to the following statements, which would be more commonly used:

```
proc macontrol data=pistons;
   machart diameter*hour / span=3;
run;
```

For more information about reading control limit parameters from a LIMITS= data set, see the READLIMITS option later in this list.

READALPHA

specifies that the variable _ALPHA_, rather than the variable _SIGMAS_, is to be read from a LIMITS= data set when both variables are available in the data set. Thus the limits displayed are probability limits. If you do not specify the READALPHA option, then _SIGMAS_ is read by default.

READINDEX='*value*'

reads control limit parameters from a LIMITS= data set (specified in the PROC MA-CONTROL statement) for each *process* listed in the MACHART statement. The control limit parameters for a particular *process* are read from the first observation in the LIMITS= data set for which

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

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

READLIMITS

specifies that control limit parameters are to be read from a LIMITS= data set specified in the PROC MACONTROL statement. The parameters for a particular *process* are read from the first observation in the LIMITS= data set for which

- the value of _VAR_ matches *process*
- the value of _SUBGRP_ matches the *subgroup variable*

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

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

- **In Release 6.09 and earlier releases, you must specify the READLIMITS option to read control limits parameters as described previously.** If you specify a LIMITS= data set without specifying the READLIMITS option (or the READINDEX= option), the control limits are computed from the data and the span of the moving averages is specified with the SPAN= option.

SIGMA0=`value`

specifies a known (standard) value σ_0 for the process standard deviation σ. The *value* must be positive. By default, the MACONTROL procedure estimates σ from the data using the formulas given in "Methods for Estimating the Standard Deviation" on page 699.

Note: As an alternative to specifying SIGMA0=σ_0, you can read a predetermined value for σ_0 from the variable _STDDEV_ in a LIMITS= data set.

SIGMAS=`value`

specifies the width of the control limits in terms of the multiple k of the standard error of the plotted moving averages on the chart. The value of k must be positive. By default, $k = 3$ and the control limits are 3σ limits.

SPAN=`value`

specifies the number of terms used to calculate the moving average (*value* is an integer greater than 1). The SPAN= option is required unless you read control limit parameters from a LIMITS= data set or a TABLE= data set. See "Plotted Points" on page 687 and "Choosing the Span of the Moving Average" on page 689 for details.

Details

Constructing Uniformly Weighted Moving Average Charts

The following notation is used in this section:

A_i	uniformly weighted moving average for the ith subgroup
w	span parameter (number of terms in moving average)
μ	process mean (expected value of the population of measurements)
σ	process standard deviation (standard deviation of the population of measurements)
x_{ij}	j^{th} measurement in i^{th} subgroup, with $j =1, 2, 3, \ldots, n_i$
n_i	sample size of ith subgroup
$\overline{X}_i$	mean of measurements in i^{th} subgroup. If $n_i = 1$, then the subgroup mean reduces to the single observation in the subgroup.
$\overline{\overline{X}}$	weighted average of subgroup means
$\Phi^{-1}(\cdot)$	inverse standard normal function

Plotted Points

Each point on the chart indicates the value of the uniformly weighted moving average for that subgroup. The moving average for the ith subgroup (A_i) is defined as

$$A_i = (\overline{X}_1 + \ldots + \overline{X}_i)/i \qquad \text{if } i < w$$
$$A_i = (\overline{X}_i + \ldots + \overline{X}_{i-w+1})/w \quad \text{if } i \geq w$$

where w is the span, or number of terms, of the moving average. You can specify the span with the SPAN= option in the MACHART statement or with the value of _SPAN_ in a LIMITS= data set.

Central Line

By default, the central line on a moving average chart indicates an estimate for 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.

Control Limits

You can compute the limits in the following ways:
- as a specified multiple (k) of the standard error of A_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 A_i exceeds the limits

The following table presents the formulas for the limits:

Table 21.19. Limits for Moving Average Chart

Control Limits
$\text{LCL} = \overline{\overline{X}} - k(\hat{\sigma}/\min(i, w))\sqrt{(1/n_i) + (1/n_{i-1}) + \ldots + (1/n_{1+\max(i-w,0)})}$
$\text{UCL} = \overline{\overline{X}} + k(\hat{\sigma}/\min(i, w))\sqrt{(1/n_i) + (1/n_{i-1}) + \ldots + (1/n_{1+\max(i-w,0)})}$

Probability Limits
$\text{LCL} = \overline{\overline{X}} - \Phi^{-1}(1 - \alpha/2)(\hat{\sigma}/\min(i, w))\sqrt{(1/n_i) + (1/n_{i-1}) + \ldots + (1/n_{1+\max(i-w,0)})}$
$\text{UCL} = \overline{\overline{X}} + \Phi^{-1}(1 - \alpha/2)(\hat{\sigma}/\min(i, w))\sqrt{(1/n_i) + (1/n_{i-1}) + \ldots + (1/n_{1+\max(i-w,0)})}$

These formulas 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 replace $\hat{\sigma}$ with σ_0 in Table 21.19. Note that the limits vary with both n_i and i.

If the subgroup sample sizes are constant ($n_i = n$), the formulas for the control limits simplify to

$$\text{LCL} = \overline{\overline{X}} - \frac{k\hat{\sigma}}{\sqrt{n \min(i, w)}}$$

$$\text{UCL} = \overline{\overline{X}} + \frac{k\hat{\sigma}}{\sqrt{n \min(i, w)}}$$

Refer to Montgomery (1996) for more details. When the subgroup sample sizes are constant, the width of the control limits for the first w moving averages decreases monotonically because each of the first w moving averages includes one more term than the preceding moving average.

If you specify the ASYMPTOTIC option, constant control limits with the following values are displayed:

$$\text{LCL} = \overline{\overline{X}} - \frac{k\hat{\sigma}}{\sqrt{nw}}$$

$$\text{UCL} = \overline{\overline{X}} + \frac{k\hat{\sigma}}{\sqrt{nw}}$$

For asymptotic probability limits, replace k with $\Phi^{-1}(1 - \alpha/2)$ in these equations. You can display asymptotic limits by specifying the ASYMPTOTIC option.

You can specify parameters for the moving average 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 w with the SPAN= option or with the variable _SPAN_ 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.

Choosing the Span of the Moving Average

There are few published guidelines for choosing the span w. In some applications, practical experience may dictate the choice of w. A more systematic approach is to choose w by considering its effect on the average run length (the expected number of points plotted before a shift is detected). This effect was studied by Roberts (1959), who used simulation methods.

You can use Table 21.20 and Table 21.21 to find a combination of k and w that yields a desired ARL for an in-control process ($\delta = 0$) and for a specified shift of δ.

Table 21.20. Average Run Lengths for One-Sided Uniformly Weighted Moving Average Charts

k	δ	w (span) 2	3	4	5	6	8	10
2.0	0.00	51.58	60.97	70.58	80.18	89.78	108.65	127.47
2.0	0.25	25.01	26.47	28.00	29.33	30.76	33.08	35.18
2.0	0.50	13.41	13.31	13.40	13.69	14.01	14.66	15.17
2.0	0.75	8.00	7.75	7.78	7.97	8.15	8.60	9.06
2.0	1.00	5.27	5.20	5.29	5.45	5.67	6.15	6.69
2.0	1.50	2.90	3.03	3.24	3.50	3.73	4.23	4.66
2.0	2.00	2.04	2.27	2.51	2.73	2.95	3.32	3.65
2.0	2.50	1.68	1.91	2.11	2.31	2.48	2.78	3.04
2.0	3.00	1.46	1.68	1.85	2.01	2.16	2.40	2.63
2.0	4.00	1.20	1.38	1.52	1.64	1.75	1.94	2.10
2.0	5.00	1.06	1.18	1.31	1.41	1.50	1.65	1.79
2.5	0.00	179.92	204.43	230.32	259.32	287.08	339.71	394.43
2.5	0.25	72.62	71.56	72.48	72.93	73.40	75.54	77.47
2.5	0.50	33.67	30.13	28.54	27.49	26.93	26.29	26.03
2.5	0.75	17.28	15.01	13.91	13.42	13.13	13.00	13.10
2.5	1.00	9.94	8.66	8.20	8.01	7.96	8.24	8.63
2.5	1.50	4.43	4.13	4.21	4.39	4.64	5.17	5.69
2.5	2.00	2.65	2.77	3.03	3.29	3.54	4.01	4.43
2.5	2.50	1.98	2.24	2.50	2.74	2.95	3.32	3.67
2.5	3.00	1.70	1.95	2.17	2.37	2.55	2.86	3.14
2.5	4.00	1.37	1.59	1.76	1.90	2.03	2.28	2.49
2.5	5.00	1.15	1.35	1.51	1.62	1.73	1.92	2.08

689

Table 21.20. (continued)

k	δ	2	3	4	5	6	8	10
3.0	0.00	792.24	867.57	963.95	1051.77	1150.79	1345.96	1539.75
3.0	0.25	269.28	244.26	231.50	226.25	220.89	209.87	204.74
3.0	0.50	104.18	83.86	72.84	65.43	60.85	54.62	50.34
3.0	0.75	45.69	34.45	28.79	25.69	23.66	21.24	20.15
3.0	1.00	22.73	16.74	14.20	12.89	12.12	11.52	11.45
3.0	1.50	7.65	6.16	5.70	5.64	5.75	6.23	6.78
3.0	2.00	3.77	3.49	3.63	3.89	4.17	4.71	5.20
3.0	2.50	2.46	2.63	2.90	3.18	3.43	3.88	4.28
3.0	3.00	1.96	2.23	2.50	2.74	2.95	3.33	3.65
3.0	4.00	1.57	1.81	2.00	2.18	2.34	2.62	2.87
3.0	5.00	1.30	1.55	1.72	1.85	1.97	2.20	2.40
3.5	0.00	4275.15	4536.99	4853.63	5168.75	5485.97	6088.03	6613.01
3.5	0.25	1281.12	1078.59	964.86	886.26	830.03	751.66	684.98
3.5	0.50	413.30	294.47	235.00	197.27	169.50	136.01	115.48
3.5	0.75	153.50	98.31	73.49	59.29	50.49	40.45	34.53
3.5	1.00	63.68	39.34	29.37	24.06	20.88	17.70	16.12
3.5	1.50	15.84	10.44	8.50	7.78	7.47	7.51	7.97
3.5	2.00	6.06	4.73	4.49	4.61	4.86	5.43	6.01
3.5	2.50	3.27	3.13	3.34	3.63	3.92	4.45	4.91
3.5	3.00	2.31	2.54	2.83	3.11	3.36	3.80	4.19
3.5	4.00	1.77	2.02	2.25	2.45	2.64	2.97	3.27
3.5	5.00	1.48	1.74	1.91	2.06	2.21	2.48	2.71

Table 21.21. Average Run Lengths for Two-Sided Uniformly Weighted Moving Average Charts

		w (span)						
k	δ	2	3	4	5	6	8	10
2.0	0.00	25.46	29.62	33.94	38.08	42.35	51.20	59.48
2.0	0.25	20.43	22.38	24.21	25.87	27.35	30.08	32.33
2.0	0.50	12.73	12.80	13.02	13.29	13.57	14.19	14.84
2.0	0.75	7.87	7.68	7.71	7.86	8.03	8.44	8.90
2.0	1.00	5.24	5.14	5.22	5.40	5.59	6.09	6.60
2.0	1.50	2.90	3.02	3.24	3.48	3.71	4.19	4.63
2.0	2.00	2.04	2.26	2.51	2.73	2.94	3.31	3.63
2.0	2.50	1.67	1.91	2.12	2.30	2.47	2.77	3.03
2.0	3.00	1.46	1.67	1.85	2.01	2.15	2.40	2.63
2.0	4.00	1.20	1.38	1.52	1.64	1.75	1.94	2.10
2.0	5.00	1.06	1.19	1.31	1.41	1.50	1.65	1.79

Table 21.21. (continued)

k	δ	2	3	4	5	6	8	10
2.5	0.00	89.48	101.24	114.35	127.74	140.88	166.98	192.93
2.5	0.25	63.12	64.91	67.00	68.75	69.84	72.22	74.49
2.5	0.50	32.46	29.54	28.20	27.33	26.72	25.92	25.72
2.5	0.75	17.28	14.97	13.85	13.29	13.02	12.81	12.98
2.5	1.00	9.94	8.61	8.16	7.99	8.01	8.23	8.63
2.5	1.50	4.42	4.14	4.20	4.38	4.62	5.16	5.67
2.5	2.00	2.65	2.77	3.03	3.29	3.54	4.00	4.43
2.5	2.50	1.99	2.24	2.50	2.73	2.95	3.33	3.65
2.5	3.00	1.69	1.95	2.17	2.37	2.54	2.86	3.14
2.5	4.00	1.37	1.59	1.76	1.90	2.04	2.27	2.49
2.5	5.00	1.15	1.35	1.51	1.63	1.73	1.92	2.09
3.0	0.00	397.12	436.27	481.16	527.14	574.05	667.68	762.89
3.0	0.25	245.51	228.67	222.75	216.07	213.79	207.03	201.71
3.0	0.50	103.15	83.49	72.47	65.67	60.67	53.93	50.30
3.0	0.75	45.56	34.25	29.01	25.72	23.59	21.12	19.93
3.0	1.00	22.68	16.81	14.19	12.92	12.18	11.54	11.48
3.0	1.50	7.68	6.14	5.71	5.65	5.77	6.23	6.77
3.0	2.00	3.74	3.49	3.63	3.88	4.17	4.71	5.21
3.0	2.50	2.46	2.63	2.90	3.18	3.43	3.89	4.29
3.0	3.00	1.96	2.23	2.50	2.73	2.95	3.32	3.66
3.0	4.00	1.57	1.81	2.00	2.18	2.34	2.62	2.88
3.0	5.00	1.30	1.55	1.72	1.85	1.97	2.20	2.40
3.5	0.00	2217.61	2372.09	2567.27	2775.06	2983.70	3398.08	3810.50
3.5	0.25	1186.27	1027.67	940.30	875.91	826.53	744.59	676.61
3.5	0.50	411.69	295.62	232.68	195.65	169.21	135.73	116.06
3.5	0.75	152.52	97.33	72.30	58.98	50.59	40.22	34.71
3.5	1.00	64.03	39.46	29.18	24.08	20.80	17.54	16.16
3.5	1.50	15.83	10.36	8.47	7.73	7.46	7.56	8.00
3.5	2.00	6.05	4.71	4.49	4.61	4.85	5.44	6.00
3.5	2.50	3.27	3.12	3.34	3.64	3.92	4.44	4.91
3.5	3.00	2.32	2.54	2.83	3.11	3.36	3.80	4.19
3.5	4.00	1.77	2.02	2.25	2.46	2.65	2.97	3.26
3.5	5.00	1.49	1.74	1.91	2.06	2.21	2.48	2.71

For example, suppose you want to construct a two-sided moving average chart with an in-control ARL of 100 and an ARL of 9 for detecting a shift of $\delta = 1$. Table 21.21 shows that the combination $w = 3$ and $k = 2.5$ yields an in-control ARL of 101.24 and an ARL of 8.61 for $\delta = 1$.

Note that you can also use Table 21.20 and Table 21.21 to evaluate an existing moving average chart (see Example 21.2 on page 703).

The following SAS program computes the average run length for a two-sided moving average chart for various shifts in the mean. This program can be adapted to compute averages run lengths for various combinations of k and w.

```
data sim;
   drop span delta time j y x;
   span=4;
   do shift=0,.25,.5,.75,1,1.5,2,2.5,3,4,5;
      do j=1 to 50000;
         do time=1 to 15000;
            if time<=100 then
               delta=0;
            else
               delta=shift;
            y=delta+rannor(234);
            if time<span then
               x=.;
            else
               x=(y+lag1(y)+lag2(y)+lag3(y))/span;
            if time>=101 and abs(x)>3/sqrt(span)
              then leave;
         end;
         arl=time-100;
         output;
         end;
      end;

proc means;
   class shift;
run;
```

In the preceding program, the size of the span w (SPAN) is 4 and the shifts in the mean are introduced to the variable (Y) $y \sim N(0, 1)$ after the first 100 observations. The first DO loop specifies shifts of various magnitude, the second DO loop performs 50000 simulations for each shift, and the third DO loop counts the run length (TIME), that is, the number of samples observed before the control chart signals. A large upper bound (15000) for TIME is specified so that the run length is uncensored.

The program can be generalized for various span sizes by assigning a different value for the variable SPAN and changing the expression for X appropriately. Optionally, you can compute the ARL for a one-sided chart by changing the limits, that is, `x>3/sqrt(span)`. This was the technique used to construct Table 21.20 and Table 21.21.

Output Data Sets

OUTLIMITS= Data Set

The OUTLIMITS= data set saves the control limit parameters. The following variables can be saved:

Variable	Description
ALPHA	probability (α) of exceeding limits
INDEX	optional identifier for the control limits specified with the OUTINDEX= option
LIMITN	sample size associated with the control limits
MEAN	process mean ($\overline{\overline{X}}$ or μ_0)
SIGMAS	multiple (k) of standard error of A_i
SPAN	number of terms in the moving average
STDDEV	process standard deviation ($\hat{\sigma}$ or σ_0)
SUBGRP	*subgroup-variable* specified in the MACHART statement
TYPE	type (estimate or standard value) of _MEAN_ and _STDDEV_
VAR	*process* specified in the MACHART statement

The OUTLIMITS= data set does not contain the control limits; instead, it contains control limit parameters that can be used to recompute the control limits.

Notes:

1. If the control limits vary with subgroup sample size, the special missing value *V* is assigned to the variable _LIMITN_.

2. If the limits are defined in terms of a multiple k of the standard error of A_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. Optional BY variables are saved in the OUTLIMITS= data set.

The OUTLIMITS= data set contains one observation for each *process* specified in the MACHART statement.

You can use OUTLIMITS= data sets

- to keep a permanent record of the control limit parameters

- to write reports. You may prefer to use OUTTABLE= data sets for this purpose.

- as LIMITS= data sets in subsequent runs of PROC MACONTROL

For an example of an OUTLIMITS= data set, see "Saving Control Limit Parameters" on page 670.

OUTHISTORY= Data Set

The OUTHISTORY= data set saves subgroup summary statistics. The following variables can be 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 moving average variable named by *process* suffixed with *A*
- 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 MACHART statement. For example, consider the following statements:

```
proc macontrol data=clips;
   machart (gap yldstren)*day / span      =3
                                outhistory=cliphist;
run;
```

The data set CLIPHIST would contain nine variables named DAY, GAPX, GAPS, GAPA, GAPN, YLDSRENX, YLDSRENS, YLDSRENA, and YLDSRENN.

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 669.

OUTTABLE= Data Set

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

Variable	Description
ALPHA	probability (α) of exceeding control limits
EXLIM	control limit exceeded on moving average chart
LCLA	lower control limit for moving average
LIMITN	nominal sample size associated with the control limits
MEAN	process mean
SIGMAS	multiple (k) of the standard error associated with control limits
SPAN	number of terms in the moving average
subgroup	values of the subgroup variable
SUBN	subgroup sample size
SUBS	subgroup standard deviation
SUBX	subgroup mean
UCLA	upper control limit for moving average
UWMA	uniformly weighted moving average
VAR	*process* specified in MACHART statement

In addition, the following variables, if specified, are included:
- BY variables
- *block-variables*
- ID variables
- _PHASE_ (if the READPHASES= option is specified)
- *symbol-variable*

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 variables _VAR_ and _EXLIM_ are character variables of length 8. The variable _PHASE_ is a character variable of length 16. All other variables are numeric.

For an example of an OUTTABLE= data set, see "Saving Control Limit Parameters" on page 670.

ODS Tables

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

Table 21.22. ODS Tables Produced with the MACHART Statement

Table Name	Description	Options
MACHART	uniformly weighted moving average chart summary statistics	TABLE, TABLEALL, TABLEC, TABLEID, TABLEOUT
Parameters	uniformly weighted moving average parameters	TABLE, TABLEALL, TABLEC, TABLEID, TABLEOUT

Input Data Sets

DATA= Data Set

You can read raw data (process measurements) from a DATA= data set specified in the PROC MACONTROL statement. Each *process* specified in the MACHART 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 MACHART 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 ith 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 ith 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 MACONTROL procedure reads all of the observations in a DATA= data set. However, if the data set includes the variable _PHASE_, you can read selected groups of observations (referred to as *phases*) with the READPHASES= option (for an example, see "Displaying Stratification in Phases" on page 1689).

For an example of a DATA= data set, see "Creating Moving Average Charts from Raw Data" on page 664.

LIMITS= Data Set

You can read preestablished control limits parameters from a LIMITS= data set specified in the PROC MACONTROL statement. The LIMITS= data set used by the MACONTROL procedure does not contain the actual control limits, but rather it contains the parameters required to compute the limits. For example, the following statements read control limit parameters from the data set PARMS:*

```
proc macontrol data=parts limits=parms;
    machart gap*day;
run;
```

The LIMITS= data set can be an OUTLIMITS= data set that was created in a previous run of the MACONTROL procedure. Such data sets always contain the variables required for a LIMITS= data set; see page 693. The LIMITS= data set can also be created directly using a DATA step.

When you create a LIMITS= data set, you must provide the variable _SPAN_, which specifies the number of terms to use in the moving average. 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, STDMEAN, and STDSIGMA.

- BY variables are required if specified with a BY statement.

Some advantages of working with a LIMITS= data set are that

- it facilitates reusing a permanently saved set of parameters

- a distinct set of parameters can be read for each *process* specified in the MACHART statement

- it facilitates keeping track of multiple sets of parameters that accumulate for the same *process* as the process evolves over time

For an example, see "Reading Preestablished Control Limit Parameters" on page 672.

HISTORY= Data Set

You can read subgroup summary statistics from a HISTORY= data set specified in the PROC MACONTROL statement. This allows you to reuse OUTHISTORY= data sets that have been created in previous runs of the MACONTROL, SHEWHART,

---

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

or CUSUM procedures or to read output data sets created with SAS summarization procedures such as PROC MEANS.

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

- the *subgroup-variable*
- a subgroup mean variable for each *process*
- a subgroup sample size variable for each *process*
- a subgroup standard deviation 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 suffix characters X, S, and N, respectively.

For example, consider the following statements:

```
proc macontrol history=cliphist;
    machart (gap diameter)*day  / span=3;
run;
```

The data set CLIPHIST must include the variables DAY, GAPX, GAPS, GAPN, DIAMTERX, DIAMTERS, and DIAMTERN.

Although a moving average variable (named by the *process* name suffixed with A) is saved in an OUTHISTORY= data set, it is not required in a HISTORY= data set, because the subgroup mean variable is sufficient to compute the moving averages.

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

For an example of a HISTORY= data set, see "Creating Moving Average Charts from Subgroup Summary Data" on page 667.

TABLE= Data Set

You can read summary statistics and control limits from a TABLE= data set specified in the PROC MACONTROL statement. This enables you to reuse an OUTTABLE= data set created in a previous run of the MACONTROL procedure.

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

698

Variable	Description
LCLE	lower control limit for Moving Average
LIMITN	nominal sample size associated with the control limits
MEAN	process mean
SPAN	number of terms in the moving average
subgroup-variable	values of the *subgroup-variable*
SUBN	subgroup sample size
SUBS	subgroup standard deviation
SUBX	subgroup mean
UCLA	upper control limit for moving average
UWMA	uniformly weighted moving average

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.
- _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 Limit Parameters" on page 670.

Methods for Estimating the Standard Deviation

When control limits are computed from the input data, four methods are available for estimating the process standard deviation σ. Three methods (referred to as the default, MVLUE, and RMSDF) are available with subgrouped data. A fourth method is used if the data are individual measurements (see "Default Method for Individual Measurements" on page 700).

Default Method for Subgroup Samples

This method is the default for moving average charts using subgrouped data. The default estimate of σ is

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

where N is the number of subgroups for which $n_i \geq 2$, s_i is the sample standard deviation of the ith subgroup

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

and

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

699

Here $\Gamma(\cdot)$ denotes the gamma function, and $\bar{X}_i$ denotes the ith 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 for Subgroup Samples

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 $s_i/c_4(n_i)$, and it is computed as

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

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 for Subgroup Samples

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

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

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.

Default Method for Individual Measurements

When each subgroup sample contains a single observation ($n_i \equiv 1$), the process standard deviation σ is estimated as

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

where N is the number of observations, and $x_1, x_2, \ldots, x_N$ are the individual measurements. This formula is given by Wetherill (1977), who states that the estimate of the variance is biased if the measurements are autocorrelated.

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	DATA=	*process*
Vertical	HISTORY=	subgroup mean variable
Vertical	TABLE=	_UWMA_

For example, the following sets of statements specify the label *Moving Average of Clip Gaps* for the vertical axis and the label *Day* for the horizontal axis of the moving average chart:

```
proc macontrol data=clips1;
   machart gap*day / span=4;
   label gap = 'Moving Average of Clip Gaps';
   label day = 'Day';
run;

proc macontrol history=cliphist;
   machart gap*day / span=4;
   label gapx = 'Moving Average of Clip Gaps';
   label day  = 'Day';
run;

proc macontrol table=cliptab;
   machart gap*day;
   label _uwma_   = 'Moving Average of Clip Gaps';
   label day      = 'Day';
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 MACHART statement.

Example 21.1. Specifying Standard Values for the Process Mean and Process Standard Deviation

By default, the MACHART statement estimates the process mean (μ) and standard deviation (σ) from the data. This is illustrated in the "Getting Started" section of this chapter. 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 metal clip manufacturing process (introduced on page 664) has a mean of 15 and standard deviation of 0.2. The following statements specify these standard values:

```
title 'Specifying Standard Process Mean and Standard Deviation';
symbol v=dot;
proc macontrol data=clips1;
   machart gap*day / mu0    = 15
                     sigma0 = 0.2
                     span   = 4
                     xsymbol= mu0;
run;
```

The XSYMBOL= option specifies the label for the central line. The resulting chart is shown in Output 21.1.1.

Output 21.1.1. Specifying Standard Values with MU0= and SIGMA0=
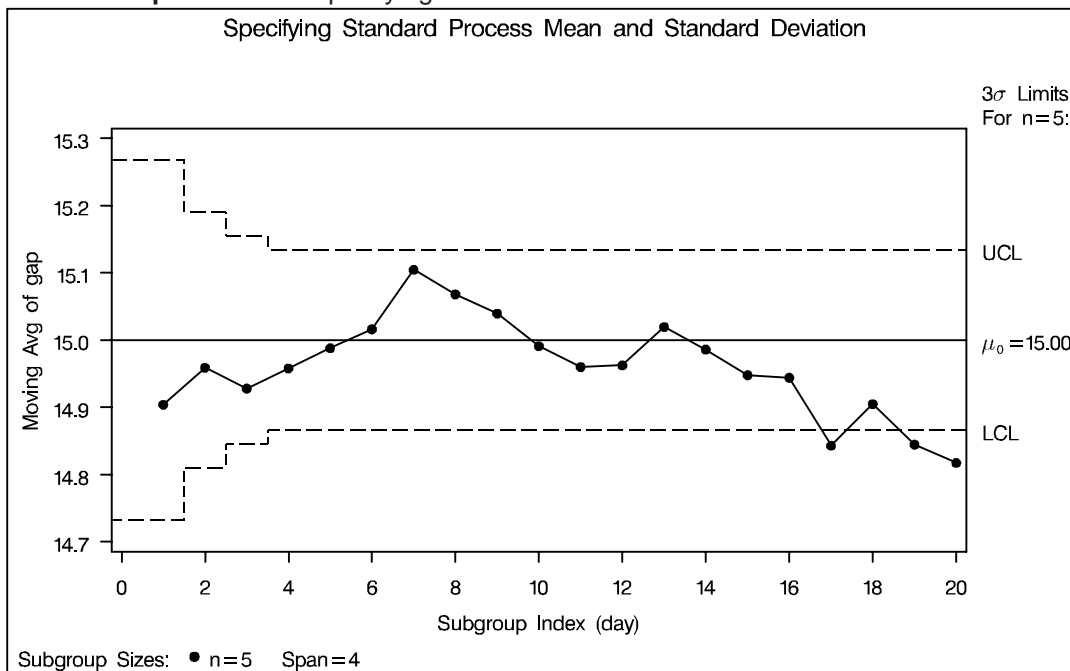
The central line and control limits are determined using μ_0 and σ_0 (see the equations in Table 21.19 on page 688). Output 21.1.1 indicates that the process is out-of-control since the moving averages for DAY=17, DAY=19, and DAY=20 lie below the lower control limit.

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 cliplim;
   length _var_ _subgrp_ _type_ $8;
   _var_    = 'gap';
   _subgrp_ = 'day';
   _type_   = 'STANDARD';
   _limitn_ = 5;
   _mean_   = 15;
   _stddev_ = 0.2;
   _span_   = 4;
run;

proc macontrol data=clips1 limits=cliplim;
   machart gap*day / xsymbol=mu0;
run;
```

The variable _SPAN_ is required, and its value provides the number of terms in the moving average. The variables _VAR_ and _SUBGRP_ are also required, and their values must match the *process* and *subgroup-variable*, respectively, specified in the MACHART 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 chart (not shown here) is identical to the one shown in Output 21.1.1.

Example 21.2. Annotating Average Run Lengths on the Chart

You can use Table 21.20 on page 689 and Table 21.21 on page 690 to find a moving average chart scheme with the desired average run length properties. Specifically, you can find a combination of k and w that yields a desired ARL for an in-control process ($\delta = 0$) and for a specified shift of δ.

See MACMA3 in the SAS/QC Sample Library

You can also use these tables to evaluate an existing moving average chart scheme. For example, the moving average chart shown in Output 21.1.1 has a two-sided scheme with $w = 4$ and $k = 3$. Suppose you want to detect a shift of $\delta = .5$. From Table 21.21, the average run length with $w = 4$, $k = 3$, and $\delta = .5$ is 72.47. The in-control average run length ($\delta = 0$) for this scheme is 481.16.

The following statements create an annotate data set that can be read to display these ARL values on the moving average chart:

```
data arlanno;
   length function style color $ 8 text $ 20;
   function = 'label';    hsys = '1';  xsys = '1'; ysys = '1';
                          size     =   5;
                          position = 'c';
                          style    = 'swiss';
                          color    = 'black';
                          x        =  78;
                          y        =  92;
                          text     = 'ARL Evaluation'; output;
                          y        =  84;
                          text     = 'ARL In';          output;
                          y        =  78;
                          text     = 'ARL Out';         output;
                          position = 'a';
                          x        =  97;
                          y        =  84;
                          text     = '481.16';          output;
                          y        =  78;
                          text     = ' 72.47' ;         output;
   function = 'poly';     style    = 'empty';
                          x = 98;  y = 98;              output;
   function = 'polycont'; x = 98;  y = 77;              output;
                          x = 77;  y = 77;              output;
                          x = 77;  y = 98;              output;
                          x = 98;  y = 98;              output;
   function = 'move';     x = 77;  y = 91;              output;
   function = 'draw';     size  = 1;
                          hsys  = '4';
                          x = 98; y = 91;              output;
run;
```

The following statements create the moving average chart shown in Output 21.2.1.

```
title 'Moving Average Chart to Detect a Shift of 0.5';
symbol v=dot h=.8;
proc macontrol data=clips1 annotate=arlanno;
   machart gap*day /
        mu0     = 15
        sigma0 = 0.2
        span    = 4
       xsymbol= mu0
        haxis  = axis1
        vaxis  = axis2;
axis1 offset=(2 pct, 2pct);
axis2 offset=(2 pct, 2pct);
run;
```

The average run lengths in this example (481.16 and 72.27) are simply copied from Table 21.21. You can generalize the preceding program so that it computes the average run lengths by incorporating the simulation program on page 692.

Output 21.2.1. Displaying Average Run Lengths on Chart
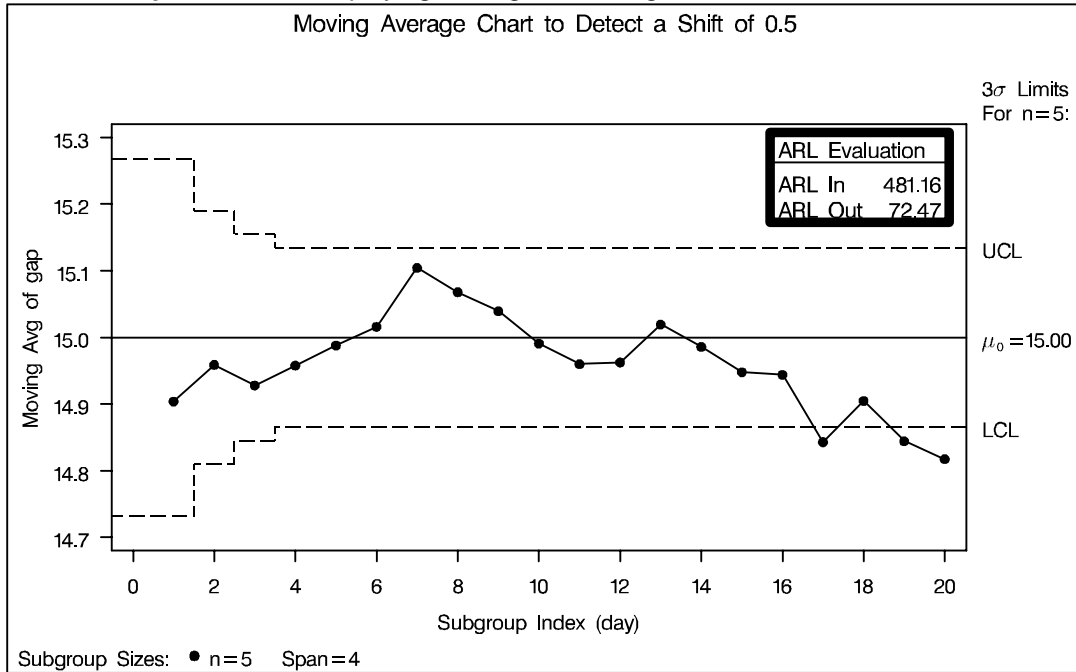
For more information on ANNOTATE= data sets, refer to *SAS/GRAPH Software: Reference*.

SAS/QC® User's Guide, Version 8