SAS/QC® 13.2 User's Guide

The ANOM Procedure

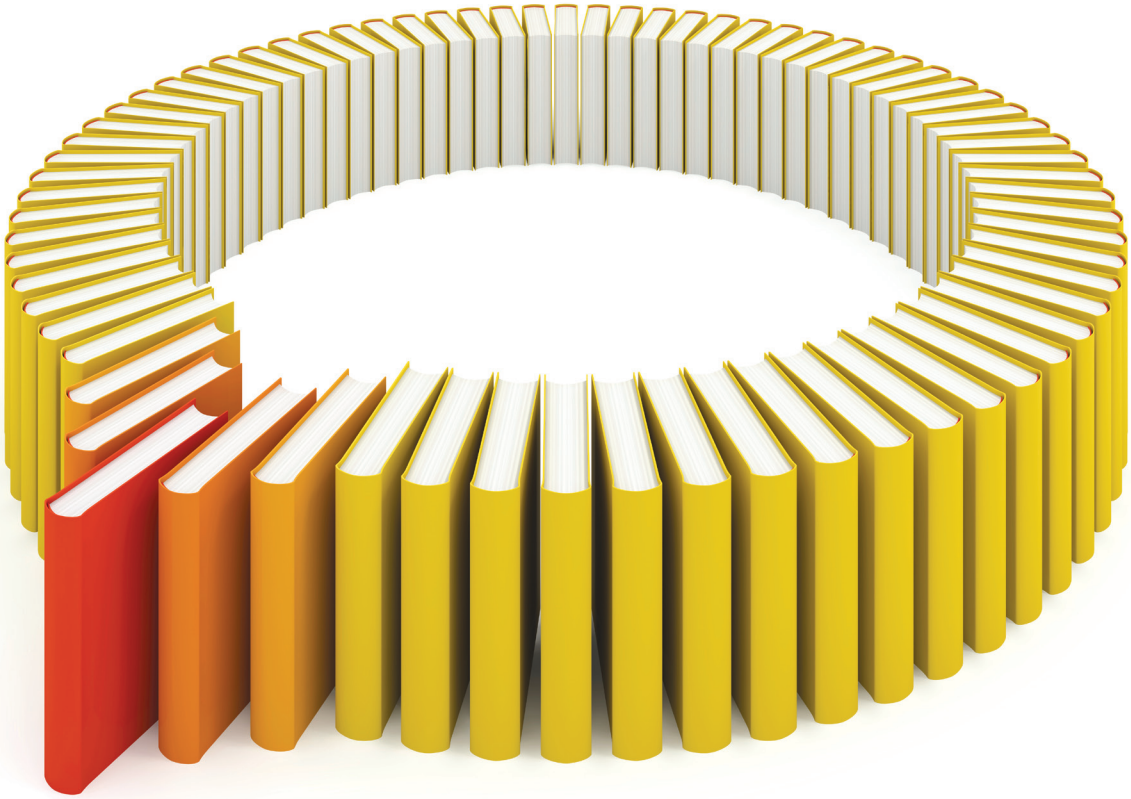
This document is an individual chapter from *SAS/QC® 13.2 User's Guide*.

The correct bibliographic citation for the complete manual is as follows: SAS Institute Inc. 2014. *SAS/QC® 13.2 User's Guide*. Cary, NC: SAS Institute Inc.

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

August 2014

SAS provides a complete selection of books and electronic products to help customers use SAS® software to its fullest potential. For more information about our offerings, visit **support.sas.com/bookstore** or call 1-800-727-3228.



Chapter 4
The ANOM Procedure

Contents

PROC ANOM and General Statements

Overview: ANOM Procedure

Analysis of means (ANOM) is a graphical and statistical method for simultaneously comparing k treatment means with their overall mean at a specified significance level α. You can use the ANOM procedure to create ANOM charts for various types of response data, including continuous measurements, proportions, and rates.

In addition, you can use the ANOM procedure to do the following:

- create charts from either response values or summarized data
- analyze multiple response variables
- specify decision limits in terms of the significance level (α)
- compute decision limits from the data and automatically adjust decision limits for unequal sample sizes
- save chart statistics and decision limits in output data sets
- tabulate chart statistics and decision limits.

See Chapter 3, "SAS/QC Graphics," for a detailed discussion of the alternatives available for producing charts with SAS/QC procedures.

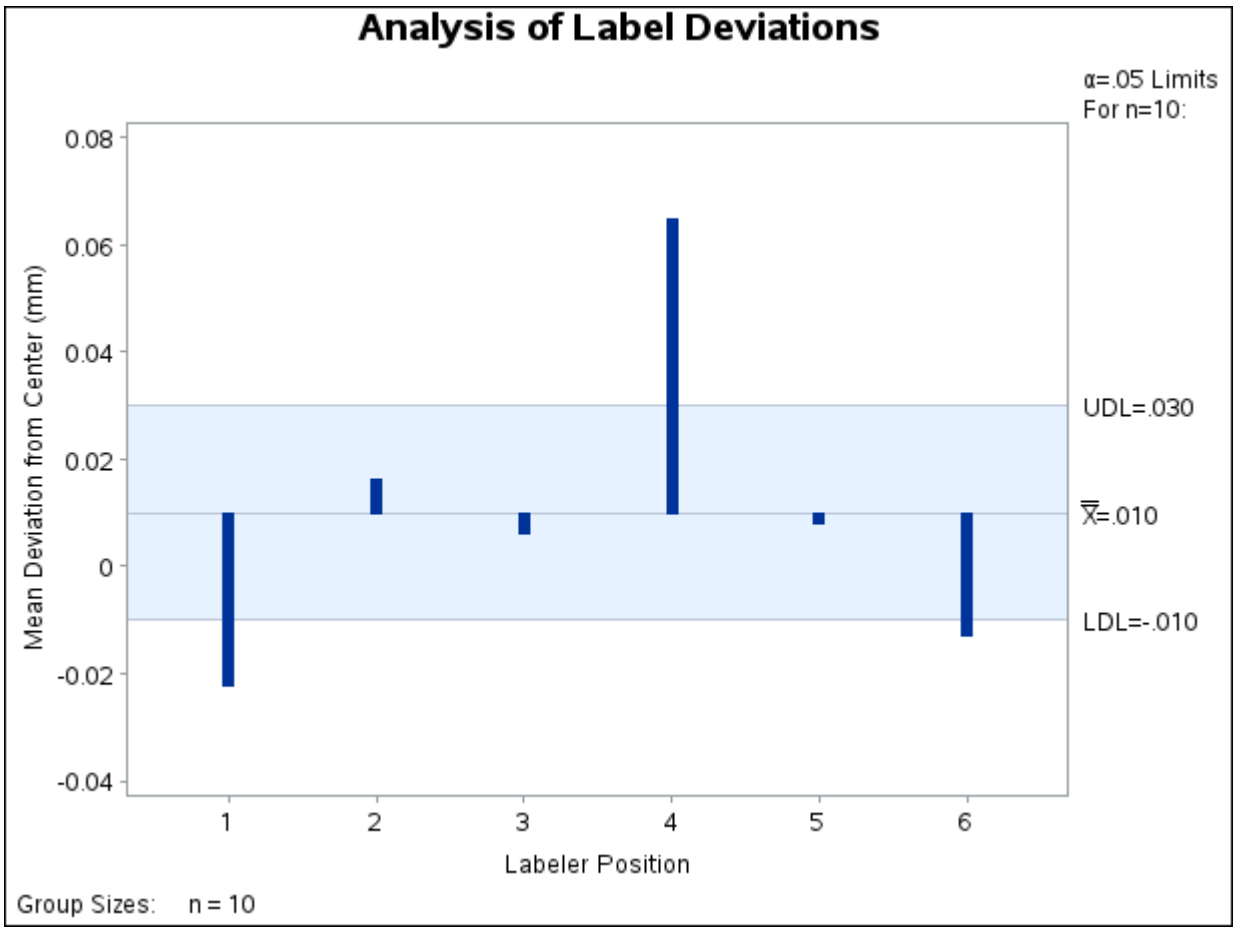
Uses of Analysis of Means

Many statistical quality improvement applications involve a comparison of treatment means to determine which are significantly different from the overall average. For example, a manufacturing engineer might run an experiment to investigate which of six positions on a machine are producing different output, in the sense that the average measurement for each position differs from the overall average. Likewise, a health care system administrator might ask which clinics in the system have a higher or lower rate of admissions than the average for all clinics.

Questions of this type can be answered with *analysis of means*, which is an alternative to one-way analysis of variance (ANOVA) for a fixed effects model. However, unlike ANOVA, which simply determines whether there is a statistically significant difference in the treatment means, ANOM identifies the means that are significantly different from the overall mean. As a statistical technique, ANOM is a method for making multiple comparisons that is sometimes referred to as a "multiple comparison with the weighted mean." Analysis of means lends itself to quality improvement applications because it has a simple graphical representation that is similar to a Shewhart chart and requires little training to interpret. This representation is also useful for assessing practical significance.

Figure 4.1 illustrates a typical ANOM chart. The central line represents the overall average. The treatment means, plotted as deviations from the overall average are compared with upper and lower decision limits to identify which are significantly different from the overall mean (in this case, the means corresponding to the first, fourth, and sixth positions).

Figure 4.1 Typical ANOM Chart



Although the term "analysis of means" suggests that the method is intended for means of continuous response measurements, the method is also applicable to means of attributes data, including proportions and rates.

Analysis of means was introduced as a tool for statistical quality control by Ellis Ott in 1967, and it became popular during the early 1980s, when it was applied to experimental data in manufacturing. In this setting, measurements are taken at a number of treatment levels (factor levels). During the 1990s, the use of ANOM spread to service industry applications and, in particular, health care quality improvement. In these settings, data (such as utilization rates) are observed for a number of groups (such as hospitals or clinics).

Terminology

In order to accommodate the growing variety of modern applications for analysis of means, the term *group* is used instead of treatment level throughout the documentation for the ANOM procedure. Likewise, the term *group-variable* is used to refer to the variable in the input data set that classifies the observations into treatment levels. In the ANOM procedure, a *group-variable* plays the same role as a CLASS variable in the GLM and ANOVA procedures, and it is syntactically the same as a *subgroup-variable* in the SHEWHART procedure.

The nomenclature for ANOM charts is the same as that for Shewhart charts: $\bar{X}$ charts for means, p charts for proportions, and u charts for rates. Consequently, the syntax for the ANOM procedure is patterned after

the syntax for the SHEWHART procedure. However, there are some important differences between ANOM charts and Shewhart charts:

- Analysis of means is formally a test of hypothesis, whereas a Shewhart chart is used to distinguish between special and common causes of variation.
- In an ANOM chart, the horizontal axis corresponds to the *group-variable*, and it identifies the groups, which can be displayed in any order. In a Shewhart chart, the horizontal axis corresponds to the *subgroup-variable*, and it identifies the order in which the subgroup measurements were taken.
- An ANOM chart displays response summary statistics for a set of groups (treatments) at a specific time. A Shewhart chart displays subgroup summary statistics for a specific process where the subgroups are made up of measurements taken over successive points in time.
- In an ANOM chart, the decision limits are determined by a specified significance level (α), which is the probability that under the null hypothesis of no treatment differences, at least one of the response summary statistics will exceed the decision limits. In a Shewhart chart, control limits are typically computed as 3σ limits.

History

Analysis of means compares the absolute deviations of group means from their overall mean, an approach that was initially studied by Laplace in 1827. Halperin et al. derived a version of this method in the form of a multiple significance test in 1955. Ott developed a graphical representation for the test and introduced the term "analysis of means" in 1967. Refer to Ott (1967) and Ott (1975).

P. R. Nelson (1982a) and L. S. Nelson (1983) provided exact critical values for ANOM when the groups have equal sample sizes. P. R. Nelson (1991) developed a method for computing exact critical values for ANOM when the group sample sizes are not equal. Refer to Nelson, Coffin, and Copeland (2003) for more information on the use of ANOM in engineering experimentation.

Using the ANOM Procedure

The PROC ANOM statement invokes the ANOM procedure and it optionally identifies various data sets.

To create an ANOM chart, you specify a chart statement (after the PROC ANOM statement) that specifies the type of ANOM chart you want to create and the variables in the input data set that you want to analyze. For example, the following statements request a basic ANOM chart for treatment means:

```
proc anom data=Values;
   xchart Weight*Treatment;
run;
```

Here, the DATA= option specifies an input data set (Values) that contains the *response* measurement variable (Weight) and the *group-variable* (Treatment). You can use options in the PROC ANOM statement to

- specify input data sets containing variables to be analyzed, decision limits, and annotation information
- specify a graphics catalog for saving graphical output

NOTE: If you are learning to use the ANOM procedure, you should read both this section and the "Getting Started" subsection in the section for the chart statement that corresponds to the chart you want to create.

Syntax: ANOM Procedure

The following are the primary statements that control the ANOM procedure:

```
PROC ANOM <options> ;
BOXCHART (responses) * group-variable <(block-variables)>
          <=symbol-variable> </ options> ;
PCHART (responses) * group-variable
        <(block-variables)>
        <=symbol-variable> </ options> ;
UCHART (responses) * group-variable
        <(block-variables)>
        <=symbol-variable> </ options> ;
XCHART (responses) * group-variable <(block-variables)>
        <=symbol-variable> </ options> ;
INSET keyword-list </ options> ;
```

The PROC ANOM statement invokes the procedure and specifies the input data set. The chart statements create different types of charts. You can specify one or more of each of the chart statements. For details, read the section on the chart statement that corresponds to the type of chart that you want to produce.

BY Statement

```
BY variables ;
```

You can specify a BY statement with PROC ANOM to obtain separate analyses of observations in groups that are defined by the BY variables. When a BY statement appears, the procedure expects the input data set to be sorted in order of the BY variables. If you specify more than one BY statement, only the last one specified is used.

If your input data set is not sorted in ascending order, use one of the following alternatives:

- Sort the data by using the SORT procedure with a similar BY statement.
- Specify the NOTSORTED or DESCENDING option in the BY statement for the ANOM procedure. The NOTSORTED option does not mean that the data are unsorted but rather that the data are arranged in groups (according to values of the BY variables) and that these groups are not necessarily in alphabetical or increasing numeric order.
- Create an index on the BY variables by using the DATASETS procedure (in Base SAS software).

For more information about BY-group processing, see the discussion in *SAS Language Reference: Concepts*. For more information about the DATASETS procedure, see the discussion in the *Base SAS Procedures Guide*.

ID Statement

In addition, you can optionally specify the following statement:

```
ID variables ;
```

The ID statement specifies variables used to identify observations. The ID variables must be variables in the DATA= or SUMMARY= input data sets.

The ID variables are used in the following ways:

- If you create an OUTSUMMARY= or OUTTABLE= data set, the ID variables are included. If the input data set is a DATA= data set, only the values of the ID variables from the first observation in each group are passed to the output data set.
- If you specify the TABLEID or TABLEALL options in a chart statement, the table produced is augmented by a column for each of the ID variables. Only the values of the ID variables from the first observation in each group are tabulated.
- If you specify the BOXSTYLE= SCHEMATICID option or the BOXSTYLE= SCHEMATICIDFAR option in the BOXCHART statement, the value of the first variable listed in the ID statement is used to label each extreme observation.

Graphical Enhancement Statements

You can use TITLE, FOOTNOTE, and NOTE statements to enhance graphical and printed output. You can also use AXIS, LEGEND, and SYMBOL statements to enhance traditional graphics. For details, refer to *SAS/GRAPH: Reference* and see the section for the chart statement that you are using.

PROC ANOM Statement

The syntax for the PROC ANOM statement is as follows:

PROC ANOM *options* **;**

The PROC ANOM statement starts the ANOM procedure and optionally identifies various data sets. The following options can appear in the PROC ANOM statement.

ANNOTATE=*SAS-data-set*

ANNO=*SAS-data-set*

specifies an input data set containing ANNOTATE= variables as described in *SAS/GRAPH: Reference*. You can use this data set to add features to ANOM charts produced as traditional graphics. Features provided in this data set are displayed on every chart produced in the current run of the ANOM procedure. This option is ignored if you are not producing traditional graphics.

BOX=*SAS-data-set*

names an input data set that contains group summary statistics, decision limits, and outlier values in "strung out" form, with more than one observation per group. Each observation corresponds to one feature of one group's box-and-whisker plot. Typically, this data set is created as an OUTBOX= data set in a previous run of the ANOM procedure with a BOXCHART statement. The BOX= data set is the only kind of summary data set you can use to produce schematic box-and-whisker plots. The BOXCHART statement is the only chart statement you can use with a BOX= input data set.

DATA=*SAS-data-set*

names an input data set that contains response values (typically, measurements or counts) as observations. Note that the DATA= data set may need sorting. If the values of the *group-variable* are numeric, you must sort the data set so that these values are in increasing order (within BY groups). Use PROC SORT if the data are not already sorted.

The DATA= data set may contain more than one observation for each value of the *group-variable*. This happens, for example, when you produce a chart for means and ranges with the XCHART statement.

You cannot use a DATA= data set together with a SUMMARY= or a TABLE= data set. If you do not specify one of these three input data sets, the ANOM procedure uses the most recently created data set as a DATA= data set. For more information, see the "DATA= Data Set" subsection in the section for the chart statement you are using.

GOUT=*graphics-catalog*

specifies the graphics catalog for traditional graphics output from the ANOM procedure. This is useful if you want to save the output. This option is ignored if you are not producing traditional graphics.

SUMMARY=*SAS-data-set*

names an input data set that contains group summary statistics. For example, you can read sample sizes, means, and standard deviations for the groups to create an ANOM chart. Typically, this data set is created as an OUTSUMMARY= data set in a previous run of the ANOM procedure, but it can also be created using a SAS summarization procedure such as PROC MEANS.

Note that the SUMMARY= data sets may need sorting. If the values of the *group-variable* are numeric, you need to sort the data set so that these values are in increasing order (within BY groups). Use PROC SORT if the data are not already sorted. The SUMMARY= data set can contain only one observation for each value for the *group-variable*.

You cannot use a SUMMARY= data set with a DATA= or a TABLE= data set. If you do not specify one of these three input data sets, the ANOM procedure uses the most recently created data set as a DATA= data set. For more information, see the "SUMMARY= Data Set" subsection in the section for the chart statement you are using.

LIMITS=*SAS-data-set*

names an input data set that contains preestablished decision limits or the parameters from which decision limits can be computed. Each observation in a LIMITS= data set provides decision limit information for a *response*. Typically, this data set is created as an OUTLIMITS= data set in a previous run of the ANOM procedure.

If you omit the LIMITS= option, then decision limits are computed from the data in the DATA= or SUMMARY= input data sets. For details about the variables needed in a LIMITS= data set, see the "LIMITS= Data Set" subsection in the section for the chart statement you are using.

TABLE=*SAS-data-set*

names an input data set that contains group summary statistics and decision limits. Each observation in a TABLE= data set provides information for a particular group and *response*. Typically, this data set is created as an OUTTABLE= data set in a previous run of the ANOM procedure.

You cannot use a TABLE= data set with a DATA= or a SUMMARY= data set. If you do not specify one of these three input data sets, the ANOM procedure uses the most recently created data set as a DATA= data set. For more information, see the "TABLE= Data Set" subsection in the section for the chart statement that you are using.

BOXCHART Statement: ANOM Procedure

Overview: BOXCHART Statement

The BOXCHART statement creates an ANOM chart for group (treatment level) means of response values superimposed with box-and-whisker plots of the measurements in each group. Throughout this chapter, a chart of this type is referred to as an *ANOM boxchart*. You can use options in the BOXCHART statement to

- compute decision limits from the data based on a specified parameters, such as the significance level (α)
- tabulate group sample sizes, group means, decision limits, and other information
- save decision limits in an output data set
- save group sample sizes and group means in an output data set
- read decision limits and decision limit parameters from a data set
- display distinct sets of decision limits for different sets of groups
- specify one of several methods for calculating quantile statistics (percentiles)
- control the style of the box-and-whisker plots
- add block legends and symbol markers to identify special groups
- 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

You have two alternatives for producing ANOM boxcharts with the BOXCHART statement:

- ODS Graphics output is produced if ODS Graphics is enabled, for example by specifying the ODS GRAPHICS ON statement prior to the PROC statement.
- Otherwise, traditional graphics are produced if SAS/GRAPH® is licensed.

See Chapter 3, “SAS/QC Graphics,” for more information about producing these different kinds of graphs.

Getting Started: BOXCHART Statement

This section introduces the BOXCHART statement with simple examples that illustrate the most commonly used options. Complete syntax for the BOXCHART statement is presented in the section "Syntax: BOXCHART Statement" on page 53, and advanced examples are given in the section "Examples: BOXCHART Statement" on page 74.

Creating ANOM Boxcharts from Response Values

NOTE: See *Creating ANOM BOXCHARTS from Response Values* in the SAS/QC Sample Library.

A manufacturing engineer carries out a study to determine the source of excessive variation in the positioning of labels on shampoo bottles.[1] A labeling machine removes bottles from the line, attaches the labels, and returns the bottles to the line. There are six positions on the machine, and the engineer suspects that one or more of the position heads might be faulty.

A sample of 60 bottles, 10 per position, is run through the machine. For each bottle, the deviation of each label is measured in millimeters, and the machine position is recorded. The following statements create a SAS data set named LabelDeviations, which contains the deviation measurements for the 60 bottles:

```
data LabelDeviations;
   input Position @;
   do i = 1 to 5;
      input Deviation @;
      output;
   end;
   drop i;
   datalines;
1  -0.02386  -0.02853  -0.03001  -0.00428  -0.03623
1  -0.04222  -0.00144  -0.06466   0.00944  -0.00163
2  -0.02014  -0.02725   0.02268  -0.03323   0.03661
2   0.04378   0.05562   0.00977   0.05641   0.01816
3  -0.00728   0.02849  -0.04404  -0.02214  -0.01394
3   0.04855   0.03566   0.02345   0.01339  -0.00203
4   0.06694   0.10729   0.05974   0.06089   0.07551
4   0.03620   0.05614   0.08985   0.04175   0.05298
5   0.03677   0.00361   0.03736   0.01164  -0.00741
5   0.02495  -0.00803   0.03021  -0.00149  -0.04640
6   0.00493  -0.03839  -0.02037  -0.00487  -0.01202
6   0.00710  -0.03075   0.00167  -0.02845  -0.00697
;
```

A partial listing of LabelDeviations is shown in Figure 4.2.

[1] This example is based on a case study described by Hansen (1990).

Figure 4.2 Listing of the Data Set LabelDeviations

The Data Set LabelDeviations

Position	Deviation
1	-0.02386
1	-0.02853
1	-0.03001
1	-0.00428
1	-0.03623
1	-0.04222
1	-0.00144
1	-0.06466
1	0.00944
1	-0.00163
2	-0.02014
2	-0.02725

The data set LabelDeviations is said to be in "strung-out" form, because each observation contains the position and the deviation measurement for a single bottle. The first 10 observations contain the measurements for the first position, the second 10 observations contain the measurements for the second position, and so on. Because the variable Position classifies the observations into groups (treatment levels), it is referred to as the *group-variable*. The variable Deviation contains the deviation measurements and is referred to as the *response variable* (or *response* for short).

The following statements create the ANOM boxchart shown in Figure 4.3:

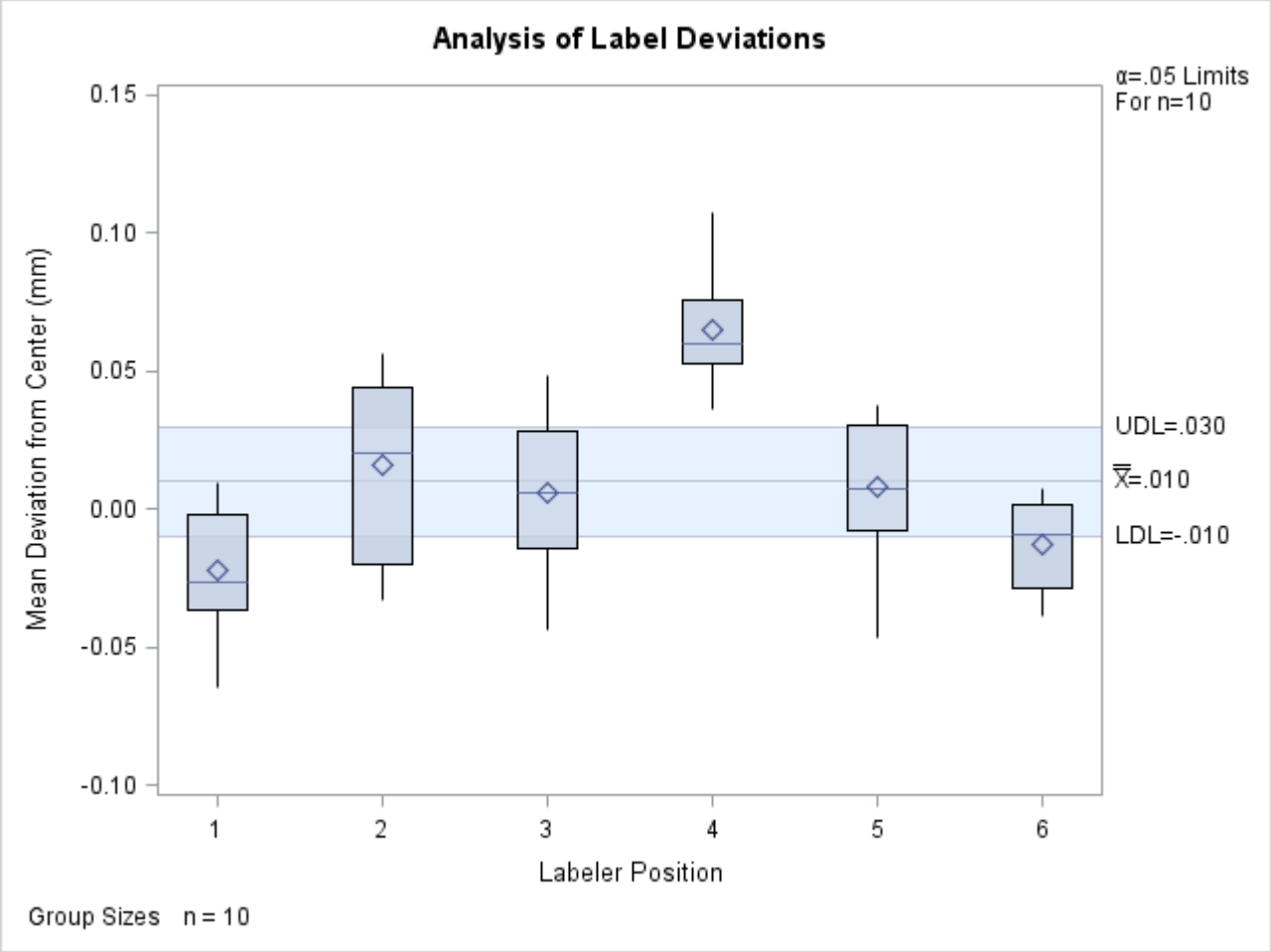
```
ods graphics on;
title 'Analysis of Label Deviations';
proc anom data=LabelDeviations;
   boxchart Deviation*Position / alpha    = 0.05
                                 odstitle = title;
   label Deviation = 'Mean Deviation from Center (mm)';
   label Position  = 'Labeler Position';
run;
```

The ODS GRAPHICS ON statement specified before the PROC ANOM statement enables ODS Graphics, so the boxchart is created by using ODS Graphics instead of traditional graphics. This example illustrates the basic form of the BOXCHART statement. After the keyword BOXCHART, you specify the *response* to analyze (in this case, Deviation) followed by an asterisk and the *group-variable* (Position). Options are specified after the slash (/) in the BOXCHART statement. A complete list of options is presented in the section "Syntax: BOXCHART Statement" on page 53.

The input data set is specified with the DATA= option in the PROC ANOM statement when it contains raw measurements for the *response*.

Each point on the ANOM chart represents the average (mean) of the response measurements for a particular sample.

Figure 4.3 ANOM Chart for Means of Labeler Position Data



The average for Position 1 is below the lower decision limit (LDL), and the average for Position 6 is slightly below the lower decision limit. The average for Position 4 exceeds the upper decision limit (UDL). The conclusion is that Positions 1, 4, and 6 are operating differently.

By default, the decision limits shown correspond to a significance level of $\alpha = 0.05$; the formulas for the limits are given in the section "Decision Limits" on page 63. You can also read decision limits from an input data set.

For computational details, see "Constructing ANOM Boxcharts" on page 62. For details on reading raw measurements, see "DATA= Data Set" on page 70.

Creating ANOM Boxcharts from Group Summary Data

NOTE: See *Creating BOXCHARTS from Group Summary Data* in the SAS/QC Sample Library.

The previous example illustrates how you can create ANOM charts for means using measurement data. However, in many applications, the data are provided as group summary statistics. This example illustrates how you can use the BOXCHART statement with data of this type.

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

```
data Labels;
   input Position DeviationL Deviation1 DeviationX
         DeviationM Deviation3 DeviationH DeviationS;
   DeviationN = 10;
   datalines;
1  -0.0647  -0.0362  -0.02234  -0.02620  -0.0016  0.0094  0.02281
2  -0.0332  -0.0201   0.01625   0.02045   0.0438  0.0564  0.03347
3  -0.0440  -0.0139   0.00604   0.00570   0.0285  0.0486  0.02885
4   0.0362   0.0530   0.06473   0.06030   0.0755  0.1073  0.02150
5  -0.0464  -0.0074   0.00813   0.00760   0.0302  0.0374  0.02593
6  -0.0384  -0.0285  -0.01283  -0.00950   0.0017  0.0071  0.01599
;
```

A listing of Labels is shown in Figure 4.4. There is exactly one observation for each group (note that the groups are still indexed by Position). There are eight summary variables in Labels.

- DeviationL contains the group minimums (low values).
- Deviation1 contains the 25th percentile (first quartile) of each group.
- DeviationX contains the group means.
- DeviationM contains the group medians.
- Deviation3 contains the 75th percentile (third quartile) of each group.
- DeviationH contains the group maximums (high values).
- DeviationS contains the group standard deviations.
- DeviationN contains the group sample sizes (these are all 10 in this case).

Figure 4.4 The Summary Data Set Labels

The Data Set Labels

Position	DeviationL	Deviation1	DeviationX	DeviationM	Deviation3	DeviationH	DeviationS	DeviationN
1	-0.0647	-0.0362	-0.02234	-0.02620	-0.0016	0.0094	0.02281	10
2	-0.0332	-0.0201	0.01625	0.02045	0.0438	0.0564	0.03347	10
3	-0.0440	-0.0139	0.00604	0.00570	0.0285	0.0486	0.02885	10
4	0.0362	0.0530	0.06473	0.06030	0.0755	0.1073	0.02150	10
5	-0.0464	-0.0074	0.00813	0.00760	0.0302	0.0374	0.02593	10
6	-0.0384	-0.0285	-0.01283	-0.00950	0.0017	0.0071	0.01599	10

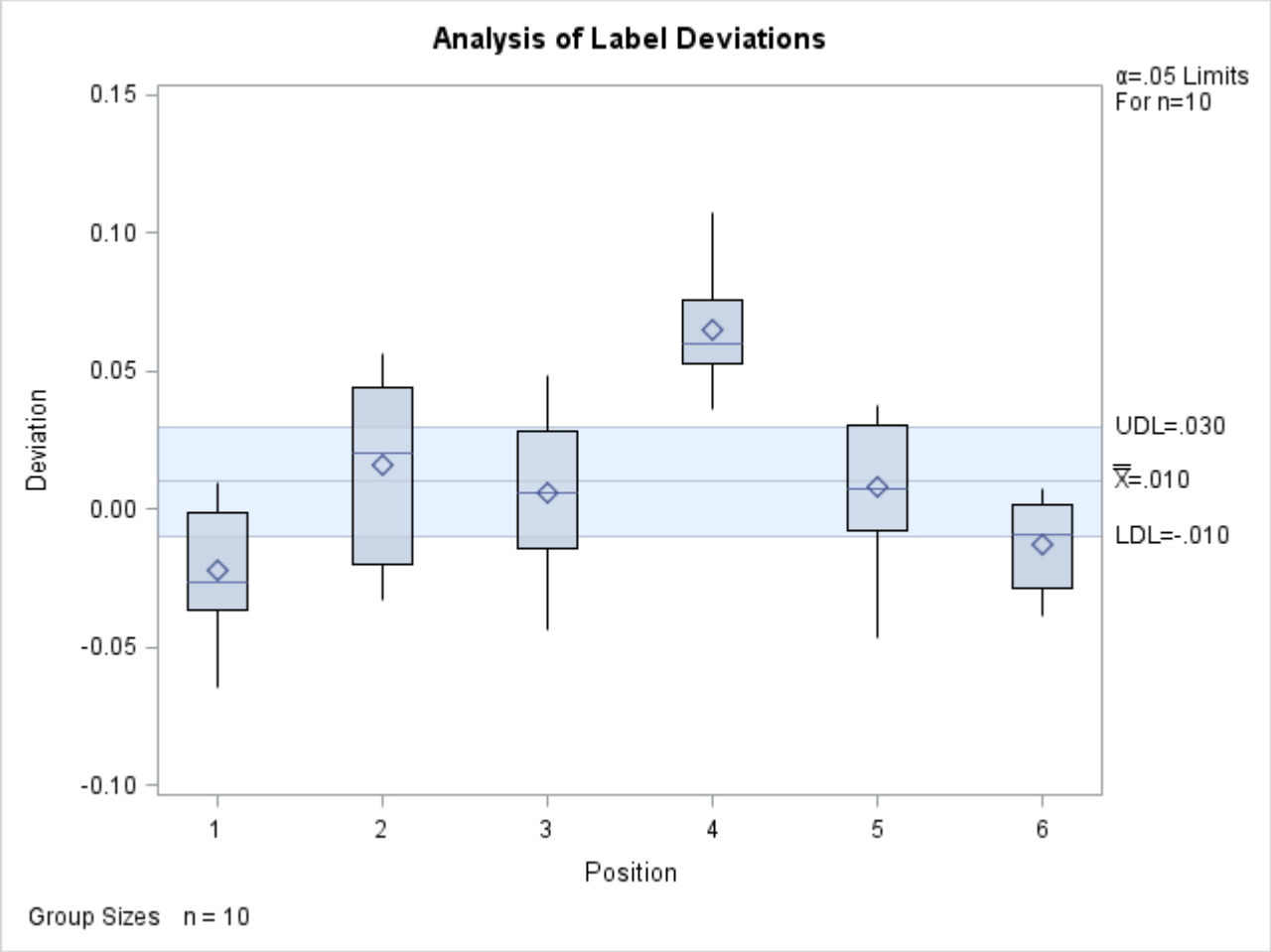
You can read this data set by specifying it as a SUMMARY= data set in the PROC ANOM statement, as follows:

```
ods graphics on;
title 'Analysis of Label Deviations';
proc anom summary=Labels;
   boxchart Deviation*Position / odstitle=title1;
run;
```

The resulting ANOM boxchart is shown in Figure 4.5.

Note that Deviation is *not* the name of a SAS variable in the data set but is, instead, the common prefix for the names of the eight summary variables. The suffix characters *L*, *1*, *X*, *M*, *3*, *H*, *S*, and *N* indicate the contents of the variable. For example, the suffix characters *1* and *3* indicate first and third quartiles. Thus, you can specify three group summary variables in a SUMMARY= data set with a single name (Deviation), which is referred to as the *response*. The name Position specified after the asterisk is the name of the *group-variable*.

Figure 4.5 ANOM Chart for Means in Data Set Labels



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

- group variable
- group minimum variable
- group first quartile variable
- group mean variable
- group median variable
- group third quartile variable
- group maximum variable
- group standard deviation variable
- group sample size variable

Furthermore, the names of the summary variables must begin with the *response* name specified in the BOXCHART statement and end with the appropriate suffix characters. If the names do not follow this convention, you can use the RENAME option in the PROC ANOM statement to rename the variables for the duration of the ANOM procedure step. If a label is associated with the group mean variable, it is used to label the vertical axis.

In summary, the interpretation of *response* depends on the input data set.

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

For more information, see "SUMMARY= Data Set" on page 71.

Saving Summary Statistics for Groups

NOTE: See *Saving Summary Statistics for Groups* in the SAS/QC Sample Library.

In this example, the BOXCHART statement is used to create a data set containing group summary statistics that can be read later by the ANOM procedure (as in the preceding example). The following statements read measurements from the data set LabelDeviations and create a summary data set named LabelSummary:

```
proc anom data=LabelDeviations;
   boxchart Deviation*Position / outsummary=LabelSummary
                                 nochart;
run;
```

The OUTSUMMARY= option names the output data set, and the NOCHART option suppresses the display of the chart, which would be identical to Figure 4.3.

Figure 4.6 contains a listing of LabelSummary.

Figure 4.6 The Summary Data Set LabelSummary

The Data Set LabelSummary

Position	DeviationL	Deviation1	DeviationX	DeviationM	Deviation3	DeviationH	DeviationS	DeviationN
1	-0.06466	-0.03623	-0.022342	-0.026195	-0.00163	0.00944	0.022805	10
2	-0.03323	-0.02014	0.016241	0.020420	0.04378	0.05641	0.033478	10
3	-0.04404	-0.01394	0.006011	0.005680	0.02849	0.04855	0.028847	10
4	0.03620	0.05298	0.064729	0.060315	0.07551	0.10729	0.021492	10
5	-0.04640	-0.00741	0.008121	0.007625	0.03021	0.03736	0.025920	10
6	-0.03839	-0.02845	-0.012812	-0.009495	0.00167	0.00710	0.015974	10

There are nine variables in the data set LabelSummary.

- Position identifies the group.
- DeviationL contains the group minimums.
- Deviation1 contains the first quartile for each group.
- DeviationX contains the group means.
- DeviationM contains the group medians.
- Deviation3 contains the third quartile for each group.
- DeviationH contains the group maximums.
- DeviationS contains the group standard deviations.
- DeviationN contains the group sizes.

Note that the summary statistic variables are named by adding the suffix characters *L*, *1*, *X*, *M*, *3*, *H*, *S*, and *N* to the *response* Deviation specified in the BOXCHART statement. In other words, the variable naming convention for OUTSUMMARY= data sets is the same as that for SUMMARY= data sets.

For more information, see "OUTSUMMARY= Data Set" on page 66.

Saving Decision Limits

NOTE: See *Saving Decision Limits Using ANOM BOXCHART* in the SAS/QC Sample Library.

You can save the decision limits for an ANOM chart, together with the parameters used to compute the limits, in a SAS data set.

The following statements read measurements from the data set LabelDeviations (see "Creating ANOM Boxcharts from Response Values" on page 45.) and save the decision limits displayed in Figure 4.3 in a data set named LabelLimits:

```
proc anom data=LabelDeviations;
   boxchart Deviation*Position / outlimits=LabelLimits
                                 nochart;
run;
```

The OUTLIMITS= option names the data set containing the decision limits, and the NOCHART option suppresses the display of the chart. The data set LabelLimits is listed in Figure 4.7.

Figure 4.7 The Data Set LabelLimits Containing Decision Limit Information

Decision Limits for Labler Position Deviations

VAR	_GROUP_	_TYPE_	_LIMITN_	_ALPHA_	_LDLX_	_MEAN_	_UDLX_	_MSE_	_DFE_	_LIMITK_
Deviation	Position	ESTIMATE	10	0.05	-.009878975	.009991333	0.029862	.000643646	54	6

The data set LabelLimits contains one observation with the limits for *response* Deviation. The values of _LDLX_ and _UDLX_ are the lower and upper decision limits for the means, and the value of _MEAN_ is the weighted average of the group means, which is represented by the central line.

The values of _MEAN_, _MSE_, _DFE_, _LIMITK_, _LIMITN_, and _ALPHA_ are the parameters used to compute the decision limits. The value of _MSE_ is the mean square error, and the value of _DFE_ is the associated degrees of freedom. The value of _LIMITK_ is the group size (k), the value of _LIMITN_ is the nominal sample size associated with the decision limits, and the value of _ALPHA_ is the value of the significance level (α). The variables _VAR_ and _GROUP_ are bookkeeping variables that save the *response* and *group-variable*. The variable _TYPE_ is a bookkeeping variable that indicates whether the values of _MEAN_ and _MSE_ are estimates computed from the data or standard (known) values specified with procedure options. In most applications, the value of _TYPE_ will be 'ESTIMATE.'

NOTE: See *Saving Decision Limits and Summary Statistics* in the SAS/QC Sample Library.

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

```
proc anom data=LabelDeviations;
   boxchart Deviation*Position / outtable=LabelTab
                                 nochart;
run;
```

The data set LabelTab is listed in Figure 4.8.

Figure 4.8 The Data Set LabelTab

Summary Statistics and Decision Limits

VAR	Position	_ALPHA_	_LIMITN_	_SUBN_	_LDLX_	_SUBX_	_MEAN_	_UDLX_
Deviation	1	0.05	10	10	-.009878975	-0.022342	.009991333	0.029862
Deviation	2	0.05	10	10	-.009878975	0.016241	.009991333	0.029862
Deviation	3	0.05	10	10	-.009878975	0.006011	.009991333	0.029862
Deviation	4	0.05	10	10	-.009878975	0.064729	.009991333	0.029862
Deviation	5	0.05	10	10	-.009878975	0.008121	.009991333	0.029862
Deviation	6	0.05	10	10	-.009878975	-0.012812	.009991333	0.029862

EXLIM	_SUBMIN_	_SUBQ1_	_SUBMED_	_SUBQ3_	_SUBMAX_
LOWER	-0.06466	-0.03623	-0.026195	-0.00163	0.00944
	-0.03323	-0.02014	0.020420	0.04378	0.05641
	-0.04404	-0.01394	0.005680	0.02849	0.04855
UPPER	0.03620	0.05298	0.060315	0.07551	0.10729
	-0.04640	-0.00741	0.007625	0.03021	0.03736
LOWER	-0.03839	-0.02845	-0.009495	0.00167	0.00710

This data set contains one observation for each group sample. The variable _SUBMIN_ contains the group minimums, and the variable _SUBQ1_ contains the first quartile for each group. The variables _SUBX_ and _SUBMED_ contain the group means and medians. The variable _SUBQ3_ contains the third quartiles, _SUBMAX_ contains the group maximums, and _SUBN_ contains the group sample sizes. The variables _LDLX_ and _UDLX_ contain the lower and upper decision limits, and the variable _MEAN_ contains the central line. The variables _VAR_ and Position contain the *response* name and values of the *group-variable*, respectively. For more information, see "OUTTABLE= Data Set" on page 67.

An OUTTABLE= data set can be read later as a TABLE= data set. For example, the following statements read LabelTab and display an ANOM boxchart (not shown here) identical to the chart in Figure 4.3:

```
title 'Analysis of Label Deviations';
proc anom table=LabelTab;
   boxchart Deviation*Position / odstitle=title;
label _SUBX_ = 'Mean Deviation from Center (mm)';
run;
```

Because the ANOM procedure simply displays the information in a TABLE= data set, you can use TABLE= data sets to create specialized ANOM boxcharts.

For more information, see "TABLE= Data Set" on page 73.

Syntax: BOXCHART Statement

The basic syntax for the BOXCHART statement is as follows:

BOXCHART *response* ∗ *group-variable* **;**

The general form of this syntax is as follows:

BOXCHART *(responses)* ∗ *group-variable* < **(***block-variables***)** >
< **=***symbol-variable* > < *options* > **;**

You can use any number of BOXCHART statements in the ANOM procedure. The components of the BOXCHART statement are described as follows.

responses

identify one or more responses to be analyzed. The specification of *response* depends on the input data set specified in the PROC ANOM statement.

- If response values (raw data) are read from a DATA= data set, *response* must be the name of the variable containing the values.For an example, see "Creating ANOM Boxcharts from Response Values" on page 45.
- If summary data are read from a SUMMARY= data set, *response* must be the common prefix of the summary variables in the SUMMARY= data set. For an example, see "Creating ANOM Boxcharts from Group Summary Data" on page 47.
- If summary data and decision limits are read from a TABLE= data set, *response* must be the value of the variable _VAR_ in the TABLE= data set. For an example, see "Saving Decision Limits" on page 51.

A *response* is required. If you specify more than one response, enclose the list in parentheses. For example, the following statements request distinct ANOM charts for the means of Weight, Length, and Width:

```
proc anom data=Measures;
   xchart (Weight Length Width)*Day;
run;
```

group-variable
is the variable that identifies groups in the data. The *group-variable* is required. In the preceding BOXCHART statement, Day is the group variable.

block-variables
are optional variables that group the data into blocks of consecutive groups. The blocks are labeled in a legend, and each *block-variable* provides one level of labels in the legend.

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

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

Summary of Options

The following tables list the BOXCHART statement options by function. Options unique to the ANOM procedure are listed in Table 4.1, and are described in detail in "Dictionary of ANOM Chart Statement Options" on page 180. Options that are common to both the ANOM and SHEWHART procedures are listed in Table 4.2, and are described in detail in "Dictionary of Options: SHEWHART Procedure" on page 1946.

Table 4.1 BOXCHART Statement Special Options

Option	Description
Options for Specifying Parameters for Decision Limits	
ALPHA=	specifies the probability of a Type I error
DFE=	specifies the degrees of freedom associated with the root mean square error
LIMITK=	specifies number of groups for decision limits
LIMITN=	specifies either nominal sample size for fixed decision limits or varying limits
MEAN=	specifies the mean
MSE=	specifies the mean square error
NOREADLIMITS	computes decision limits for each *response* from the data rather than a LIMITS= data set
READINDEXES=	reads multiple sets of decision limits for each *response* from a LIMITS= data set

Table 4.1 *continued*

Option	Description
TYPE=	identifies parameters as estimates or standard values and specifies value of _TYPE_ in the OUTLIMITS= data set
Options for Displaying Decision Limits	
CINFILL=	specifies color for area inside decision limits
CLIMITS=	specifies color of decision limits, central line, and related labels
LDLLABEL=	specifies label for lower decision limit
LIMLABSUBCHAR=	specifies a substitution character for labels provided as quoted strings; the character is replaced with the value of the decision limit
LLIMITS=	specifies line type for decision limits
NDECIMAL=	specifies number of digits to right of decimal place in default labels for decision limits and central line
NOCTL	suppresses display of central line
NOLDL	suppresses display of lower decision limit
NOLIMITLABEL	suppresses labels for decision limits and central line
NOLIMITS	suppresses display of decision limits
NOLIMITSFRAME	suppresses default frame around decision limit information when multiple sets of decision limits are read from a LIMITS= data set
NOLIMITSLEGEND	suppresses legend for decision limits
NOUDL	suppresses display of upper decision limit
UDLLABEL=	specifies label for upper decision limit
WLIMITS=	specifies width for decision limits and central line
XSYMBOL=	specifies label for central line
Output Data Set Option	
OUTSUMMARY=	creates output data set containing group summary statistics

Table 4.2 BOXCHART Statement General Options

Option	Description
Options for Controlling Box Appearance	
BOXCONNECT=	connects group means, medians, maximum values, minimum values, or quartiles in box-and-whisker plots
BOXSTYLE=	specifies style of box-and-whisker plots
BOXWIDTH=	specifies width of box-and-whisker plots
BOXWIDTHSCALE=	specifies that widths of box-and-whisker plots vary proportionately to group sample size
CBOXES=	specifies color for outlines of box-and-whisker plots
CBOXFILL=	specifies fill color for interior of box-and-whisker plots
IDCOLOR=	specifies outlier symbol color in schematic box-and-whisker plots

Table 4.2 *continued*

Option	Description
IDCTEXT=	specifies text color to label outliers or response variable values
IDFONT=	specifies text font to label outliers or response variable values
IDHEIGHT=	specifies text height to label outliers or response variable values
IDSYMBOL=	specifies outlier symbol in schematic box-and-whisker plots
LBOXES=	specifies line types for outlines of box-and-whisker plots
NOTCHES	specifies that box-and-whisker plots are to be notched
PCTLDEF=	specifies percentile definition used for box-and-whisker plots
SERIFS	adds serifs to the whiskers of skeletal box-and-whisker plots
Options for Plotting and Labeling Points	
ALLLABEL=	labels every point on ANOM boxchart
CLABEL=	specifies color for labels
CCONNECT=	specifies color for line segments that connect points on chart
CFRAMELAB=	specifies fill color for frame around labeled points
COUT=	specifies color for portions of line segments that connect points outside decision limits
LABELANGLE=	specifies angle at which labels are drawn
LABELFONT=	specifies software font for labels
LABELHEIGHT=	specifies height of labels
OUTLABEL=	labels points outside decision limits
SYMBOLLEGEND=	specifies LEGEND statement for levels of *symbol-variable*
SYMBOLORDER=	specifies order in which symbols are assigned for levels of *symbol-variable*
TURNALL|TURNOUT	turns point labels so that they are strung out vertically
Axis and Axis Label Options	
CAXIS=	specifies color for axis lines and tick marks
CFRAME=	specifies fill colors for frame for plot area
CTEXT=	specifies color for tick mark values and axis labels
DISCRETE	produces horizontal axis for discrete numeric group values
HAXIS=	specifies major tick mark values for horizontal axis
HEIGHT=	specifies height of axis label and axis legend text
HMINOR=	specifies number of minor tick marks between major tick marks on horizontal axis
HOFFSET=	specifies length of offset at both ends of horizontal axis

Table 4.2 *continued*

Option	Description
NOHLABEL	suppresses label for horizontal axis
NOTICKREP	specifies that only the first occurrence of repeated, adjacent group values is to be labeled on horizontal axis
NOVANGLE	requests vertical axis labels that are strung out vertically
SKIPHLABELS=	specifies thinning factor for tick mark labels on horizontal axis
TURNHLABELS	requests horizontal axis labels that are strung out vertically
VAXIS=	specifies major tick mark values for vertical axis of ANOM boxchart
VFORMAT=	specifies format for vertical axis tick mark labels
VMINOR=	specifies number of minor tick marks between major tick marks on vertical axis
VOFFSET=	specifies length of offset at both ends of vertical axis
VZERO	forces origin to be included in vertical axis for ANOM boxchart
WAXIS=	specifies width of axis lines
Plot Layout Options	
ALLN	plots means for all groups
BILEVEL	creates ANOM boxchart using half-screens and half-pages
EXCHART	creates ANOM boxchart for a response only when a group mean exceeds the decision limits
INTERVAL=	natural time interval between consecutive group positions when time, date, or datetime format is associated with a numeric group variable
MAXPANELS=	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 decision limits
NOCHART	suppresses creation of chart
NOFRAME	suppresses frame for plot area
NOLEGEND	suppresses legend for group sample sizes
NPANELPOS=	specifies number of group positions per panel on each chart
REPEAT	repeats last group position on panel as first group position of next panel
TOTPANELS=	specifies number of pages or screens to be used to display chart
ZEROSTD	displays ANOM boxchart regardless of whether root mean square error is zero

Table 4.2 *continued*

Option	Description
Reference Line Options	
CHREF=	specifies color for lines requested by HREF= option
CVREF=	specifies color for lines requested by VREF= option
HREF=	specifies position of reference lines perpendicular to horizontal axis on ANOM boxchart
HREFDATA=	specifies position of reference lines perpendicular to horizontal axis on ANOM boxchart
HREFLABELS=	specifies labels for HREF= lines
HREFLABPOS=	specifies position of HREFLABELS= labels
LHREF=	specifies line type for HREF= lines
LVREF=	specifies line type for VREF= lines
NOBYREF	specifies that reference line information in a data set applies uniformly to charts created for all BY groups
VREF=	specifies position of reference lines perpendicular to vertical axis on ANOM boxchart
VREFLABELS=	specifies labels for VREF= lines
VREFLABPOS=	specifies position of VREFLABELS= labels
Grid Options	
CGRID=	specifies color for grid requested with GRID or ENDGRID option
ENDGRID	adds grid after last plotted point
GRID	adds grid to control chart
LENDGRID=	specifies line type for grid requested with the ENDGRID option
LGRID=	specifies line type for grid requested with the GRID option
WGRID=	specifies width of grid lines
Clipping Options	
CCLIP=	specifies color for plot symbol for clipped points
CLIPFACTOR=	determines extent to which extreme points are clipped
CLIPLEGEND=	specifies text for clipping legend
CLIPLEGPOS=	specifies position of clipping legend
CLIPSUBCHAR=	specifies substitution character for CLIPLEGEND= text
CLIPSYMBOL=	specifies plot symbol for clipped points
CLIPSYMBOLHT=	specifies symbol marker height for clipped points
Graphical Enhancement Options	
ANNOTATE=	specifies annotate data set that adds features to ANOM boxchart
DESCRIPTION=	specifies description of ANOM boxchart's GRSEG catalog entry
FONT=	specifies software font for labels and legends on chart

Table 4.2 *continued*

Option	Description
NAME=	specifies name of ANOM boxchart's GRSEG catalog entry
PAGENUM=	specifies the form of the label used in pagination
PAGENUMPOS=	specifies the position of the page number requested with the PAGENUM= option
Options for Producing Graphs Using ODS Styles	
BLOCKVAR=	specifies one or more variables whose values define colors for filling background of *block-variable* legend
BOXES=	specifies variables whose values define colors box outlines
BOXFILL=	specifies variables whose values define colors for filling boxes
CFRAMELAB	draws a frame around labeled points
CPHASEBOX	requests boxes enclosing all plotted points for a phase
CPHASEBOXCONNECT	requests lines connecting adjacent enclosing boxes
CPHASEBOXFILL	fills boxes enclosing all plotted points for a phase
CPHASEMEANCONNECT	requests lines connecting phase average value points
Options for ODS Graphics	
BLOCKREFTRANSPARENCY=	specifies the wall fill transparency for blocks and phases
BOXTRANSPARENCY=	specifies the box fill transparency for box-and-whisker charts
INFILLTRANSPARENCY=	specifies the decision limit infill transparency
NOBLOCKREF	suppresses block and phase reference lines
NOBLOCKREFFILL	suppresses block and phase wall fills
NOBOXFILLLEGEND	suppresses legend for levels of a BOXFILL= variable
NOFILLLEGEND	suppresses legend for levels of a BOXFILL= variable
NOPHASEREF	suppresses block and phase reference lines
NOPHASEREFFILL	suppresses block and phase wall fills
NOREF	suppresses block and phase reference lines
NOREFFILL	suppresses block and phase wall fills
NOTRANSPARENCY	disables transparency in ODS Graphics output
ODSFOOTNOTE=	specifies a graph footnote
ODSLEGENDEXPAND	specifies that legend entries contain all levels observed in the data
ODSTITLE=	specifies a graph title
OVERLAYURL=	specifies URLs to associate with overlay points
PHASEBOXLABELS	draws phase labels as titles along the top of phase boxes
PHASEPOS=	specifies vertical position of phase legend
PHASEREFLEVEL=	associates phase and block reference lines with either innermost or the outermost level
PHASEREFTRANSPARENCY=	specifies the wall fill transparency for blocks and phases
REFFILLTRANSPARENCY=	specifies the wall fill transparency for blocks and phases

Table 4.2 *continued*

Option	Description
SIMULATEQCFONT	draws central line labels using a simulated software font
URL=	specifies a variable whose values are URLs to be associated with groups
WBOXES=	specifies width of box outlines for box-and-whisker charts
Input Data Set Options	
MISSBREAK	specifies that observations with missing values are not to be processed
Output Data Set Options	
OUTBOX=	creates output data set containing group summary statistics, decision limits, and outlier values
OUTINDEX=	specifies value of _INDEX_ in the OUTLIMITS= data set
OUTLIMITS=	creates output data set containing decision limits
OUTTABLE=	creates output data set containing group summary statistics and decision limits
Tabulation Options	
NOTE: specifying (EXCEPTIONS) after a tabulation option creates a table for exceptional points only.	
TABLE	creates a basic table of group means, group sample sizes, and decision 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
TABLEOUTLIM	augments basic table with columns indicating decision limits exceeded
Block Variable Legend Options	
BLOCKLABELPOS=	specifies position of label for *block-variable* legend
BLOCKLABTYPE=	specifies text size of *block-variable* legend
BLOCKPOS=	specifies vertical position of *block-variable* legend
BLOCKREP	repeats identical consecutive labels in *block-variable* legend
CBLOCKLAB=	specifies fill colors for frames enclosing variable labels in *block-variable* legend
CBLOCKVAR=	specifies one or more variables whose values are colors for filling background of *block-variable* legend

Table 4.2 *continued*

Option	Description
Phase Options	
CPHASELEG=	specifies text color for *phase* legend
NOPHASEFRAME	suppresses default frame for *phase* legend
OUTPHASE=	specifies value of _PHASE_ in the OUTSUMMARY= data set
PHASEBREAK	disconnects last point in a *phase* from first point in next *phase*
PHASELABTYPE=	specifies text size of *phase* legend
PHASELEGEND	displays *phase* labels in a legend across top of chart
PHASELIMITS	labels decision limits for each phase, provided they are constant within that phase
PHASEREF	delineates *phases* with vertical reference lines
READPHASES=	specifies *phases* to be read from an input data set
Overlay Options	
CCOVERLAY=	specifies colors for overlay line segments
COVERLAY=	specifies colors for overlay plots
COVERLAYCLIP=	specifies color for clipped points on overlays
LOVERLAY=	specifies line types for overlay line segments
NOOVERLAYLEGEND	suppresses legend for overlay plots
OVERLAY=	specifies variables to overlay on chart
OVERLAYCLIPSYM=	specifies symbol for clipped points on overlays
OVERLAYCLIPSYMHT=	specifies symbol height for clipped points on overlays
OVERLAYHTML=	specifies links to associate with overlay points
OVERLAYID=	specifies labels for overlay points
OVERLAYLEGLAB=	specifies label for overlay legend
OVERLAYSYM=	specifies symbols for overlays
OVERLAYSYMHT=	specifies symbol heights for overlays
WOVERLAY=	specifies widths of overlay line segments
Options for Interactive ANOM Charts	
HTML=	specifies a variable whose values create links to be associated with groups
HTML_LEGEND=	specifies a variable whose values create links to be associated with symbols in the symbol legend
WEBOUT=	creates an OUTTABLE= data set with additional graphics coordinate data

Details: BOXCHART Statement

Constructing ANOM Boxcharts

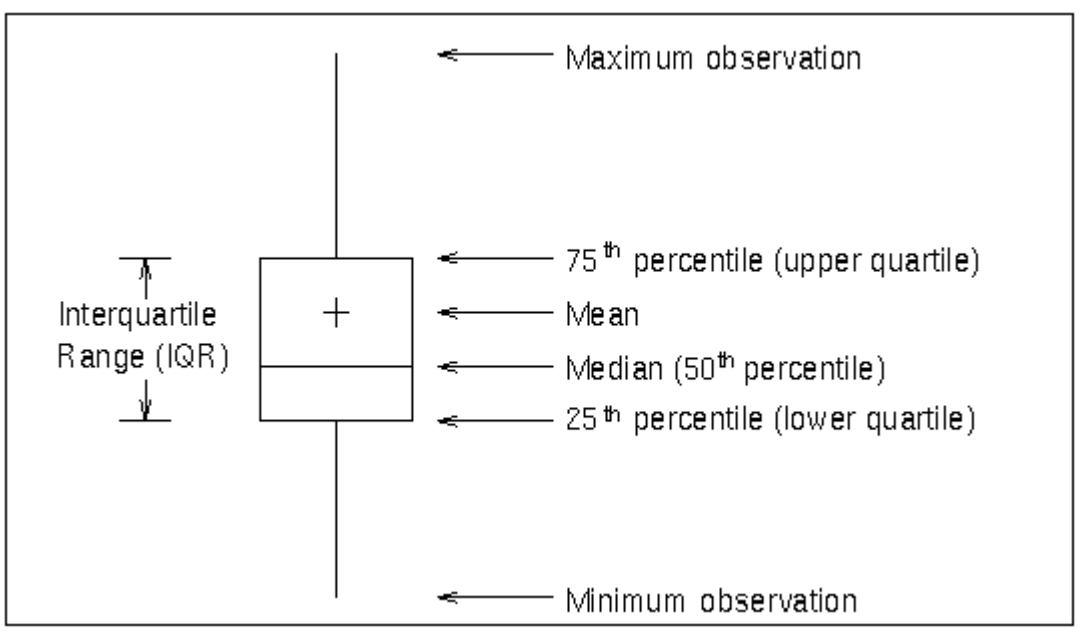
The following notation is used in this section:

X_{ij}	jth response in the ith group
k	number of groups
n_i	sample size of ith group
N	total sample size $= n_1 + \cdots + n_k$
μ_i	expected value of a response in the ith group
σ	standard deviation of response
$\bar{X}_i$	average response in ith group
$\overline{\overline{X}}$	weighted average of k group means
s_i^2	sample variance of the responses in the ith group
$\widehat{\sigma^2}$	mean square error (MSE)
ν	degrees of freedom associated with the mean square error
α	significance level
$h(\alpha; k, n, \nu)$	critical value for analysis of means when the sample sizes n_i are equal ($n_i \equiv n$)
$h(\alpha; k, n_1, \ldots, n_k, \nu)$	critical value for analysis of means when the sample sizes n_i are not equal

Elements of Box-and-Whisker Plots

A box-and-whisker plot is displayed for the measurements in each group on the ANOM boxchart. Figure 4.9 illustrates the elements of each plot.

Figure 4.9 Box-and-Whisker Plot



The skeletal style of the box-and-whisker plot shown in Figure 4.9 is the default. You can specify alternative styles with the BOXSTYLE= option; see the entry for the BOXSTYLE= option in "Dictionary of Options: SHEWHART Procedure" on page 1946.

Central Line

By default, the central line on an ANOM chart for means represents the weighted average of the group means, which is computed as

$$\overline{\overline{X}} = \frac{n_1\bar{X}_1 + \cdots + n_k\bar{X}_k}{n_1 + \cdots + n_k}$$

You can specify a value for $\overline{\overline{X}}$ with the MEAN= option in the BOXCHART statement or with the variable _MEAN_ in a LIMITS= data set.

Decision Limits

In the analysis of means for continuous data, it is assumed that the responses in the ith group are at least approximately normally distributed with a constant variance:

$$X_{ij} \sim N(\mu_i, \sigma^2), \qquad j = 1, \ldots, n_i$$

When the group sizes are constant ($n_i \equiv n$), then $\nu = N - k = k(n-1)$ and the decision limits are computed as follows:

$$\begin{aligned}
\text{lower decision limit (LDL)} &= \overline{\overline{X}} - h(\alpha; k, n, \nu)\sqrt{\text{MSE}}\sqrt{\frac{k-1}{N}} \\
\text{upper decision limit (UDL)} &= \overline{\overline{X}} + h(\alpha; k, n, \nu)\sqrt{\text{MSE}}\sqrt{\frac{k-1}{N}}
\end{aligned}$$

Here the mean square error (MSE) is computed as follows:

$$\text{MSE} = \widehat{\sigma^2} = \frac{1}{k}\sum_{j=1}^{k} s_j^2$$

For details concerning the function $h(\alpha; k, n, \nu)$, see Nelson (1981, 1982a, 1993).

When the group sizes n_i are not constant (the unbalanced case), $\nu = N - k$ and the decision limits for the ith group are computed as follows:

$$\begin{aligned}
\text{lower decision limit (LDL)} &= \overline{\overline{X}} - h(\alpha; k, n_1, \ldots, n_k, \nu)\sqrt{\text{MSE}}\sqrt{\frac{N-n_i}{Nn_i}} \\
\text{upper decision limit (UDL)} &= \overline{\overline{X}} + h(\alpha; k, n_1, \ldots, n_k, \nu)\sqrt{\text{MSE}}\sqrt{\frac{N-n_i}{Nn_i}}
\end{aligned}$$

Here the mean square error (MSE) is computed as follows:

$$\text{MSE} = \widehat{\sigma^2} = \frac{(n_1 - 1)s_1^2 + \cdots + (n_k - 1)s_k^2}{n_1 + \cdots + n_k - k}$$

This requires that ν be positive. A chart is not produced if $\nu > 0$ but MSE is equal to zero (unless you specify the ZEROSTD option). For details concerning the function $h(\alpha; k, n_1, \ldots, n_k, \nu)$, see Fritzsch and Hsu (1997), Nelson (1982b, 1991), and Soong and Hsu (1997).

You can specify parameters for the limits as follows:

- Specify α with the ALPHA= option or with the variable _ALPHA_ in a LIMITS= data set. By default, $\alpha = 0.05$.
- Specify a constant nominal sample size $n_i \equiv n$ for the decision limits in the balanced case with the LIMITN= option or with the variable _LIMITN_ in a LIMITS= data set. By default, n is the observed sample size in the balanced case.
- Specify k with the LIMITK= option or with the variable _LIMITK_ in a LIMITS= data set. By default, k is the number of groups.
- Specify $\overline{\overline{X}}$ with the MEAN= option or with the variable _MEAN_ in a LIMITS= data set. By default, $\overline{\overline{X}}$ is the weighted average of the responses.
- Specify $\widehat{\sigma^2}$ with the MSE= option or with the variable _MSE_ in a LIMITS= data set. By default, $\widehat{\sigma^2}$ is computed as indicated above.
- Specify ν with the DFE= option or with the variable _DFE_ in a LIMITS= data set. By default, ν is determined as indicated above.

Output Data Sets

OUTBOX= Data Set

The OUTBOX= data set saves group summary statistics, decision limits, and outlier values. The following variables can be saved:

- the *group-variable*
- the variable _VAR_, containing the analysis variable name
- the variable _TYPE_, identifying features of box-and-whisker plots
- the variable _VALUE_, containing values of box-and-whisker plot features
- the variable _ID_, containing labels for outliers
- the variable _HTML_, containing links associated with box-and-whisker plot features

_ID_ is included in the OUTBOX= data set only if one of the keywords SCHEMATICID or SCHEMATICID-FAR is specified with the BOXSTYLE= option. _HTML_ is present only if the HTML= or HTML2= option is specified.

Each observation in an OUTBOX= data set records the value of a single feature of one group's box-and-whisker plot, such as its mean. The _TYPE_ variable identifies the feature whose value is recorded in _VALUE_. The following table lists valid _TYPE_ variable values:

Table 4.3 Valid _TYPE_ Values in an OUTBOX= Data Set

TYPE Value	Description
N	group size
ALPHA	significance level
LIMITN	nominal sample size associated with decision limits
LDLX	lower decision limit for group mean
UDLX	upper decision limit for group mean
RESPMEAN	overall response variable mean
MIN	group minimum value
Q1	group first quartile
MEDIAN	group median
MEAN	group mean
Q3	group third quartile
MAX	group maximum value
LOW	low outlier value
HIGH	high outlier value
LOWHISKR	low whisker value, if different from MIN
HIWHISKR	high whisker value, if different from MAX
FARLOW	low far outlier value
FARHIGH	high far outlier value

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

- *block-variables*
- *symbol-variable*
- BY variables
- ID variables

OUTLIMITS= Data Set

The OUTLIMITS= data set saves decision limits and decision limit parameters. The following variables can be saved:

Table 4.4 OUTLIMITS= Data Set

Variable	Description
ALPHA	significance level
DFE	degrees of freedom for mean square error
GROUP	*group-variable* specified in the BOXCHART statement
INDEX	optional identifier for the decision limits specified with the OUTINDEX= option
LDLX	lower decision limit for group means
LIMITK	number of groups
LIMITN	sample size associated with the decision limits
MEAN	weighted average of group means ($\overline{\overline{X}}$)
MSE	mean square error
TYPE	type (estimate or standard value) of _MEAN_ and _MSE_
UDLX	upper decision limit for group means
VAR	*response* specified in the BOXCHART statement

Notes:

1. In the unbalanced case, the special missing value *V* is assigned to the variables _LIMITN_, _LDLX_, and _UDLX_.
2. Optional BY variables are saved in the OUTLIMITS= data set.

The OUTLIMITS= data set contains one observation for each *response* specified in the BOXCHART statement. For an example, see "Saving Decision Limits" on page 51.

OUTSUMMARY= Data Set

The OUTSUMMARY= data set saves group summary statistics. The following variables can be saved:

- the *group-variable*
- a group minimum variable named by *response* suffixed with *L*
- a group first-quartile variable named by *response* suffixed with *1*
- a group mean variable named by *response* suffixed with *X*
- a group median variable named by *response* suffixed with *M*
- a group third-quartile variable named by *response* suffixed with *3*
- a group maximum variable named by *response* suffixed with *H*
- a group standard deviation variable named by *response* suffixed with *S*
- a group sample size variable named by *response* suffixed with *N*

Given a *response* name that contains 32 characters, the procedure first shortens the name to its first 16 characters and its last 15 characters, and then it adds the suffix.

Group summary variables are created for each *response* specified in the BOXCHART statement. For example, consider the following statements:

```
proc anom data=Steel;
   xchart (Width Diameter)*Lot / outsummary=Summary;
run;
```

The data set Summary contains variables named Lot, WidthL, Width1, WidthX, WidthM, Width3, WidthH, WidthS, WidthN, DiameterL, Diameter1, DiameterX, DiameterM, Diameter3, DiameterH, DiameterS, and DiameterN. 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 OUTSUMMARY= data set, see "Saving Summary Statistics for Groups" on page 50.

OUTTABLE= Data Set

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

Variable	Description
ALPHA	significance level
EXLIM	decision limit exceeded (if any)
group	values of the group variable
LDLX	lower decision limit for group mean
LIMITN	nominal sample size associated with the decision limits
MEAN	central line
SUBMAX	group maximum
SUBMED	group median
SUBMIN	group minimum
SUBN	group sample size
SUBQ1	group first quartile
SUBQ3	group third quartile
SUBX	group mean
UDLX	upper decision limit for group mean
VAR	*response* specified in the BOXCHART 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)

NOTE: The variable _EXLIM_ is a character variable of length 8. The variable _PHASE_ is a character variable of length 48. The variable _VAR_ is a character variable whose length is no greater than 32. All other variables are numeric.

For an example, see "Saving Decision Limits" on page 51.

ODS Tables

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

Table 4.5 ODS Tables Produced with the BOXCHART Statement

Table Name	Description	Options
BOXCHART	ANOM chart summary statistics	TABLE, TABLEALL, TABLEC, TABLEID, TABLEOUT

ODS Graphics

Before you create ODS Graphics output, ODS Graphics must be enabled (for example, by using the ODS GRAPHICS ON statement). For more information about enabling and disabling ODS Graphics, see the section "Enabling and Disabling ODS Graphics" (Chapter 21, *SAS/STAT User's Guide*). **NOTE:** In SAS/QC 13.1 the ANOM procedure does not support the creation of graphs that are editable with the ODS Graphics Editor.

The appearance of a graph produced with ODS Graphics is determined by the style associated with the ODS destination where the graph is produced. BOXCHART options used to control the appearance of traditional graphics are ignored for ODS Graphics output. Options for Producing Graphs Using ODS Styles lists options that can be used to control the appearance of graphs produced with ODS Graphics or with traditional graphics using ODS styles. Options for ODS Graphics lists options to be used exclusively with ODS Graphics. Detailed descriptions of these options are provided in "Dictionary of Options: SHEWHART Procedure" on page 1946.

When ODS Graphics is in effect, the BOXCHART statement assigns a name to the graph it creates. You can use this name to reference the graph when using ODS. The name is listed in Table 4.6.

Table 4.6 ODS Graphics Produced by the BOXCHART Statement

ODS Graph Name	Plot Description
BoxChart	ANOM boxchart

See Chapter 3, "SAS/QC Graphics," for more information about ODS Graphics and other methods for producing charts.

Input Data Sets

BOX= Data Set

You can read summary statistics, decision limits, and outlier values from a BOX= data set specified in the PROC ANOM statement. This enables you to reuse an OUTBOX= data set created in a previous run of the ANOM procedure to display a box chart.

A BOX= data set must contain the following variables:

- the group variable
- _VAR_, containing the analysis variable name
- _TYPE_, identifying features of box-and-whisker plots
- _VALUE_, containing values of those features

Each observation in a BOX= data set records the value of a single feature of one group's box-and-whisker plot, such as its mean. The _TYPE_ variable identifies the feature whose value is recorded in a given observation. The following table lists valid _TYPE_ variable values:

Table 4.7 Valid _TYPE_ Values in a BOX= Data Set

TYPE Value	Description
N	group size
ALPHA	significance level
LIMITN	nominal sample size associated with decision limits
LDLX	lower decision limit for group mean
UDLX	upper decision limit for group mean
RESPMEAN	overall response variable mean
MIN	group minimum value
Q1	group first quartile
MEDIAN	group median
MEAN	group mean
Q3	group third quartile
MAX	group maximum value
LOW	low outlier value
HIGH	high outlier value
LOWHISKR	low whisker value, if different from MIN
HIWHISKR	high whisker value, if different from MAX
FARLOW	low far outlier value
FARHIGH	high far outlier value

The features identified by _TYPE_ values N, LDLX, UDLX, RESPMEAN, MIN, Q1, MEDIAN, MEAN, Q3, and MAX are required for each group.

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

- the variable _ID_, containing labels for outliers

- the variable _HTML_, containing links to be associated with features on box plots
- *block-variables*
- *symbol-variable*
- BY variables
- ID variables

When you specify one of the keywords SCHEMATICID or SCHEMATICIDFAR with the BOXSTYLE= option, values of _ID_ are used as outlier labels. If _ID_ does not exist in the BOX= data set, the values of the first variable listed in the ID statement are used.

DATA= Data Set

You can read raw data (response values) from a DATA= data set specified in the PROC ANOM statement. Each *response* specified in the BOXCHART statement must be a SAS variable in the DATA= data set. This variable provides measurements that must be grouped into group samples indexed by the *group-variable*. The *group-variable*, which is specified in the BOXCHART 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 *response* and a value for the *group-variable*. If the *i*th group contains n_i items, there should be n_i consecutive observations for which the value of the *group-variable* is the index of the *i*th group. For example, if each group contains five items and there are 10 groups, the DATA= data set should contain 50 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 ANOM 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 of a DATA= data set, see "Creating ANOM Boxcharts from Response Values" on page 45.

LIMITS= Data Set

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

```
proc anom data=Info limits=Conlims;
   xchart Weight*Batch;
run;
```

The LIMITS= data set can be an OUTLIMITS= data set that was created in a previous run of the ANOM procedure. Such data sets always contain the variables required for a LIMITS= data set; see Table 4.4. 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 _LDLX_, _MEAN_, and _UDLX_, which specify the decision limits directly
- the variables _MEAN_, _MSE_, and _DFE_, which are used to calculate the decision limits according to the equations in the section "Decision Limits" on page 63.

In addition, note the following:

- The variables _VAR_ and _GROUP_ are required. These must be character variables whose lengths are no greater than 32.
- _DFE_ is optional. The default is $\nu = N - k$, and in the case of equal group sizes, $\nu = k(n-1)$.
- _MSE_ is optional if _LDLX_ and _UDLX_ are specified; otherwise it is required.
- _LDLX_ and _UDLX_ must be specified together; otherwise their values are computed.
- _ALPHA_ is optional but is recommended in order to maintain a complete set of decision limit information. The default value is 0.05.
- _LIMITK_ is optional. The default value is k, the number of groups. A group must have at least one nonmissing value ($n_i \geq 1$) and there must be at least one group with $n_i \geq 2$. If specified, _LIMITK_ overrides the value of k.
- _LIMITN_ is optional. The default value is the common group size (n), in the balanced case $n_i \equiv n$. If specified, _LIMITN_ overrides the value of n.
- The variable _TYPE_ is optional, but is recommended to maintain a complete set of decision limit information. The variable _TYPE_ must be a character variable of length 8. Valid values are 'ESTIMATE,' 'STANDARD,' 'STDMEAN,' and 'STDRMS.' The default is 'ESTIMATE.'
- The variable _INDEX_ is required if you specify the READINDEX= option; this must be a character variable whose length is no greater than 48.
- BY variables are required if specified with a BY statement.

SUMMARY= Data Set

You can read group summary statistics from a SUMMARY= data set specified in the PROC ANOM statement. This enables you to reuse OUTSUMMARY= data sets that have been created in previous runs of the ANOM procedure or to read output data sets created with SAS summarization procedures, such as PROC MEANS.

A SUMMARY= data set used with the BOXCHART statement must contain the following:

- the *group-variable*
- a group minimum variable for each *response*

- a group first-quartile variable for each *response*
- a group mean variable for each *response*
- a group median variable for each *response*
- a group third-quartile variable for each *response*
- a group maximum variable for each *response*
- a group standard deviation variable for each *response*
- a group sample size variable for each *response*

The names of the group summary statistics variables must be the *response* name concatenated with the following special suffix characters:

Group Summary Statistic	Suffix Character
group minimum	L
group first-quartile	1
group median	M
group mean	X
group third-quartile	3
group maximum	H
group standard deviation	S
group sample size	N

For example, consider the following statements:

```
proc anom summary=Summary;
   xchart (Weight Yieldstrength)*Batch;
run;
```

The data set Summary must include the variables Batch, WeightL, Weight1, WeightX, WeightM, Weight3, WeightH, WeightS, WeightN, YieldstrengthL, Yieldstrength1, YieldstrengthX, YieldstrengthM, Yieldstrength3, YieldstrengthH, YieldstrengthS, and YieldstrengthN. Note that if you specify a *response* name that contains 32 characters, the names of the summary variables must be formed from the first 16 characters and the last 15 characters of the *response* name, suffixed with the appropriate character.

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

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

By default, the ANOM procedure reads all of the observations in a SUMMARY= 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.

For an example of a SUMMARY= data set, see "Creating ANOM Boxcharts from Group Summary Data" on page 47.

TABLE= Data Set

You can read summary statistics and decision limits from a TABLE= data set specified in the PROC ANOM statement. This enables you to reuse an OUTTABLE= data set created in a previous run of the ANOM procedure. Because the ANOM procedure simply displays the information in a TABLE= data set, you can use TABLE= data sets to create specialized ANOM charts.

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

Table 4.8 Variables Required in a TABLE= Data Set

Variable	Description
group-variable	values of the *group-variable*
LDLX	lower decision limit for mean
LIMITN	nominal sample size associated with the decision limits
MEAN	central line
SUBMAX	group maximum
SUBMED	group median
SUBMIN	group minimum
SUBN	group sample size
SUBQ1	group first quartile
SUBQ3	group third quartile
SUBX	group mean
UDLX	upper decision 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 whose length is no greater than 48.
- _VAR_. This variable is required if more than one *response* is specified or if the data set contains information for more than one *response*. This variable must be a character variable whose length is no greater than 32.

For an example of a TABLE= data set, see "Saving Decision Limits" on page 51.

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	*group-variable*
Vertical	DATA=	*response*
Vertical	SUMMARY=	group mean variable
Vertical	TABLE=	_SUBX_

Missing Values

An observation read from a DATA=, SUMMARY=, or TABLE= data set is not analyzed if the value of the group variable is missing. For a particular response variable, an observation read from a DATA= data set is not analyzed if the value of the response variable is missing. Missing values of response variables generally lead to unequal group sample sizes. For a particular response variable, an observation read from a SUMMARY= or TABLE= data set is not analyzed if the values of any of the corresponding summary variables are missing.

Examples: BOXCHART Statement

This section provides an advanced example of the BOXCHART statement.

Example 4.1: ANOM Boxcharts with Unequal Group Sizes

NOTE: See *ANOM BOXCHARTS With Unequal Group Sizes* in the SAS/QC Sample Library.

Consider the example described in "Creating ANOM Boxcharts from Response Values" on page 45. Suppose that four of the 10 measurements were missing for the third and fourth labeler positions. The following statements create a SAS data set named LabelDev2, which contains the resulting deviation measurements:

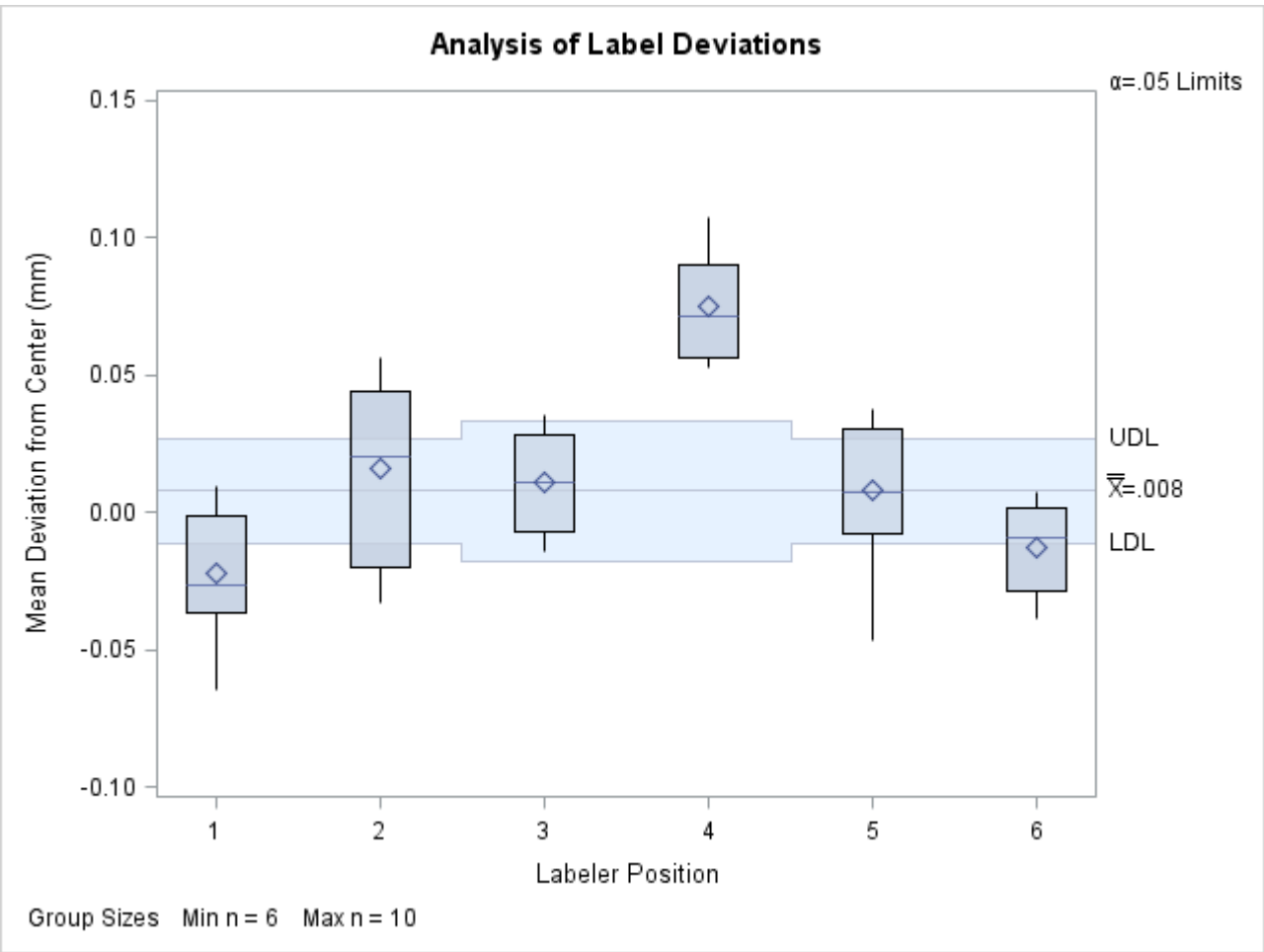
```
data LabelDev2;
   input Position @;
   do i = 1 to 5;
      input Deviation @;
      output;
   end;
   drop i;
   datalines;
1  -0.0239  -0.0285  -0.0300  -0.0043  -0.0362
1  -0.0422  -0.0014  -0.0647   0.0094  -0.0016
2  -0.0201  -0.0273   0.0227  -0.0332   0.0366
2   0.0438   0.0556   0.0098   0.0564   0.0182
3  -0.0073   0.0285    .        .      -0.0139
3    .       0.0357   0.0235    .      -0.0020
4   0.0669   0.1073    .        .       0.0755
4    .       0.0561   0.0899    .       0.0530
5   0.0368   0.0036   0.0374   0.0116  -0.0074
```

```
5   0.0250  -0.0080   0.0302  -0.0015  -0.0464
6   0.0049  -0.0384  -0.0204  -0.0049  -0.0120
6   0.0071  -0.0308   0.0017  -0.0285  -0.0070
;
```

The following statements create the ANOM chart shown in Output 4.1.1:

```
ods graphics on;
title 'Analysis of Label Deviations';
proc anom data=LabelDev2;
   boxchart Deviation*Position / odstitle=title;
   label Deviation = 'Mean Deviation from Center (mm)';
   label Position  = 'Labeler Position';
run;
```

Output 4.1.1 ANOM Chart with Unequal Group Sizes



Note that the decision limits are automatically adjusted for the varying group sizes. The legend reports the minimum and maximum group sizes.

PCHART Statement: ANOM Procedure

Overview: PCHART Statement

The PCHART statement creates ANOM charts for group (treatment level) proportions, also referred to as ANOM *p charts*.

You can use options in the PCHART statement to

- compute decision limits from the data based on specified parameters, such as the significance level (α)
- tabulate group sample sizes, group proportions, decision limits, and other information
- save decision limits in an output data set
- save group sample sizes and group proportions in an output data set
- read decision limits and decision limit parameters from a data set
- display distinct sets of decision limits for different sets of groups on the same chart
- add block legends and symbol markers to identify special groups
- 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

You have two alternatives for producing ANOM *p* charts with the PCHART statement:

- ODS Graphics output is produced if ODS Graphics is enabled, for example by specifying the ODS GRAPHICS ON statement prior to the PROC statement.
- Otherwise, traditional graphics are produced if SAS/GRAPH® is licensed.

See Chapter 3, "SAS/QC Graphics," for more information about producing these different kinds of graphs.

Getting Started: PCHART Statement

This section introduces the PCHART statement with simple examples that illustrate commonly used options. Complete syntax for the PCHART statement is presented in the section "Syntax: PCHART Statement" on page 84.

Creating ANOM Charts for Proportions from Group Counts

NOTE: See *Creating ANOM p Charts from Group Counts* in the SAS/QC Sample Library.

A health care system administrator uses ANOM to compare cesarean section rates for a set of medical groups. For more background concerning this application, refer to Rodriguez (1996).

The following statements create a SAS data set named Csection, which contains the number of c-sections and the total number of deliveries for each medical group over a one-year period.

```
data Csection;
   length ID $ 2;
   input ID Csections Total @@;
   label ID = 'Medical Group Identification Number';
   datalines;
1A  150  923 1K   45  298 1B   34  170 1D   18  132
3I   20  106 3M   12  105 1E   10   77 1N   19   74
1Q    7   69 3H   11   65 1R   11   49 1H    9   48
3J    7   20 1C    8   43 3B    6   43 1M    4   29
3C    5   28 1O    4   27 1J    6   22 1T    3   22
3E    4   18 1G    4   15 3D    4   13 3G    1   11
1L    2   10 1I    1    8 1P    0    3 1F    0    3
1S    1    3
;
```

A partial listing of Csection is shown in Figure 4.10.

Figure 4.10 The Data Set Csection

Cesarean Section Data

ID	Csections	Total
1A	150	923
1K	45	298
1B	34	170
1D	18	132
3I	20	106
3M	12	105
1E	10	77
1N	19	74
1Q	7	69
3H	11	65

The variable ID identifies the medical groups and is referred to as the *group-variable*. The variable Csections provides the number of c-sections, and is referred to as the *response variable* (or *response* for short). The variable Total provides the total number of deliveries.

The following statements create the *p* chart shown in Figure 4.11:

```
ods graphics off;
title 'Analysis of C-Sections';
proc anom data=Csection;
   pchart Csections*ID / groupn   = Total
                         hoffset  = 2
                         nolegend
                         turnhlabels;
   label Csections = 'Proportion of Cesarean Sections';
run;
```

This example illustrates the basic form of the PCHART statement. After the keyword PCHART, you specify the *response* to analyze (in this case, Csections, followed by an asterisk and the *group-variable* ID.

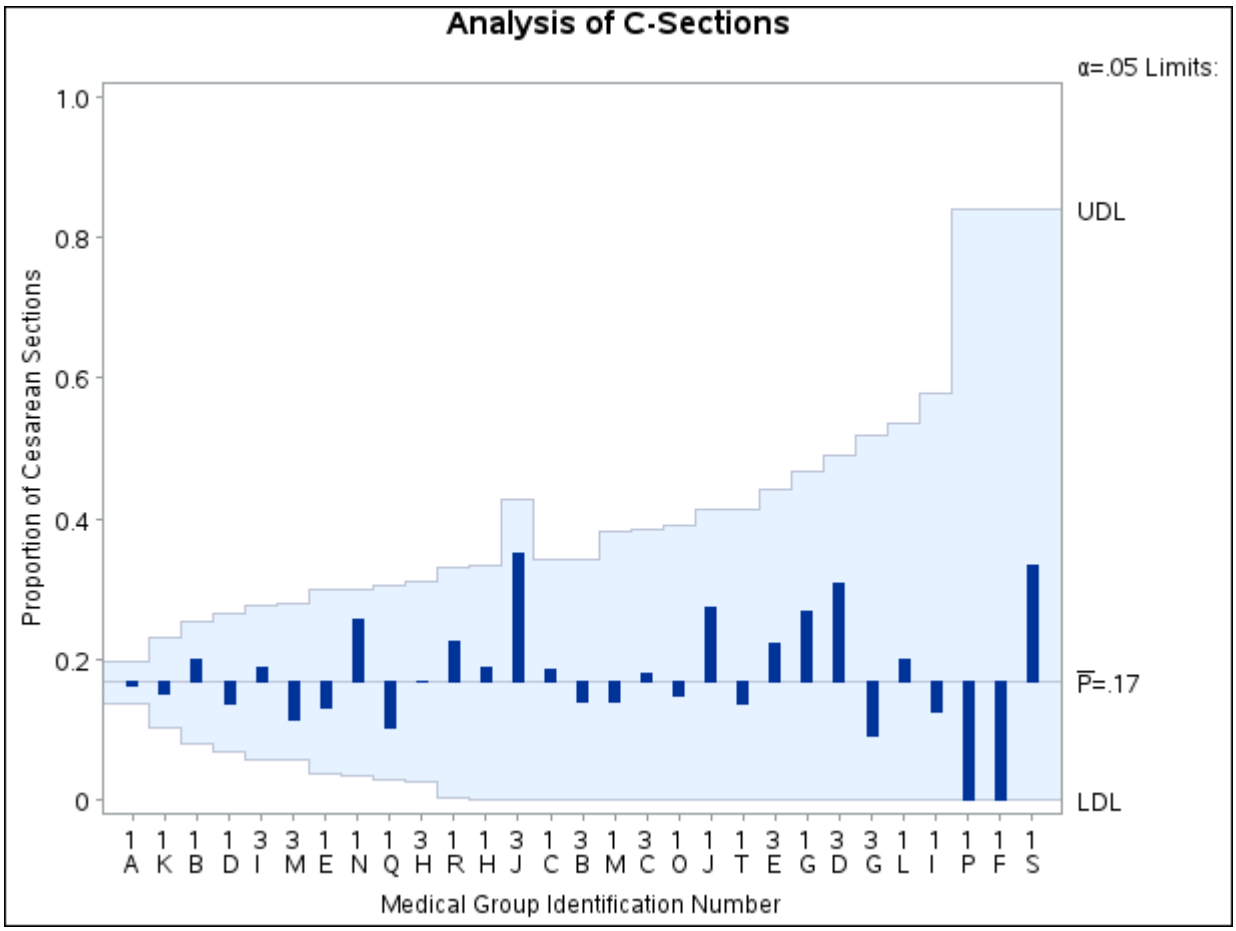
The input data set is specified with the DATA= option in the PROC ANOM statement. The GROUPN= option specifies the sample size in each group and is required with a DATA= input data set. The GROUPN= option specifies one of the following:

- a constant group sample size
- a variable in the input data set whose values provide the group sample sizes (in this case, Total)

The TURNHLABELS option turns the horizontal axis labels since the default labeling skips labels if the characters exceed the space allotted. See Axis and Axis Label Options. To angle the axis labels, see Example 4.2.

Options such as GROUPN= and TURNHLABELS are specified after the slash (/) in the PCHART statement. A complete list of options is presented in the section "Syntax: PCHART Statement" on page 84.

Figure 4.11 ANOM p Chart for Cesarean Sections



Each point on the p chart represents the proportion of c-sections for a particular group. For instance, the value plotted for group 1A is $150/923 = 0.163$.

Since all the points fall within the decision limits, it can be concluded that the variation in proportions of c-sections across medical groups is strictly due to chance.

By default, the decision limits shown correspond to a significance level of $\alpha = 0.05$. This means that, assuming all groups have the same proportion of c-sections, there is a 0.05 probability that one or more of the decision limits would be exceeded purely by chance. The formulas for the limits are given in "Decision Limits" on page 93. Note that the decision limits vary with the number of deliveries in each group, and the widest limits correspond to the group with the smallest number of deliveries.

For more details on reading group counts, see "DATA= Data Set" on page 98.

Creating ANOM Charts for Proportions from Group Summary Data

NOTE: See *Creating ANOM p Charts from Group Summary Data* in the SAS/QC Sample Library.

The previous example illustrates how you can create ANOM charts for proportions using count data. However, in many applications, the group data are provided in summarized form as proportions or percentages. This example illustrates how you can use the PCHART statement with data of this type.

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

```
data CsectProp;
   length ID $ 2;
   input ID CsectionsP CsectionsN @@;
   datalines;
1A  0.163  923   1K  0.151  298   1B  0.200  170   1D  0.136  132
3I  0.189  106   3M  0.114  105   1E  0.130   77   1N  0.257   74
1Q  0.101   69   3H  0.169   65   1R  0.224   49   1H  0.188   48
3J  0.350   20   1C  0.186   43   3B  0.140   43   1M  0.138   29
3C  0.179   28   1O  0.148   27   1J  0.273   22   1T  0.136   22
3E  0.222   18   1G  0.267   15   3D  0.308   13   3G  0.091   11
1L  0.200   10   1I  0.125    8   1P  0.000    3   1F  0.000    3
1S  0.333    3
;
```

A partial listing of CsectProp is shown in Figure 4.12. The groups are still indexed by ID. The variable CsectionsP contains the proportions of c-sections, and the variable CsectionsN contains the group sample sizes.

Figure 4.12 The Data Set CsectProp

Proportions of Cesarean Sections

ID	CsectionsP	CsectionsN
1A	0.163	923
1K	0.151	298
1B	0.200	170
1D	0.136	132
3I	0.189	106
3M	0.114	105
1E	0.130	77
1N	0.257	74
1Q	0.101	69
3H	0.169	65

You can analyze this data set by specifying it as a SUMMARY= data set in the PROC ANOM statement.

Note that Csections is *not* the name of a SAS variable in the data set but is, instead, the common prefix for the names of the two SAS variables CsectionsP and CsectionsN. The suffix characters *P* and *N* indicate *proportion* and *sample size*, respectively. Thus, you can specify two group variables in a SUMMARY= data set with a single name Csections, which is referred to as the *response*. The name ID specified after the asterisk is the name of the *group-variable*.

A SUMMARY= data set used with the PCHART statement must contain the following variables:

- group variable
- group proportion variable
- group sample size variable

Furthermore, the names of the group proportion and sample size variables must begin with the *response* name specified in the PCHART statement and end with the special suffix characters *P* and *N*, respectively.

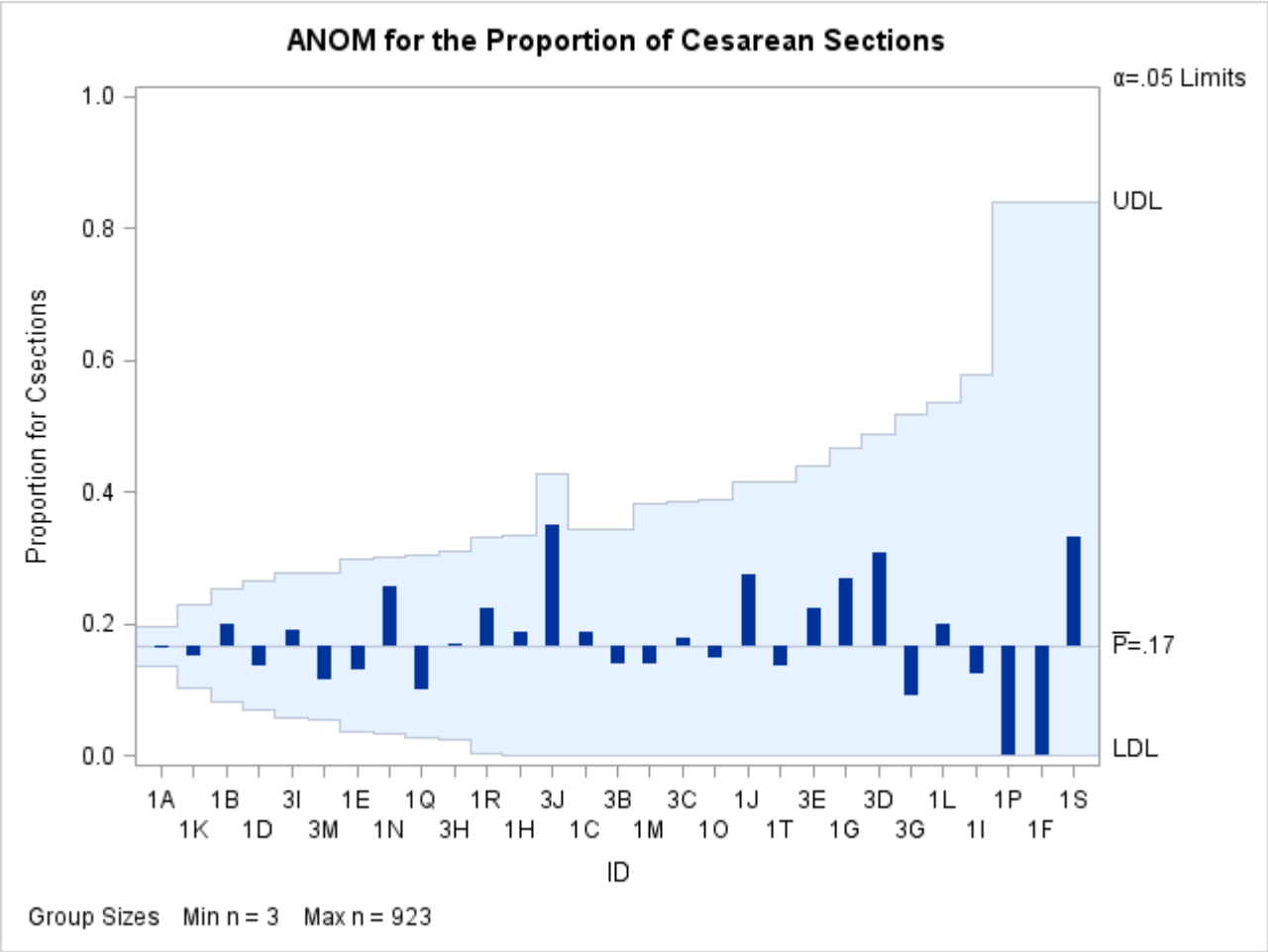
For more information, see "SUMMARY= Data Set" on page 99.

The following statements create a *p* Chart for C-Sections using the SUMMARY= data set CsectProp:

```
ods graphics on;
title 'ANOM for the Proportion of Cesarean Sections';
proc anom summary=CsectProp;
   pchart Csections*ID / odstitle = title1;
run;
```

The ODS GRAPHICS ON statement specified before the PROC ANOM statement enables ODS Graphics, so the *p* chart is created using ODS Graphics instead of traditional graphics. The resulting ANOM *p* chart is shown in Figure 4.13.

Figure 4.13 ANOM *p* Chart from Group Proportions



Saving Group Proportions

NOTE: See *Saving Group Proportions Using ANOM PCHART* in the SAS/QC Sample Library.

In this example, the PCHART statement is used to create a summary data set that can later be read by the ANOM procedure (as in the preceding example). The following statements read the data set CSection (see

"Creating ANOM Charts for Proportions from Group Counts" on page 77) and create a summary data set named CSummary:

```
proc anom data=Csection;
   pchart Csections*ID / groupn     = Total
                         outsummary = CSummary
                         nochart;
run;
```

The OUTSUMMARY= 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 4.11. Figure 4.14 contains a partial listing of CSummary.

Figure 4.14 The Data Set CSummary

Group Proportions and Decision Limit Information

ID	CsectionsP	CsectionsN
1A	0.16251	923
1K	0.15101	298
1B	0.20000	170
1D	0.13636	132
3I	0.18868	106
3M	0.11429	105
1E	0.12987	77
1N	0.25676	74
1Q	0.10145	69
3H	0.16923	65

There are three variables in the data set CSummary:

- ID identifies the groups.
- CSectionsP contains the group proportions.
- CSectionsN contains the group sample sizes.

Note that the variables containing the group proportions and group sample sizes are named by adding the suffix characters *P* and *N* to the *response* CSections specified in the PCHART statement. In other words, the variable naming convention for OUTSUMMARY= data sets is the same as that for SUMMARY= data sets. For more information, see "OUTSUMMARY= Data Set" on page 95.

Saving Decision Limits

NOTE: See *Saving Decision Limits Using ANOM PCHART* in the SAS/QC Sample Library.

You can save the decision limits for an ANOM *p* chart in a SAS data set.

The following statements read the number of c-sections per group from the data set CSection (see "Creating ANOM Charts for Proportions from Group Counts" on page 77) and save the decision limits displayed in Figure 4.11 in a data set named CSectionLim:

```
proc anom data=Csection;
   pchart Csections*ID / groupn    = Total
                         outlimits = CsectionLim
                         nochart;
run;
```

The OUTLIMITS= option names the data set containing the decision limits, and the NOCHART option suppresses the display of the chart. The data set CSectionLim is listed in Figure 4.15.

Figure 4.15 The Data Set CSectionLim with Decision Limits

Decision Limits for the Proportion of Cesarean Sections

VAR	_GROUP_	_TYPE_	_LIMITN_	_ALPHA_	_LDLP_	_P_	_UDLP_	_LIMITK_
Csections	ID	ESTIMATE	V	0.05	V 0.16680		V	29

The data set CSectionLim contains one observation with the limits for the *response* CSections. The variables _LDLP_ and _UDLP_ contain the lower and upper decision limits, and the variable _P_ contains the central line. The value of _LIMITN_ is the nominal sample size associated with the decision limits, the value of _LIMITK_ is the number of groups, and the value of _ALPHA_ is the significance level associated with the decision limits. The variables _VAR_ and _GROUP_ are bookkeeping variables that save the *response* and *group-variable*. The variable _TYPE_ is a bookkeeping variable that indicates whether the value of _P_ is an estimate or a known (standard) value. Typically, the value of _TYPE_ is 'ESTIMATE.'

For more information, see the section "OUTLIMITS= Data Set" on page 95.

NOTE: See *Saving ANOM PCHART Summary Statistics and Decision Limits* in the SAS/QC Sample Library.

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

```
proc anom data=Csection;
   pchart Csections*ID / groupn   = Total
                         outtable = CsectionTab
                         nochart;
run;
```

A partial listing of the data set CSectionTab is shown in Figure 4.16.

Figure 4.16 The Data Set CSectionTab

Proportions and Decision Limits for Cesarean Sections

VAR	ID	_ALPHA_	_LIMITN_	_SUBN_	_LDLP_	_SUBP_	_P_	_UDLP_	_EXLIM_
Csections	1A	0.05	923	923	0.13658	0.16251	0.16680	0.19703	
Csections	1K	0.05	298	298	0.10355	0.15101	0.16680	0.23006	
Csections	1B	0.05	170	170	0.08059	0.20000	0.16680	0.25302	
Csections	1D	0.05	132	132	0.06814	0.13636	0.16680	0.26547	
Csections	3I	0.05	106	106	0.05608	0.18868	0.16680	0.27752	
Csections	3M	0.05	105	105	0.05553	0.11429	0.16680	0.27807	
Csections	1E	0.05	77	77	0.03609	0.12987	0.16680	0.29752	
Csections	1N	0.05	74	74	0.03338	0.25676	0.16680	0.30023	
Csections	1Q	0.05	69	69	0.02849	0.10145	0.16680	0.30512	
Csections	3H	0.05	65	65	0.02417	0.16923	0.16680	0.30943	

This data set contains one observation for each group sample. The variables _SUBP_ and _SUBN_ contain the group proportions and group sample sizes. The variables _LDLP_ and _UDLP_ contain the lower and upper decision limits, and the variable _P_ contains the central line. The variables _VAR_ and ID contain the *response* name and values of the *group-variable*, respectively. For more information, see "OUTTABLE= Data Set" on page 96.

An OUTTABLE= data set can be read later as a TABLE= data set. For example, the following statements read the information in CSectionTab and display an ANOM *p* chart (not shown here) identical to the chart in Figure 4.11:

```
title 'Analysis of C-Sections';
proc anom table=CSectionTab;
   pchart CSections*id;
label _subp_ = 'Proportion of Cesarean Sections';
run;
```

Because the ANOM procedure simply displays the information in a TABLE= data set, you can use TABLE= data sets to create specialized ANOM charts. For more information, see "TABLE= Data Set" on page 100.

Syntax: PCHART Statement

The basic syntax for the PCHART statement is as follows:

PCHART *response* ∗ *group-variable* **;**

The general form of this syntax is as follows:

PCHART *responses* ∗ *group-variable* < **(***block-variables***)** >
< **=***symbol-variable* | **=***'character'* > < *options* > **;**

You can use any number of PCHART statements in the ANOM procedure. The components of the PCHART statement are described as follows.

response
responses

identify one or more responses to be analyzed. The specification of *response* depends on the input data set specified in the PROC ANOM statement.

- If response counts are read from a DATA= data set, *response* must be the name of the variable containing the counts. For an example, see "Creating ANOM Charts for Proportions from Group Summary Data" on page 79.
- If response proportions are read from a SUMMARY= data set, *response* must be the common prefix of the summary variables in the SUMMARY= data set. For an example, see "Creating ANOM Charts for Proportions from Group Summary Data" on page 79.
- If response proportions and decision limits are read from a TABLE= data set, *response* must be the value of the variable _VAR_ in the TABLE= data set. For an example, see "Saving Decision Limits" on page 83.

A *response* is required. If you specify more than one response, enclose the list in parentheses. For example, the following statements request distinct ANOM *p* charts for the responses Rejects and Reworks:

```
proc anom data=Measures;
   pchart (Rejects Reworks)*Sample / groupn=100;
run;
```

Note that when data are read from a DATA= data set, the GROUPN= option, which specifies group sample sizes, is required.

group-variable

is the variable that identifies groups in the data. The *group-variable* is required. In the preceding PCHART statement, Sample is the group variable.

block-variables

are optional variables that identify sets of consecutive groups on the chart. The blocks are labeled in a legend, and each *block-variable* provides one level of labels in the legend.

symbol-variable

is an optional variable whose levels (unique values) determine the symbol marker used to plot proportions. 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.

options

control the analysis, enhance the appearance of the chart, save results in data sets, and so on. The section "Summary of Options" lists all options by function.

Summary of Options

The following tables list the PCHART statement options by function. Options unique to the ANOM procedure are listed in Table 4.9, and are described in detail in "Dictionary of ANOM Chart Statement Options" on page 180. Options that are common to both the ANOM and SHEWHART procedures are listed in Table 4.10, and are described in detail in "Dictionary of Options: SHEWHART Procedure" on page 1946.

Table 4.9 PCHART Statement Special Options

Option	Description
Options for Specifying Decision Limits	
ALPHA=	specifies significance level
LIMITK=	specifies number of groups for decision limits
LIMITN=	specifies either nominal sample size for fixed decision limits or varying limits
NOREADLIMITS	computes decision limits for each *response* from the data rather than a LIMITS= data set
P=	specifies the weighted average of group proportions
READINDEXES=	reads multiple sets of decision limits for each *response* from a LIMITS= data set
TYPE=	identifies parameters as estimates or standard values and specifies value of _TYPE_ in the OUTLIMITS= data set
Options for Displaying Decision Limits	
CINFILL=	specifies color for area inside decision limits
CLIMITS=	specifies color of decision limits, central line, and related labels
LDLLABEL=	specifies label for lower decision limit
LIMLABSUBCHAR=	specifies a substitution character for labels provided as quoted strings; the character is replaced with the value of the decision limit
LLIMITS=	specifies line type for decision limits
NDECIMAL=	specifies number of digits to right of decimal place in default labels for decision limits and central line
NOCTL	suppresses display of central line
NOLDL	suppresses display of lower decision limit
NOLIMIT0	suppresses display of lower decision limit if it is 0
NOLIMIT1	suppresses display of upper decision limit if it is 1 (100%)
NOLIMITLABEL	suppresses labels for decision limits and central line
NOLIMITS	suppresses display of decision limits
NOLIMITSFRAME	suppresses default frame around decision limit information when multiple sets of decision limits are read from a LIMITS= data set
NOLIMITSLEGEND	suppresses legend for decision limits
NOUDL	suppresses display of upper decision limit
PSYMBOL=	specifies label for central line
UDLLABEL=	specifies label for upper decision limit
WLIMITS=	specifies width for decision limits and central line
Input Data Set Option	
GROUPN=	specifies group sample sizes as constant number *n* or as values of variable in a DATA= data set
Output Data Set Option	
OUTSUMMARY=	creates output data set containing group summary statistics

Table 4.10 PCHART Statement General Options

Option	Description
Options for Plotting and Labeling Points	
ALLLABEL=	labels every point on ANOM *p* chart
CLABEL=	specifies color for labels
CCONNECT=	specifies color for line segments that connect points on chart
CFRAMELAB=	specifies fill color for frame around labeled points
CNEEDLES=	specifies color for needles that connect points to central line
COUT=	specifies color for portions of line segments that connect points outside decision limits
COUTFILL=	specifies color for shading areas between the connected points and decision limits outside the limits
LABELANGLE=	specifies angle at which labels are drawn
LABELFONT=	specifies software font for labels
LABELHEIGHT=	specifies height of labels
NONEEDLES	suppresses vertical needles connecting points to central line
OUTLABEL=	labels points outside decision limits
SYMBOLLEGEND=	specifies LEGEND statement for levels of *symbol-variable*
SYMBOLORDER=	specifies order in which symbols are assigned for levels of *symbol-variable*
TURNALL|TURNOUT	turns point labels so that they are strung out vertically
WNEEDLES=	specifies width of needles
Axis and Axis Label Options	
CAXIS=	specifies color for axis lines and tick marks
CFRAME=	specifies fill colors for frame for plot area
CTEXT=	specifies color for tick mark values and axis labels
DISCRETE	produces horizontal axis for discrete numeric group values
HAXIS=	specifies major tick mark values for horizontal axis
HEIGHT=	specifies height of axis label and axis legend text
HMINOR=	specifies number of minor tick marks between major tick marks on horizontal axis
HOFFSET=	specifies length of offset at both ends of horizontal axis
INTSTART=	specifies first major tick mark value for numeric horizontal axis with date, time, or datetime format
NOHLABEL	suppresses label for horizontal axis
NOTICKREP	specifies that only the first occurrence of repeated, adjacent group values is to be labeled on horizontal axis
NOTRUNC	suppresses vertical axis truncation at zero applied by default
NOVANGLE	requests vertical axis labels that are strung out vertically

Table 4.10 *continued*

Option	Description
SKIPHLABELS=	specifies thinning factor for tick mark labels on horizontal axis
TURNHLABELS	requests horizontal axis labels that are strung out vertically
VAXIS=	specifies major tick mark values for vertical axis of ANOM *p* chart
VFORMAT=	specifies format for vertical axis tick mark labels
VMINOR=	specifies number of minor tick marks between major tick marks on vertical axis
VOFFSET=	specifies length of offset at both ends of vertical axis
VZERO	forces origin to be included in vertical axis for ANOM *p* chart
WAXIS=	specifies width of axis lines
YSCALE=	scales vertical axis in percent units (rather than proportions)
Plot Layout Options	
ALLN	plots means for all groups
BILEVEL	creates ANOM *p* chart using half-screens and half-pages
EXCHART	creates ANOM *p* chart for a response only when a group mean exceeds the decision limits
INTERVAL=	natural time interval between consecutive group positions when time, date, or datetime format is associated with a numeric group variable
MAXPANELS=	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 decision limits
NOCHART	suppresses creation of chart
NOFRAME	suppresses frame for plot area
NOLEGEND	suppresses legend for group sample sizes
NPANELPOS=	specifies number of group positions per panel on each chart
REPEAT	repeats last group position on panel as first group position of next panel
TOTPANELS=	specifies number of pages or screens to be used to display chart
ZEROSTD	displays ANOM *p* chart regardless of whether root mean square error is zero
Reference Line Options	
CHREF=	specifies color for lines requested by HREF= option
CVREF=	specifies color for lines requested by VREF= option

Table 4.10 *continued*

Option	Description
HREF=	specifies position of reference lines perpendicular to horizontal axis on ANOM *p* chart
HREFDATA=	specifies position of reference lines perpendicular to horizontal axis on ANOM *p* chart
HREFLABELS=	specifies labels for HREF= lines
HREFLABPOS=	specifies position of HREFLABELS= labels
LHREF=	specifies line type for HREF= lines
LVREF=	specifies line type for VREF= lines
NOBYREF	specifies that reference line information in a data set applies uniformly to charts created for all BY groups
VREF=	specifies position of reference lines perpendicular to vertical axis on ANOM *p* chart
VREFLABELS=	specifies labels for VREF= lines
VREFLABPOS=	specifies position of VREFLABELS= labels
Grid Options	
CGRID=	specifies color for grid requested with GRID or ENDGRID option
ENDGRID	adds grid after last plotted point
GRID	adds grid to control chart
LENDGRID=	specifies line type for grid requested with the ENDGRID option
LGRID=	specifies line type for grid requested with the GRID option
WGRID=	specifies width of grid lines
Clipping Options	
CCLIP=	specifies color for plot symbol for clipped points
CLIPFACTOR=	determines extent to which extreme points are clipped
CLIPLEGEND=	specifies text for clipping legend
CLIPLEGPOS=	specifies position of clipping legend
CLIPSUBCHAR=	specifies substitution character for CLIPLEGEND= text
CLIPSYMBOL=	specifies plot symbol for clipped points
CLIPSYMBOLHT=	specifies symbol marker height for clipped points
Graphical Enhancement Options	
ANNOTATE=	specifies annotate data set that adds features to ANOM *p* chart
DESCRIPTION=	specifies description of ANOM *p* chart's GRSEG catalog entry
FONT=	specifies software font for labels and legends on chart
NAME=	specifies name of ANOM *p* chart's GRSEG catalog entry
PAGENUM=	specifies the form of the label used in pagination

Table 4.10 *continued*

Option	Description
PAGENUMPOS=	specifies the position of the page number requested with the PAGENUM= option
Options for Producing Graphs Using ODS Styles	
BLOCKVAR=	specifies one or more variables whose values define colors for filling background of *block-variable* legend
CFRAMELAB	draws a frame around labeled points
COUT	draw portions of line segments that connect points outside decision limits in a contrasting color
CSTAROUT	specifies that portions of stars exceeding inner or outer circles are drawn using a different color
OUTFILL	shades areas between decision limits and connected points lying outside the limits
STARFILL=	specifies a variable identifying groups of stars filled with different colors
STARS=	specifies a variable identifying groups of stars whose outlines are drawn with different colors
Options for ODS Graphics	
BLOCKREFTRANSPARENCY=	specifies the wall fill transparency for blocks and phases
INFILLTRANSPARENCY=	specifies the decision limit infill transparency
MARKERS	plots group points with markers
NOBLOCKREF	suppresses block and phase reference lines
NOBLOCKREFFILL	suppresses block and phase wall fills
NOFILLLEGEND	suppresses legend for levels of a STARFILL= variable
NOPHASEREF	suppresses block and phase reference lines
NOPHASEREFFILL	suppresses block and phase wall fills
NOREF	suppresses block and phase reference lines
NOREFFILL	suppresses block and phase wall fills
NOSTARFILLLEGEND	suppresses legend for levels of a STARFILL= variable
NOTRANSPARENCY	disables transparency in ODS Graphics output
ODSFOOTNOTE=	specifies a graph footnote
ODSLEGENDEXPAND	specifies that legend entries contain all levels observed in the data
ODSTITLE=	specifies a graph title
OUTFILLTRANSPARENCY=	specifies decision limit outfill transparency
OVERLAYURL=	specifies URLs to associate with overlay points
PHASEPOS=	specifies vertical position of phase legend
PHASEREFLEVEL=	associates phase and block reference lines with either innermost or the outermost level
PHASEREFTRANSPARENCY=	specifies the wall fill transparency for blocks and phases
REFFILLTRANSPARENCY=	specifies the wall fill transparency for blocks and phases

Table 4.10 *continued*

Option	Description
SIMULATEQCFONT	draws central line labels using a simulated software font
STARTRANSPARENCY=	specifies star fill transparency
URL=	specifies a variable whose values are URLs to be associated with groups
Input Data Set Option	
DATAUNIT=	specifies that input values are proportions or percentages rather than counts
Output Data Set Options	
OUTINDEX=	specifies value of _INDEX_ in the OUTLIMITS= data set
OUTLIMITS=	creates output data set containing decision limits
OUTTABLE=	creates output data set containing group summary statistics and decision limits
Tabulation Options	
NOTE: specifying (EXCEPTIONS) after a tabulation option creates a table for exceptional points only.	
TABLE	creates a basic table of group means, group sample sizes, and decision 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
TABLEOUTLIM	augments basic table with columns indicating decision limits exceeded
Block Variable Legend Options	
BLOCKLABELPOS=	specifies position of label for *block-variable* legend
BLOCKLABTYPE=	specifies text size of *block-variable* legend
BLOCKPOS=	specifies vertical position of *block-variable* legend
BLOCKREP	repeats identical consecutive labels in *block-variable* legend
CBLOCKLAB=	specifies fill colors for frames enclosing variable labels in *block-variable* legend
CBLOCKVAR=	specifies one or more variables whose values are colors for filling background of *block-variable* legend
Phase Options	
CPHASELEG=	specifies text color for *phase* legend
NOPHASEFRAME	suppresses default frame for *phase* legend
OUTPHASE=	specifies value of _PHASE_ in the OUTHISTORY= data set
PHASEBREAK	disconnects last point in a *phase* from first point in next *phase*

Table 4.10 *continued*

Option	Description
PHASELABTYPE=	specifies text size of *phase* legend
PHASELEGEND	displays *phase* labels in a legend across top of chart
PHASELIMITS	labels decision limits for each phase, provided they are constant within that phase
PHASEREF	delineates *phases* with vertical reference lines
READPHASES=	specifies *phases* to be read from an input data set
Star Options	
CSTARCIRCLES=	specifies color for STARCIRCLES= circles
CSTARFILL=	specifies color for filling stars
CSTAROUT=	specifies outline color for stars exceeding inner or outer circles
CSTARS=	specifies color for outlines of stars
LSTARCIRCLES=	specifies line types for STARCIRCLES= circles
LSTARS=	specifies line types for outlines of STARVERTICES= stars
STARBDRADIUS=	specifies radius of outer bound circle for vertices of stars
STARCIRCLES=	specifies reference circles for stars
STARINRADIUS=	specifies inner radius of stars
STARLABEL=	specifies vertices to be labeled
STARLEGEND=	specifies style of legend for star vertices
STARLEGENDLAB=	specifies label for STARLEGEND= legend
STAROUTRADIUS=	specifies outer radius of stars
STARSPECS=	specifies method used to standardize vertex variables
STARSTART=	specifies angle for first vertex
STARTYPE=	specifies graphical style of star
STARVERTICES=	superimposes star at each point on ANOM *p* chart
WSTARCIRCLES=	specifies width of STARCIRCLES= circles
WSTARS=	specifies width of STARVERTICES= stars
Overlay Options	
CCOVERLAY=	specifies colors for overlay line segments
COVERLAY=	specifies colors for overlay plots
COVERLAYCLIP=	specifies color for clipped points on overlays
LOVERLAY=	specifies line types for overlay line segments
NOOVERLAYLEGEND	suppresses legend for overlay plots
OVERLAY=	specifies variables to overlay on chart
OVERLAYCLIPSYM=	specifies symbol for clipped points on overlays
OVERLAYCLIPSYMHT=	specifies symbol height for clipped points on overlays
OVERLAYHTML=	specifies links to associate with overlay points
OVERLAYID=	specifies labels for overlay points
OVERLAYLEGLAB=	specifies label for overlay legend
OVERLAYSYM=	specifies symbols for overlays
OVERLAYSYMHT=	specifies symbol heights for overlays

Table 4.10 *continued*

Option	Description
WOVERLAY=	specifies widths of overlay line segments
Options for Interactive ANOM Charts	
HTML=	specifies a variable whose values create links to be associated with groups
HTML_LEGEND=	specifies a variable whose values create links to be associated with symbols in the symbol legend
WEBOUT=	creates an OUTTABLE= data set with additional graphics coordinate data

Details: PCHART Statement

Constructing ANOM Charts for Proportions

The following notation is used in this section:

X_i	response number (count) in the ith group
k	number of groups
n_i	sample size of the ith group
N	total sample size $= n_1 + \cdots + n_k$
p_i	proportion in the ith group, where $p_i = X_i/n_i$
$\bar{p}$	weighted average of proportions across groups: $$\bar{p} = \frac{n_1 p_1 + \cdots + n_k p_k}{N} = \frac{X_1 + \cdots + X_k}{N}$$
α	significance level
$h(\alpha; k, n, \infty)$	critical value for ANOM for normal data in the balanced case ($n_i \equiv n$)
$h(\alpha; k, n_1, \ldots, n_k, \infty)$	critical value for ANOM for normal data in the unbalanced case

Plotted Points

Each point on an ANOM p chart represents the response proportion ($p_i = X_i/n_i$) for a group.

Central Line

By default, the central line on an ANOM p chart is computed as $\bar{p}$, the weighted average of the group proportions. You can specify $\bar{p}$ with the P= option or with the variable _P_ in a LIMITS= data set.

Decision Limits

For the ith group, the response counts are assumed to have the binomial distribution $B(n_i, p_i)$. The ANOM method for proportions tests the null hypothesis that $p_1 = p_2 = \cdots = p_k$, that is, that the proportions are the same, against the alternative that at least one of the p_i's is different from the average of the k proportions.

The decision limits are computed using the normal approximation to the binomial distribution, which is appropriate when the sample sizes for the groups are large; refer to Ramig (1983). A commonly recommended check for this assumption is that $n_i p_i > 5$ and $n_i(1 - p_i) > 5$ for all the groups. The critical values in the ANOM method for normally distributed data are adapted to the binomial case by using infinite degrees of freedom for the variance.

When the sample sizes are constant across groups ($n_i \equiv n$), the decision limits are computed as follows:

$$\begin{aligned}
\text{lower decision limit (LDL)} &= \max\left(\bar{p} - h(\alpha; k, n, \infty)\sqrt{\bar{p}(1-\bar{p})}\sqrt{\frac{k-1}{N}}, 0\right) \\
\text{upper decision limit (UDL)} &= \min\left(\bar{p} + h(\alpha; k, n, \infty)\sqrt{\bar{p}(1-\bar{p})}\sqrt{\frac{k-1}{N}}, 1\right)
\end{aligned}$$

For the theoretical derivation of the decision limits, refer to Nelson (1982a).

When the sample sizes (n_i) are different across groups (the unbalanced case), the decision limits are computed as follows:

$$\begin{aligned}
\text{lower decision limit (LDL)} &= \max\left(\bar{p} - h(\alpha; k, n_1, \ldots, n_k, \infty)\sqrt{\bar{p}(1-\bar{p})}\sqrt{\frac{N-n_i}{Nn_i}}, 0\right) \\
\text{upper decision limit (UDL)} &= \min\left(\bar{p} + h(\alpha; k, n_1, \ldots, n_k, \infty)\sqrt{\bar{p}(1-\bar{p})}\sqrt{\frac{N-n_i}{Nn_i}}, 1\right)
\end{aligned}$$

Note that the decision limits for the ith group depend on n_i. If the sample sizes are constant across groups ($n_i \equiv n$), the decision limits in the unbalanced case reduce to the formulas given for the balanced case since $n_i = n$ and $N = kn$ so

$$\sqrt{\frac{N-n_i}{Nn_i}} = \sqrt{\frac{kn-n}{Nn}} = \sqrt{\frac{k-1}{N}}$$

For the derivation of the decision limits for unequal sample sizes, refer to Nelson (1991).

Exact critical values $h(\alpha; k, n, \infty)$ were first tabulated by L. S. Nelson (1983). Refer to Nelson (1993) for derivation of critical values.

You can specify parameters for the limits as follows:

- Specify α with the ALPHA= option or with the variable _ALPHA_ in a LIMITS= data set. By default, $\alpha = 0.05$.

- Specify a constant nominal sample size $n_i \equiv n$ for the decision limits with the LIMITN= option or with the variable _LIMITN_ in a LIMITS= data set. By default, n is the observed sample size in the balanced case.
- Specify $\bar{p}$ with the P= option or with the variable _P_ in a LIMITS= data set. By default, $\bar{p}$ is the weighted average of the group proportions.

Output Data Sets

OUTLIMITS= Data Set

The OUTLIMITS= data set saves decision limits and decision limit parameters. The following variables can be saved:

Table 4.11 OUTLIMITS= Data Set

Variable	Description
ALPHA	significance level (α)
GROUP	*group-variable* specified in the PCHART statement
INDEX	optional identifier for the decision limits specified with the OUTINDEX= option
LDLP	lower decision limit for proportions
LIMITK	number of groups
LIMITN	nominal sample size associated with the decision limits
_P_	average proportion of nonconforming items ($\bar{p}$)
TYPE	type (standard or estimate) of _P_
UDLP	upper decision limit for proportions
VAR	*response* specified in the PCHART statement

Notes:

1. If the decision limits vary with group sample size, the special missing value V is assigned to the variables _LIMITN_, _LDLP_, and _UDLP_.
2. A group must have at least one nonmissing value ($n_i \geq 1$), and there must be at least one group with $n_i \geq 2$.
3. Optional BY variables are saved in the OUTLIMITS= data set.

The OUTLIMITS= data set contains one observation for each *response* specified in the PCHART statement. For an example, see "Saving Decision Limits" on page 83.

OUTSUMMARY= Data Set

The OUTSUMMARY= data set saves group summary statistics. The following variables are saved:

- the *group-variable*
- a group proportion variable named by *response* suffixed with *P*

- a group sample size variable named by *response* suffixed with *N*

Given a *response* name that contains 32 characters, the procedure first shortens the name to its first 16 characters and its last 15 characters, and then it adds the suffix.

Group summary variables are created for each *response* specified in the PCHART statement. For example, consider the following statements:

```
proc anom data=Input;
   pchart (Rework Rejected)*Batch / outsummary=Summary
                                    groupn =30;
run;
```

The data set Summary contains variables named Batch, ReworkP, ReworkN, RejetedP, and RejetedN.

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 OUTSUMMARY= data set, see "Saving Group Proportions" on page 81.

Note that an OUTSUMMARY= data set created with the PCHART statement can be reused as a SUMMARY= data set.

OUTTABLE= Data Set

The OUTTABLE= data set saves group summary statistics, decision limits, and related information. The variables shown in the following table are saved:

Variable	Description
ALPHA	significance level (α)
EXLIM	decision limit exceeded on p chart
group	values of the group variable
LDLP	lower decision limit for proportions
LIMITN	nominal sample size associated with the decision limits
SUBP	group proportion
SUBN	group sample size
UDLP	upper decision limit for proportions
VAR	*response* specified in the PCHART 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)

NOTE: The variable _EXLIM_ is a character variable of length 8. The variable _PHASE_ is a character variable of length 48. The variable _VAR_ is a character variable whose length is no greater than 32. All other variables are numeric.

For an example, see "Saving Decision Limits" on page 83.

ODS Tables

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

Table 4.12 ODS Tables Produced with the PCHART Statement

Table Name	Description	Options
PCHART	*p* chart summary statistics	TABLE, TABLEALL, TABLEC, TABLEID, TABLEOUT

ODS Graphics

Before you create ODS Graphics output, ODS Graphics must be enabled (for example, by using the ODS GRAPHICS ON statement). For more information about enabling and disabling ODS Graphics, see the section "Enabling and Disabling ODS Graphics" (Chapter 21, *SAS/STAT User's Guide*). **NOTE:** In SAS/QC 13.1 the ANOM procedure does not support the creation of graphs that are editable with the ODS Graphics Editor.

The appearance of a graph produced with ODS Graphics is determined by the style associated with the ODS destination where the graph is produced. PCHART options used to control the appearance of traditional graphics are ignored for ODS Graphics output. Options for Producing Graphs Using ODS Styles lists options that can be used to control the appearance of graphs produced with ODS Graphics or with traditional graphics using ODS styles. Options for ODS Graphics lists options to be used exclusively with ODS Graphics. Detailed descriptions of these options are provided in "Dictionary of Options: SHEWHART Procedure" on page 1946.

When ODS Graphics is in effect, the PCHART statement assigns a name to the graph it creates. You can use this name to reference the graph when using ODS. The name is listed in Table 4.13.

Table 4.13 ODS Graphics Produced by the PCHART Statement

ODS Graph Name	Plot Description
PChart	ANOM *p* chart

See Chapter 3, "SAS/QC Graphics," for more information about ODS Graphics and other methods for producing charts.

Input Data Sets

DATA= Data Set

You can read count data from a DATA= data set specified in the PROC ANOM statement. Each *response* specified in the PCHART statement must be a SAS variable in the DATA= data set. This variable provides counts for group samples indexed by the values of the *group-variable*. The *group-variable*, which is specified in the PCHART statement, must also be a SAS variable in the DATA= data set. Each observation in a DATA= data set must contain a count for each *response* and a value for the *group-variable*. The data set must contain one observation for each group. Note that you can specify the DATAUNIT= option in the PCHART statement to read proportions or percentages instead of counts. 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

When you use a DATA= data set with the PCHART statement, the GROUPN= option (which specifies the group sample size) is required.

For an example of a DATA= data set, see "Creating ANOM Charts for Proportions from Group Counts" on page 77.

LIMITS= Data Set

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

```
proc anom data=Info limits=Conlims;
   pchart Rejects*Batch / groupn= 100;
run;
```

The LIMITS= data set can be an OUTLIMITS= data set that was created in a previous run of the ANOM 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 _LDLP_, _P_, and _UDLP_, which specify the decision limits directly
- the variable _P_, without providing _LDLP_ and _UDLP_. The value of _P_ is used to calculate the decision limits according to the equations in the section "Decision Limits" on page 93.

In addition, note the following:

- The variables _VAR_ and _GROUP_ are always required. These must be character variables whose lengths are no greater than 32.

- _LDLP_ and _UDLP_ must be specified together; otherwise their values are computed.
- _ALPHA_ is optional but is recommended in order to maintain a complete set of decision limit information. The default value is 0.05.
- _LIMITK_ is optional. The default value is k, the number of groups. A group must have at least one nonmissing value ($n_i \geq 1$) and there must be at least one group with $n_i \geq 2$. If specified, _LIMITK_ overrides the value of k.
- _LIMITN_ is optional. The default value is the common group size (n), in the balanced case $n_i \equiv n$. If specified, _LIMITN_ overrides the value of n.
- The variable _TYPE_ is optional, but is recommended to maintain a complete set of decision limit information. The variable _TYPE_ must be a character variable of length 8. Valid values are 'ESTIMATE,' 'STANDARD,' 'STDMEAN,' and 'STDRMS.' The default is 'ESTIMATE.'
- The variable _INDEX_ is required if you specify the READINDEX= option; this must be a character variable whose length is no greater than 48.
- BY variables are required if specified with a BY statement.

SUMMARY= Data Set

You can read group summary statistics from a SUMMARY= data set specified in the PROC ANOM statement. This enables you to reuse OUTSUMMARY= data sets that have been created in previous runs of the ANOM procedure or to create your own SUMMARY= data set.

A SUMMARY= data set used with the PCHART statement must contain the following:

- the *group-variable*
- a group proportion variable for each *response*
- a group sample size variable for each *response*

The names of the proportion sample size variables must be the *response* name concatenated with the special suffix characters *P* and *N*, respectively.

For example, consider the following statements:

```
proc anom summary=Summary;
   pchart (Rework Rejected)*Batch / groupn=50;
run;
```

The data set Summary must include the variables Batch, ReworkP, ReworkN, RejetedP, and RejetedN.

Note that if you specify a *response* name that contains 32 characters, the names of the summary variables must be formed from the first 16 characters and the last 15 characters of the *response* name, suffixed with the appropriate character.

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

- _PHASE_ (if the READPHASES= option is specified)

- *block-variables*
- *symbol-variable*
- BY variables
- ID variables

For an example of a SUMMARY= data set, see "Creating ANOM Charts for Proportions from Group Summary Data" on page 79.

TABLE= Data Set

You can read summary statistics and decision limits from a TABLE= data set specified in the PROC ANOM statement. This enables you to reuse an OUTTABLE= data set created in a previous run of the ANOM procedure. Because the ANOM procedure simply displays the information read from a TABLE= data set, you can use TABLE= data sets to create specialized ANOM charts.

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

Table 4.14 Variables Required in a TABLE= Data Set

Variable	Description
group-variable	values of the *group-variable*
LDLP	lower decision limit for proportions
LIMITN	nominal sample size associated with the decision limits
_P_	average proportion of nonconforming items
SUBN	group sample size
SUBP	group proportion of nonconforming items
UDLP	upper decision limit for proportions

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 whose length is no greater than 48.
- _VAR_. This variable is required if more than one *response* is specified or if the data set contains information for more than one *response*. This variable must be a character variable whose length is no greater than 32.

For an example of a TABLE= data set, see "Saving Decision Limits" on page 83.

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	*group-variable*
Vertical	DATA=	*response*
Vertical	SUMMARY=	group proportion variable
Vertical	TABLE=	_SUBP_

For example, the following sets of statements specify the label *Proportion Nonconforming* for the vertical axis of the *p* chart:

```
proc anom data=Circuits;
   pchart Fail*Batch / groupn=50;
   label Fail = 'Proportion Nonconforming';
run;

proc anom summary=Cirhist;
   pchart Fail*Batch ;
   label Failp = 'Proportion Nonconforming';
run;

proc anom table=Cirtable;
   pchart Fail*batch ;
   label _SUBP_ = 'Proportion Nonconforming';
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=, SUMMARY=, or TABLE= data set is not analyzed if the value of the group variable is missing. For a particular response variable, an observation read from a DATA= data set is not analyzed if the value of the response variable is missing. Missing values of response variables generally lead to unequal group sample sizes. For a particular response variable, an observation read from a SUMMARY= or TABLE= data set is not analyzed if the values of any of the corresponding summary variables are missing.

Examples: PCHART Statement

This section provides advanced examples of the PCHART statement.

Example 4.2: ANOM p Charts with Angled Axis Labels

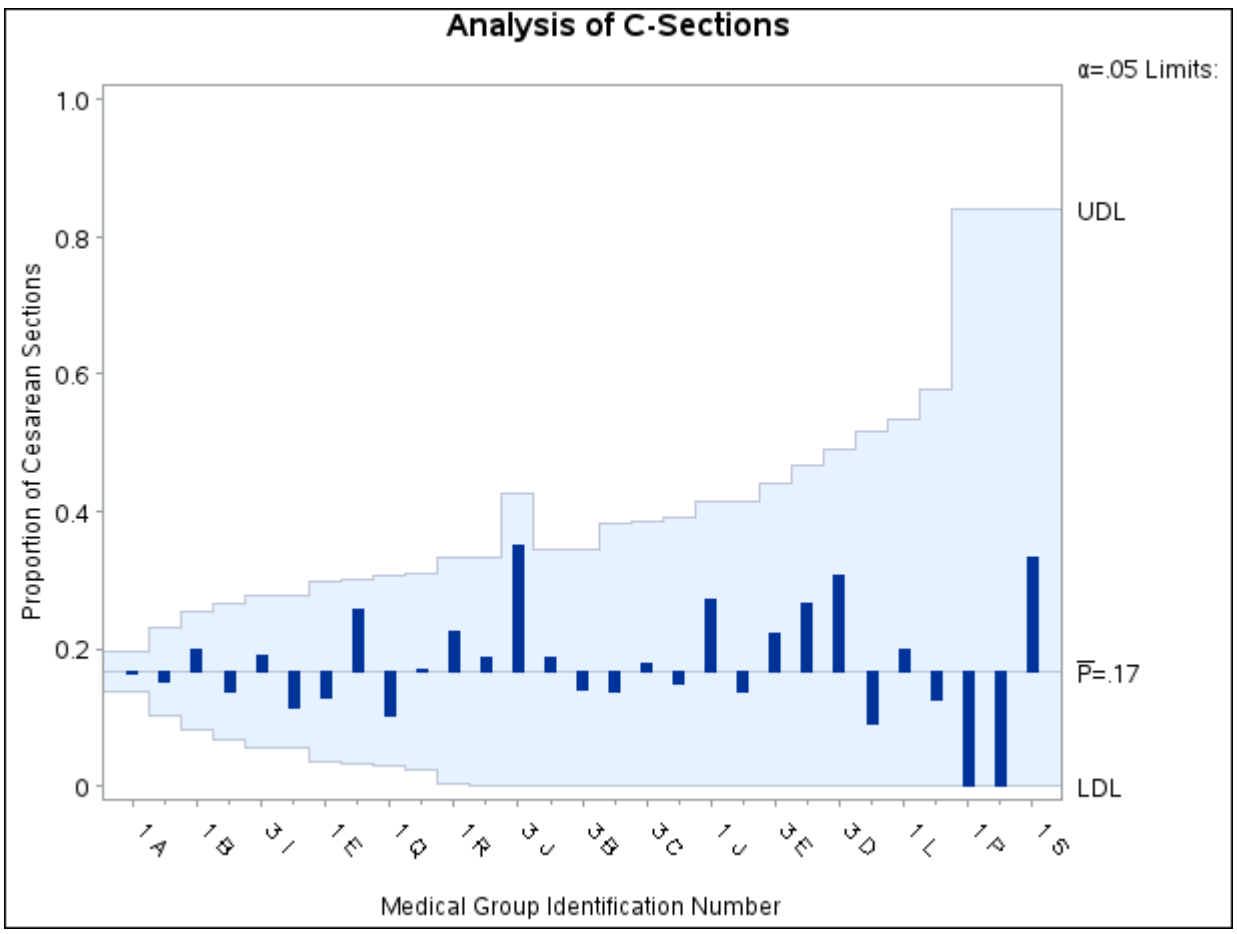
NOTE: See *ANOM p Charts with Angled Axis Labels* in the SAS/QC Sample Library.

Consider the example described in the section "Creating ANOM Charts for Proportions from Group Counts" on page 77. In the example, the option TURNHLABELS was used to vertically display the horizontal axis labels. You can also use an AXIS statement to create an angled display of the horizontal or vertical axis labels. The following statements create the p CHART shown in Output 4.2.1:

```
ods graphics off;
title 'Analysis of C-Sections';
proc anom data=Csection;
   pchart Csections*ID / groupn   = Total
                         nolegend
                         haxis    = axis1;
   axis1 value     = (a=-45 h=2.0pct);
   label Csections = 'Proportion of Cesarean Sections';
run;
```

The angle is specified with the a= option in the AXIS1 statement. Valid angle values are between -90 and 90. The height of the labels is specified with the h= option in the AXIS1 statement. See Axis and Axis Label Options.

Output 4.2.1 ANOM p Chart for C-Sections with Angled Axis Labels



UCHART Statement: ANOM Procedure

Overview: UCHART Statement

The UCHART statement creates ANOM charts for group (treatment level) rates, also referred to as ANOM *u charts*. The rate plotted on a *u* chart is the number or *count* of events occurring in a group divided by a measure of the opportunity for an event to occur.

You can use options in the UCHART statement to

- compute decision limits from the data based on specified parameters, such as the significance level (α)
- tabulate group summary statistics and decision limits
- save decision limits in an output data set
- save group summary statistics in an output data set
- read decision limits and decision limit parameters from a data set
- display distinct sets of decision limits for different sets of groups on the same chart
- add block legends and symbol markers to identify special groups
- 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

You have two alternatives for producing ANOM *u* charts with the UCHART statement:

- ODS Graphics output is produced if ODS Graphics is enabled, for example by specifying the ODS GRAPHICS ON statement prior to the PROC statement.
- Otherwise, traditional graphics are produced if SAS/GRAPH® is licensed.

See Chapter 3, "SAS/QC Graphics," for more information about producing these different kinds of graphs.

Getting Started: UCHART Statement

This section introduces the UCHART statement with simple examples that illustrate commonly used options. Complete syntax for the UCHART statement is presented in the section "Syntax: UCHART Statement" on page 108.

Creating ANOM Charts for Rates from Group Counts

NOTE: See *Creating ANOM Charts for Rates from Group Counts* in the SAS/QC Sample Library.

A health care system administrator uses ANOM to compare medical/surgical admissions rates for set of clinics. For more background concerning this application, refer to Rodriguez (1996).

The following statements create a SAS data set named MSAdmits, which contains the number of admissions and the number of member-months for each clinic during a one-year period.

```
data MSAdmits;
   length ID $ 2;
   input ID Count MemberMonths @@;
   KMemberYrs = MemberMonths/12000;
   label ID = 'Medical Group Id Number';
   datalines;
1A   1882    697204 1K    600    224715 1B    438    154720
1D    318     82254 3M    183     76450 3I    220     73529
1N    121     60169 3H    105     52886 1Q    124     52595
1E    171     51229 3B     88     34775 1C    100     31959
1H    112     28782 3C     84     27478 1R     69     26494
1T     21     25096 1M    130     24723 1O     61     24526
3D     66     22359 1J     54     19101 3J     30     16089
3G     36     13851 3E     26     10587 1G     28     10351
1I     25      6041 1L     20      5138 1S      7      2723
1F      7      2424 1P      2      2030
;
proc sort data=MSAdmits;
   by ID;
run;
```

A partial listing of MSAdmits is shown in Figure 4.17.

Figure 4.17 The Data Set MSAdmits

Medical/Surgical Admissions

ID	Count	MemberMonths	KMemberYrs
1A	1882	697204	58.1003
1B	438	154720	12.8933
1C	100	31959	2.6633
1D	318	82254	6.8545
1E	171	51229	4.2691
1F	7	2424	0.2020
1G	28	10351	0.8626
1H	112	28782	2.3985
1I	25	6041	0.5034
1J	54	19101	1.5918

There is a single observation per clinic. The variable ID identifies the clinics and is referred to as the *group-variable*. The variable Count provides the number of admissions for each clinic, which is referred to as the *response variable* (or *response* for short). The variable MemberMonths, which provides the number of member-months for each clinic, is divided by 1200 to compute the variable KMemberYrs, the number of 1000-member-years, which serves as the measure of opportunity for an admission to occur.

The following example illustrates the basic form of the UCHART statement. After the keyword UCHART, you specify the *response* to analyze (in this case, Count), followed by an asterisk and the *group-variable* (ID).

The following statements create the *u* chart shown in Figure 4.18:

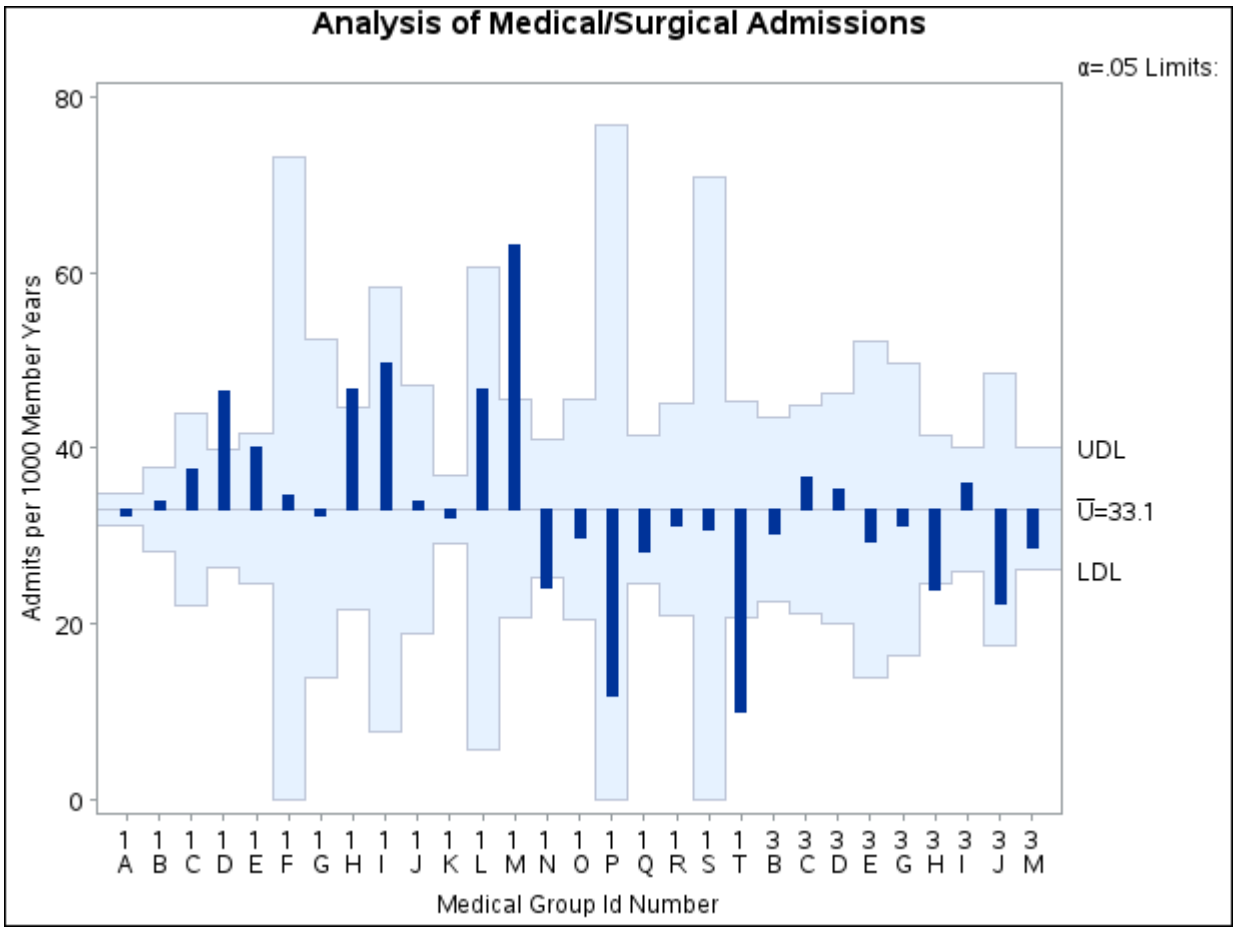
```
ods graphics off;
title 'Analysis of Medical/Surgical Admissions';
proc anom data=MSAdmits;
   uchart Count*ID / groupn = KMemberYrs
                     turnhlabels
                     nolegend;
   label Count = 'Admits per 1000 Member Years';
run;
```

The TURNHLABELS option is used to vertically display the horizontal axis labels. The GROUPN= option specifies the number of "occurrence opportunity" units in each group and is required if the input data set is a DATA= data set. In this example, 1000 member years represent one unit of opportunity. The number of units per group can be thought of as the group "sample size." You can use the GROUPN= option to specify one of the following:

- a constant number of units, which applies to all the groups
- an input variable name, which provides the number of units for each group (KMemberYrs in this example)

Options such as GROUPN= are specified after the slash (/) in the UCHART statement. A complete list of options is presented in the section "Syntax: UCHART Statement" on page 108.

Figure 4.18 u Chart Example



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

Each point on the u chart represents the rate of occurrence, computed as the count divided by the number of opportunity units. The points are displayed in the sort order for the *group-variable* ID. The chart shows that Clinics 1D, 1H, and 1M have significantly higher admissions rates, and Clinics 1N, 1T, and 3H have significantly lower admissions rates.

By default, the decision limits correspond to a significance level of $\alpha = 0.05$. This means that, assuming all clinics have the same rate of admissions, there is a 0.05 probability that one or more of the decision limits would be exceeded purely by chance. The formulas for the limits are given in the section "Decision Limits" on page 117. Note that the decision limits vary with the number of 1000-member-years for each clinic.

For more details on reading count data, see "DATA= Data Set" on page 121.

Saving Decision Limits

NOTE: See *Saving Decision Limits Using ANOM UCHART* in the SAS/QC Sample Library.

You can save the decision limits for an ANOM u chart in a SAS data set.

The following statements read the data set MSAdmits (see "Creating ANOM Charts for Rates from Group Counts" on page 104) and save the decision limits displayed in Figure 4.18 in a data set named MSLimits:

```
proc anom data=MSAdmits;
   uchart Count*ID / groupn    = KMemberYrs
                     outlimits = MSLimits
                     nochart;
run;
```

The GROUPN= option specifies the number of opportunity units for each group. The OUTLIMITS= option names the data set containing the decision limits, and the NOCHART option suppresses the display of the chart. The data set MSLimits is listed in Figure 4.19.

Figure 4.19 Data Set MSLimits Containing Decision Limits

Decision Limits for Medical/Surgical Admissions Rates

VAR	_GROUP_	_TYPE_	_LIMITN_	_ALPHA_	_LDLU_	_U_	_UDLU_	_LIMITK_
Count	ID	ESTIMATE	V	0.05	V	33.0789	V	29

The data set MSLimits contains one observation with the limits for *response* Count. The variables _LDLU_ and _UDLU_ contain the lower and upper decision limits, and the variable _U_ contains the central line. The value of _LIMITN_ is the nominal number of units associated with the decision limits (which are varying in this case), the value of _LIMITK_ is the number of groups, and the value of _ALPHA_ is the significance level of the decision limits. The variables _VAR_ and _GROUP_ are bookkeeping variables that save the *response* and *group-variable*. The variable _TYPE_ is a bookkeeping variable that indicates whether the value of _U_ is an estimate or standard (known) value. Typically, the value of _TYPE_ is 'ESTIMATE.' For more information, see "OUTLIMITS= Data Set" on page 118.

Alternatively, you can use the OUTTABLE= option to create an output data set that saves both the decision limits and the group statistics, as illustrated by the following statements:

```
proc anom data=MSAdmits;
   uchart Count*ID / groupn   = KMemberYrs
                     outtable = MSTable
                     nochart;
run;
```

The a partial listing of the data set MSTable is shown in Figure 4.20.

Figure 4.20 Data Set MSTable

Rates and Decision Limits for Medical/Surgical Admissions

VAR	ID	_ALPHA_	_LIMITN_	_SUBN_	_LDLU_	_SUBU_	_U_	_UDLU_	_EXLIM_
Count	1A	0.05	58.1003	58.1003	31.2135	32.3922	33.0789	34.9443	
Count	1B	0.05	12.8933	12.8933	28.2837	33.9710	33.0789	37.8741	
Count	1C	0.05	2.6633	2.6633	22.1550	37.5481	33.0789	44.0028	
Count	1D	0.05	6.8545	6.8545	26.3640	46.3929	33.0789	39.7938	UPPER
Count	1E	0.05	4.2691	4.2691	24.4964	40.0554	33.0789	41.6615	
Count	1F	0.05	0.2020	0.2020	0.0000	34.6535	33.0789	73.0631	
Count	1G	0.05	0.8626	0.8626	13.7710	32.4606	33.0789	52.3868	
Count	1H	0.05	2.3985	2.3985	21.5579	46.6959	33.0789	44.5999	UPPER
Count	1I	0.05	0.5034	0.5034	7.7757	49.6607	33.0789	58.3822	
Count	1J	0.05	1.5918	1.5918	18.8992	33.9249	33.0789	47.2587	

This data set contains one observation for each group. The variables _SUBU_ and _SUBN_ contain the rate of occurrence and the number of opportunity units for each group. The variables _LDLU_ and _UDLU_ contain the lower and upper decision limits, and the variable _U_ contains the central line. The variables _VAR_ and ID contain the *response* name and values of the *group-variable*, respectively. For more information, see "OUTTABLE= Data Set" on page 119.

NOTE: See *Saving ANOM UCHART Summary Statistics and Decision Limits* in the SAS/QC Sample Library.

An OUTTABLE= data set can be read later as a TABLE= data set by the ANOM procedure. For example, the following statements read MSTable and display a *u* chart (not shown here) identical to the chart in Figure 4.18:

```
ods graphics off;
title 'Analysis of Medical/Surgical Admissions';
proc anom table=MSTable;
   uchart Count*id ;
   label _subu_ = 'Admits per 1000 Member Years';
run;
```

Because the ANOM procedure simply displays the information in a TABLE= data set, you can use TABLE= data sets to create specialized ANOM charts. For more information, see the section "TABLE= Data Set" on page 123.

Syntax: UCHART Statement

The basic syntax for the UCHART statement is as follows:

UCHART *response* ∗ *group-variable* **;**

The general form of this syntax is as follows:

UCHART *responses* ∗ *group-variable* < **(***block-variables***)** >
< **=***symbol-variable* | **=***'character'* > < *options* > **;**

You can use any number of UCHART statements in the ANOM procedure. The components of the UCHART statement are described as follows.

response

responses

identify one or more responses to be analyzed. The specification of *response* depends on the input data set specified in the PROC ANOM statement.

- If counts are read from a DATA= data set, *response* must be the name of the variable containing the counts. For an example, see "Creating ANOM Charts for Rates from Group Counts" on page 104.
- If rates and numbers of opportunity units are read from a SUMMARY= data set, *response* must be the common prefix of the appropriate variables in the SUMMARY= data set.
- If rates, numbers of opportunity units, and decision limits are read from a TABLE= data set, *response* must be the value of the variable _VAR_ in the TABLE= data set.

A *response* is required. If you specify more than one response, enclose the list in parentheses. For example, the following statements request distinct ANOM *u* charts for Defects and Flaws:

```
proc anom data=Measures;
   uchart (Defects Flaws)*Sample / groupn=30;
run;
```

Note that when data are read from a DATA= data set with the UCHART statement, the GROUPN= option (which specifies the number of opportunity units per group) is required.

group-variable
is the variable that identifies groups in the data. The *group-variable* is required. In the preceding UCHART statement, sample is the group variable.

block-variables
are optional variables that identify sets of consecutive groups on the chart. The blocks are labeled in a legend, and each *block-variable* provides one level of labels in the legend.

symbol-variable
is an optional variable whose levels (unique values) determine the symbol marker used to plot the rates. 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.

options
enhance the appearance of the chart, request additional analyses, save results in data sets, and so on. "Summary of Options" on page 109 lists all options by function.

Summary of Options

The following tables list the UCHART statement options by function. Options unique to the ANOM procedure are listed in Table 4.15, and are described in detail in "Dictionary of ANOM Chart Statement Options" on page 180. Options that are common to both the ANOM and SHEWHART procedures are listed in Table 4.16, and are described in detail in "Dictionary of Options: SHEWHART Procedure" on page 1946.

Table 4.15 UCHART Statement Special Options

Option	Description
Options for Specifying Decision Limits	
ALPHA=	specifies significance level
LIMITK=	specifies number of groups for decision limits
LIMITN=	specifies either nominal sample size for fixed decision limits or varying limits
NOREADLIMITS	computes decision limits for each *response* from the data rather than a LIMITS= data set
READINDEXES=	reads multiple sets of decision limits for each *response* from a LIMITS= data set
TYPE=	identifies parameters as estimates or standard values and specifies value of _TYPE_ in the OUTLIMITS= data set
U=	specifies the weighted average of group rates
Options for Displaying Decision Limits	
CINFILL=	specifies color for area inside decision limits

Table 4.15 *continued*

Option	Description
CLIMITS=	specifies color of decision limits, central line, and related labels
LDLLABEL=	specifies label for lower decision limit
LIMLABSUBCHAR=	specifies a substitution character for labels provided as quoted strings; the character is replaced with the value of the decision limit
LLIMITS=	specifies line type for decision limits
NDECIMAL=	specifies number of digits to right of decimal place in default labels for decision limits and central line
NOCTL	suppresses display of central line
NOLDL	suppresses display of lower decision limit
NOLIMIT0	suppresses display of lower decision limit if it is 0
NOLIMITLABEL	suppresses labels for decision limits and central line
NOLIMITS	suppresses display of decision limits
NOLIMITSFRAME	suppresses default frame around decision limit information when multiple sets of decision limits are read from a LIMITS= data set
NOLIMITSLEGEND	suppresses legend for decision limits
NOUDL	suppresses display of upper decision limit
UDLLABEL=	specifies label for upper decision limit
USYMBOL=	specifies label for central line
WLIMITS=	specifies width for decision limits and central line
Input Data Set Option	
GROUPN=	specifies group sample sizes as constant number *n* or as values of variable in a DATA= data set
Output Data Set Option	
OUTSUMMARY=	creates output data set containing group summary statistics

Table 4.16 UCHART Statement General Options

Option	Description
Options for Plotting and Labeling Points	
ALLLABEL=	labels every point on ANOM *u* chart
CLABEL=	specifies color for labels
CCONNECT=	specifies color for line segments that connect points on chart
CFRAMELAB=	specifies fill color for frame around labeled points
CNEEDLES=	specifies color for needles that connect points to central line
COUT=	specifies color for portions of line segments that connect points outside decision limits
COUTFILL=	specifies color for shading areas between the connected points and decision limits outside the limits
LABELANGLE=	specifies angle at which labels are drawn

Table 4.16 *continued*

Option	Description
LABELFONT=	specifies software font for labels
LABELHEIGHT=	specifies height of labels
NONEEDLES	suppresses vertical needles connecting points to central line
OUTLABEL=	labels points outside decision limits
SYMBOLLEGEND=	specifies LEGEND statement for levels of *symbol-variable*
SYMBOLORDER=	specifies order in which symbols are assigned for levels of *symbol-variable*
TURNALL|TURNOUT	turns point labels so that they are strung out vertically
WNEEDLES=	specifies width of needles
Axis and Axis Label Options	
CAXIS=	specifies color for axis lines and tick marks
CFRAME=	specifies fill colors for frame for plot area
CTEXT=	specifies color for tick mark values and axis labels
DISCRETE	produces horizontal axis for discrete numeric group values
HAXIS=	specifies major tick mark values for horizontal axis
HEIGHT=	specifies height of axis label and axis legend text
HMINOR=	specifies number of minor tick marks between major tick marks on horizontal axis
HOFFSET=	specifies length of offset at both ends of horizontal axis
INTSTART=	specifies first major tick mark value for numeric horizontal axis with date, time, or datetime format
NOHLABEL	suppresses label for horizontal axis
NOTICKREP	specifies that only the first occurrence of repeated, adjacent group values is to be labeled on horizontal axis
NOTRUNC	suppresses vertical axis truncation at zero applied by default
NOVANGLE	requests vertical axis labels that are strung out vertically
SKIPHLABELS=	specifies thinning factor for tick mark labels on horizontal axis
TURNHLABELS	requests horizontal axis labels that are strung out vertically
VAXIS=	specifies major tick mark values for vertical axis of ANOM *u* chart
VFORMAT=	specifies format for vertical axis tick mark labels
VMINOR=	specifies number of minor tick marks between major tick marks on vertical axis
VOFFSET=	specifies length of offset at both ends of vertical axis
VZERO	forces origin to be included in vertical axis for ANOM *u* chart
WAXIS=	specifies width of axis lines

Table 4.16 *continued*

Option	Description
Plot Layout Options	
ALLN	plots means for all groups
BILEVEL	creates ANOM *u* chart using half-screens and half-pages
EXCHART	creates ANOM *u* chart for a response only when a group mean exceeds the decision limits
INTERVAL=	natural time interval between consecutive group positions when time, date, or datetime format is associated with a numeric group variable
MAXPANELS=	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 decision limits
NOCHART	suppresses creation of chart
NOFRAME	suppresses frame for plot area
NOLEGEND	suppresses legend for group sample sizes
NPANELPOS=	specifies number of group positions per panel on each chart
REPEAT	repeats last group position on panel as first group position of next panel
TOTPANELS=	specifies number of pages or screens to be used to display chart
ZEROSTD	displays ANOM *u* chart regardless of whether root mean square error is zero
Reference Line Options	
CHREF=	specifies color for lines requested by HREF= option
CVREF=	specifies color for lines requested by VREF= option
HREF=	specifies position of reference lines perpendicular to horizontal axis on ANOM *u* chart
HREFDATA=	specifies position of reference lines perpendicular to horizontal axis on ANOM *u* chart
HREFLABELS=	specifies labels for HREF= lines
HREFLABPOS=	specifies position of HREFLABELS= labels
LHREF=	specifies line type for HREF= lines
LVREF=	specifies line type for VREF= lines
NOBYREF	specifies that reference line information in a data set applies uniformly to charts created for all BY groups
VREF=	specifies position of reference lines perpendicular to vertical axis on ANOM *u* chart
VREFLABELS=	specifies labels for VREF= lines
VREFLABPOS=	specifies position of VREFLABELS= labels
Grid Options	
CGRID=	specifies color for grid requested with GRID or ENDGRID option
ENDGRID	adds grid after last plotted point

Table 4.16 *continued*

Option	Description
GRID	adds grid to control chart
LENDGRID=	specifies line type for grid requested with the ENDGRID option
LGRID=	specifies line type for grid requested with the GRID option
WGRID=	specifies width of grid lines
Clipping Options	
CCLIP=	specifies color for plot symbol for clipped points
CLIPFACTOR=	determines extent to which extreme points are clipped
CLIPLEGEND=	specifies text for clipping legend
CLIPLEGPOS=	specifies position of clipping legend
CLIPSUBCHAR=	specifies substitution character for CLIPLEGEND= text
CLIPSYMBOL=	specifies plot symbol for clipped points
CLIPSYMBOLHT=	specifies symbol marker height for clipped points
Graphical Enhancement Options	
ANNOTATE=	specifies annotate data set that adds features to ANOM *u* chart
DESCRIPTION=	specifies description of ANOM *u* chart's GRSEG catalog entry
FONT=	specifies software font for labels and legends on chart
NAME=	specifies name of ANOM *u* chart's GRSEG catalog entry
PAGENUM=	specifies the form of the label used in pagination
PAGENUMPOS=	specifies the position of the page number requested with the PAGENUM= option
Options for Producing Graphs Using ODS Styles	
BLOCKVAR=	specifies one or more variables whose values define colors for filling background of *block-variable* legend
CFRAMELAB	draws a frame around labeled points
COUT	draw portions of line segments that connect points outside decision limits in a contrasting color
CSTAROUT	specifies that portions of stars exceeding inner or outer circles are drawn using a different color
OUTFILL	shades areas between decision limits and connected points lying outside the limits
STARFILL=	specifies a variable identifying groups of stars filled with different colors
STARS=	specifies a variable identifying groups of stars whose outlines are drawn with different colors
Options for ODS Graphics	
BLOCKREFTRANSPARENCY=	specifies the wall fill transparency for blocks and phases
INFILLTRANSPARENCY=	specifies the decision limit infill transparency
MARKERS	plots group points with markers

Table 4.16 *continued*

Option	Description
NOBLOCKREF	suppresses block and phase reference lines
NOBLOCKREFFILL	suppresses block and phase wall fills
NOFILLLEGEND	suppresses legend for levels of a STARFILL= variable
NOPHASEREF	suppresses block and phase reference lines
NOPHASEREFFILL	suppresses block and phase wall fills
NOREF	suppresses block and phase reference lines
NOREFFILL	suppresses block and phase wall fills
NOSTARFILLLEGEND	suppresses legend for levels of a STARFILL= variable
NOTRANSPARENCY	disables transparency in ODS Graphics output
ODSFOOTNOTE=	specifies a graph footnote
ODSLEGENDEXPAND	specifies that legend entries contain all levels observed in the data
ODSTITLE=	specifies a graph title
OUTFILLTRANSPARENCY=	specifies decision limit outfill transparency
OVERLAYURL=	specifies URLs to associate with overlay points
PHASEPOS=	specifies vertical position of phase legend
PHASEREFLEVEL=	associates phase and block reference lines with either innermost or the outermost level
PHASEREFTRANSPARENCY=	specifies the wall fill transparency for blocks and phases
REFFILLTRANSPARENCY=	specifies the wall fill transparency for blocks and phases
SIMULATEQCFONT	draws central line labels using a simulated software font
STARTRANSPARENCY=	specifies star fill transparency
URL=	specifies a variable whose values are URLs to be associated with groups
Output Data Set Options	
OUTINDEX=	specifies value of _INDEX_ in the OUTLIMITS= data set
OUTLIMITS=	creates output data set containing decision limits
OUTTABLE=	creates output data set containing group summary statistics and decision limits
Tabulation Options	
NOTE: specifying (EXCEPTIONS) after a tabulation option creates a table for exceptional points only.	
TABLE	creates a basic table of group means, group sample sizes, and decision 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
TABLEOUTLIM	augments basic table with columns indicating decision limits exceeded
Block Variable Legend Options	
BLOCKLABELPOS=	specifies position of label for *block-variable* legend
BLOCKLABTYPE=	specifies text size of *block-variable* legend
BLOCKPOS=	specifies vertical position of *block-variable* legend

Table 4.16 *continued*

Option	Description
BLOCKREP	repeats identical consecutive labels in *block-variable* legend
CBLOCKLAB=	specifies fill colors for frames enclosing variable labels in *block-variable* legend
CBLOCKVAR=	specifies one or more variables whose values are colors for filling background of *block-variable* legend
Phase Options	
CPHASELEG=	specifies text color for *phase* legend
NOPHASEFRAME	suppresses default frame for *phase* legend
OUTPHASE=	specifies value of _PHASE_ in the OUTHISTORY= data set
PHASEBREAK	disconnects last point in a *phase* from first point in next *phase*
PHASELABTYPE=	specifies text size of *phase* legend
PHASELEGEND	displays *phase* labels in a legend across top of chart
PHASELIMITS	labels decision limits for each phase, provided they are constant within that phase
PHASEREF	delineates *phases* with vertical reference lines
READPHASES=	specifies *phases* to be read from an input data set
Star Options	
CSTARCIRCLES=	specifies color for STARCIRCLES= circles
CSTARFILL=	specifies color for filling stars
CSTAROUT=	specifies outline color for stars exceeding inner or outer circles
CSTARS=	specifies color for outlines of stars
LSTARCIRCLES=	specifies line types for STARCIRCLES= circles
LSTARS=	specifies line types for outlines of STARVERTICES= stars
STARBDRADIUS=	specifies radius of outer bound circle for vertices of stars
STARCIRCLES=	specifies reference circles for stars
STARINRADIUS=	specifies inner radius of stars
STARLABEL=	specifies vertices to be labeled
STARLEGEND=	specifies style of legend for star vertices
STARLEGENDLAB=	specifies label for STARLEGEND= legend
STAROUTRADIUS=	specifies outer radius of stars
STARSPECS=	specifies method used to standardize vertex variables
STARSTART=	specifies angle for first vertex
STARTYPE=	specifies graphical style of star
STARVERTICES=	superimposes star at each point on ANOM u chart
WSTARCIRCLES=	specifies width of STARCIRCLES= circles
WSTARS=	specifies width of STARVERTICES= stars
Overlay Options	
CCOVERLAY=	specifies colors for overlay line segments
COVERLAY=	specifies colors for overlay plots

Table 4.16 *continued*

Option	Description
COVERLAYCLIP=	specifies color for clipped points on overlays
LOVERLAY=	specifies line types for overlay line segments
NOOVERLAYLEGEND	suppresses legend for overlay plots
OVERLAY=	specifies variables to overlay on chart
OVERLAYCLIPSYM=	specifies symbol for clipped points on overlays
OVERLAYCLIPSYMHT=	specifies symbol height for clipped points on overlays
OVERLAYHTML=	specifies links to associate with overlay points
OVERLAYID=	specifies labels for overlay points
OVERLAYLEGLAB=	specifies label for overlay legend
OVERLAYSYM=	specifies symbols for overlays
OVERLAYSYMHT=	specifies symbol heights for overlays
WOVERLAY=	specifies widths of overlay line segments
Options for Interactive ANOM Charts	
HTML=	specifies a variable whose values create links to be associated with groups
HTML_LEGEND=	specifies a variable whose values create links to be associated with symbols in the symbol legend
WEBOUT=	creates an OUTTABLE= data set with additional graphics coordinate data

Details: UCHART Statement

Constructing ANOM Charts for Rates

The following notation is used in this section:

c_i	count (number of occurrences) in the ith group
k	number of groups
n_i	number of occurrence opportunity units in the ith group
N	total sample size $= n_1 + \cdots + n_k$
u_i	occurrence rate in the ith group ($u_i = c_i/n_i$)
$\bar{u}$	average of occurrence rates taken across groups. The quantity $\bar{u}$ is computed as a weighted average: $$\bar{u} = \frac{n_1 u_1 + \cdots + n_k u_k}{N} = \frac{c_1 + \cdots + c_k}{N}$$
α	significance level
$h(\alpha; k, n, \infty)$	critical value for ANOM for normal data in the balanced case ($n_i \equiv n$)
$h(\alpha; k, n_1, \ldots, n_k, \infty)$	critical value for ANOM for normal data in the unbalanced case

Plotted Points

Each point on a u chart indicates the rate of occurrence (u_i) in a group.

Central Line

In an ANOM chart for rates, the central line represents the weighted average of the group rates, which is denoted by $\bar{u}$.

Decision Limits

For the ith group, the occurrence counts are assumed to have a Poisson distribution with parameter λ_i. The ANOM method tests the null hypothesis that $\lambda_1 = \cdots = \lambda_k$, that is, that the rates are the same, against the alternative that at least one of the λ_i's is different from the average of the k rates.

The decision limits are computed using the normal approximation to the Poisson distribution, which is appropriate when the sample sizes for the groups are large; see Ramig (1983). A commonly recommended check for this assumption is that $c_i > 5$ for all the groups. The critical values in the ANOM method for normally distributed data are adapted to the Poisson case by using infinite degrees of freedom for the variance.

When the number of opportunity units is constant ($n_i \equiv n$) across groups, the decision limits are computed as follow:

$$\begin{aligned}
\text{lower decision limit (LDLU)} &= \max\left(\bar{u} - h(\alpha; k, n, \infty)\sqrt{\bar{u}}\sqrt{\frac{k-1}{N}}\ , 0\right) \\
\text{upper decision limit (UDLU)} &= \bar{u} + h(\alpha; k, n, \infty)\sqrt{\bar{u}}\sqrt{\frac{k-1}{N}}
\end{aligned}$$

For the theoretical derivation of the decision limits, refer to Nelson (1982a).

When the number of opportunity units (n_i) is different across groups (the unbalanced case), the decision limits are computed as follows:

$$\begin{aligned}
\text{lower decision limit (LDLU)} &= \max\left(\bar{u} - h(\alpha; k, n_1, \ldots, n_k, \infty)\sqrt{\bar{u}}\sqrt{\frac{N-n_i}{Nn_i}}\ , 0\right) \\
\text{upper decision limit (UDLU)} &= \bar{u} + h(\alpha; k, n_1, \ldots, n_k, \infty)\sqrt{\bar{u}}\sqrt{\frac{N-n_i 1}{Nn_i}}
\end{aligned}$$

Note that the decision limits for the ith group depend on n_i. If the sample sizes are constant across groups ($n_i \equiv n$), the decision limits in the unbalanced case reduce to the formulas given for the balanced case, since $n_i \equiv n$ and $N = kn$, so

$$\sqrt{\frac{N-n_i}{Nn_i}} = \sqrt{\frac{kn-n}{Nn}} = \sqrt{\frac{k-1}{N}}$$

For the derivation of the decision limits for unequal sample sizes, refer to Nelson (1991).

Exact critical values were first tabulated by Nelson (1982a). Refer to Nelson (1993) for a derivation of the critical values $h(\alpha; k, n, \infty)$ and Nelson (1991) for a derivation of the critical values $h(\alpha; k, n_1, \ldots, n_k, \infty)$. Note that the critical values in the unequal sample size case have not been tabulated.

You can specify parameters for the limits as follows:

- Specify α with the ALPHA= option or with the variable _ALPHA_ in a LIMITS= data set.
- Specify a nominal constant number of opportunity units $n_i \equiv n$ with the LIMITN= option or with the variable _LIMITN_ in a LIMITS= data set.
- Specify $\bar{u}$ with the U= option or with the variable _U_ in a LIMITS= data set.

Output Data Sets

OUTLIMITS= Data Set

The OUTLIMITS= data set saves decision limits and decision limit parameters. The following variables can be saved:

Table 4.17 OUTLIMITS= Data Set

Variable	Description
ALPHA	significance level (α)
GROUP	*group-variable* specified in the UCHART statement
INDEX	optional identifier for the decision limits specified with the OUTINDEX= option
LDLU	lower decision limit for occurrence rates
LIMITK	number of groups
LIMITN	number of opportunity units associated with the decision limits
TYPE	type (estimate or standard value) of _U_
_U_	value of central line of u chart ($\bar{u}$)
UDLU	upper decision limit for occurrence rates
VAR	*response* specified in the UCHART statement

Notes:

1. If the decision limits vary with the number of opportunity units, the special missing value V is assigned to the variables _LDLU_, _UDLU_, and _LIMITN_.
2. Optional BY variables are saved in the OUTLIMITS= data set.

The OUTLIMITS= data set contains one observation for each *response* specified in the UCHART statement. For an example, see "Saving Decision Limits" on page 106.

OUTSUMMARY= Data Set

The OUTSUMMARY= data set saves group summary statistics. The following variables are saved:

- the *group-variable*
- a response rate variable, whose name is *response* suffixed with *U*
- a number of opportunity units variable, whose name is *response* suffixed with *N*

Given a *response* name that contains 32 characters, the procedure first shortens the name to its first 16 characters and its last 15 characters, and then it adds the suffix.

Group summary variables are created for each *response* specified in the UCHART statement. For example, consider the following statements:

```
proc anom data=Fabric;
   uchart (Flaws Ndefects)*Treatment / outsummary=Summary
                                       groupn = 30;
run;
```

The data set summary contains the variables Treatment, FlawsU, FlawsN, NdefectsU, and NdefectsN.

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

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

OUTTABLE= Data Set

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

Variable	Description
ALPHA	significance level (α)
EXLIM	decision limit exceeded on *u* chart
group	values of the group variable
LDLU	lower decision limit for group rate
LIMITN	nominal number of opportunity units associated with the decision limits
SUBU	group rate
SUBN	number of opportunity units in group
UDLU	upper decision limit for group rate
VAR	*response* specified in the UCHART 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)

NOTE: The variable _EXLIM_ is a character variable of length 8. The variable _PHASE_ is a character variable of length 48. The variable _VAR_ is a character variable whose length is no greater than 32. All other variables are numeric.

ODS Tables

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

Table 4.18 ODS Tables Produced with the UCHART Statement

Table Name	Description	Options
UCHART	ANOM *u* chart summary statistics	TABLE, TABLEALL, TABLEC, TABLEID, TABLEOUT

ODS Graphics

Before you create ODS Graphics output, ODS Graphics must be enabled (for example, by using the ODS GRAPHICS ON statement). For more information about enabling and disabling ODS Graphics, see the section "Enabling and Disabling ODS Graphics" (Chapter 21, *SAS/STAT User's Guide*). **NOTE:** In SAS/QC 13.1 the ANOM procedure does not support the creation of graphs that are editable with the ODS Graphics Editor.

The appearance of a graph produced with ODS Graphics is determined by the style associated with the ODS destination where the graph is produced. UCHART options used to control the appearance of traditional graphics are ignored for ODS Graphics output. Options for Producing Graphs Using ODS Styles lists options that can be used to control the appearance of graphs produced with ODS Graphics or with traditional graphics using ODS styles. Options for ODS Graphics lists options to be used exclusively with ODS Graphics. Detailed descriptions of these options are provided in "Dictionary of Options: SHEWHART Procedure" on page 1946.

When ODS Graphics is in effect, the UCHART statement assigns a name to the graph it creates. You can use this name to reference the graph when using ODS. The name is listed in Table 4.19.

Table 4.19 ODS Graphics Produced by the UCHART Statement

ODS Graph Name	Plot Description
UChart	ANOM *u* chart

See Chapter 3, "SAS/QC Graphics," for more information about ODS Graphics and other methods for producing charts.

Input Data Sets

DATA= Data Set

You can read response counts for groups from a DATA= data set specified in the PROC ANOM statement. Each *response* specified in the UCHART statement must be a SAS variable in the data set. This variable provides the count (number of occurrences) for groups indexed by the *group-variable*. The *group-variable*, specified in the UCHART 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 *response* and a value for the *group-variable*. The data set should contain one observation per group. When you use a DATA= data set with the UCHART statement, the GROUPN= option (which specifies the number of inspection units per group) is required. 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

For an example of a DATA= data set, see "Creating ANOM Charts for Rates from Group Counts" on page 104.

LIMITS= Data Set

You can read decision limits (or parameters from which the decision limits can be calculated) from a LIMITS= data set specified in the PROC ANOM statement. For example, the following statements read decision limit information from the data set Conlims:

```
proc anom data=Info limits=Conlims;
   uchart Defects*Treatment / groupn = 30;
run;
```

The LIMITS= data set can be an OUTLIMITS= data set that was created in a previous run of the ANOM 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 _LDLU_, _U_, and _UDLU_, which specify the decision limits
- the variable _U_, without providing the variables _LDLU_ and _UDLU_, which is used to calculate the decision limits (see "Decision Limits" on page 117)

In addition, note the following:

- The variables _VAR_ and _GROUP_ are always required. These must be character variables whose lengths are no greater than 32.
- _LDLU_ and _UDLU_ must be specified together; otherwise their values are computed.

- _ALPHA_ is optional but is recommended in order to maintain a complete set of decision limit information. The default value is 0.05.
- _LIMITK_ is optional. The default value is k, the number of groups. A group must have at least one nonmissing value ($n_i \geq 1$) and there must be at least one group with $n_i \geq 2$. If specified, _LIMITK_ overrides the value of k.
- _LIMITN_ is optional. The default value is the common group size (n), in the balanced case $n_i \equiv n$. If specified, _LIMITN_ overrides the value of n.
- The variable _TYPE_ is optional, but is recommended to maintain a complete set of decision limit information. The variable _TYPE_ must be a character variable of length 8. Valid values are 'ESTIMATE,' 'STANDARD,' 'STDMEAN,' and 'STDRMS.' The default is 'ESTIMATE.'
- The variable _INDEX_ is required if you specify the READINDEX= option; this must be a character variable whose length is no greater than 48.
- BY variables are required if specified with a BY statement.

SUMMARY= Data Set

You can read group summary statistics from a SUMMARY= data set specified in the PROC ANOM statement. This enables you to reuse OUTSUMMARY= data sets that have been created in previous runs of the ANOM procedure or to read output data sets created with SAS summarization procedures.

A SUMMARY= data set used with the UCHART statement must contain the following variables:

- *group-variable*
- response rates for each *response*
- number of occurrence units for each *response*

The names of the variables containing the rates and number of occurrence units must be the *response* name concatenated with the special suffix characters *U* and *N*, respectively. For example, consider the following statements:

```
proc anom summary=Summary;
   uchart (Flaws Ndefects)*Treatment;
run;
```

The data set Summary must include the variables Treatment, FlawsU, FlawsN, NdefectsU, and NdefectsN.

Note that if you specify a *response* name that contains 32 characters, the names of the summary variables must be formed from the first 16 characters and the last 15 characters of the *response* name, suffixed with the appropriate character.

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

- _PHASE_ (if the READPHASES= option is specified)
- *block-variables*

- *symbol-variable*
- BY variables
- ID variables

TABLE= Data Set

You can read group statistics and decision limits from a TABLE= data set specified in the PROC ANOM statement. This enables you to reuse an OUTTABLE= data set created in a previous run of the ANOM procedure or to create your own TABLE= data set. Because the ANOM procedure simply displays the information read from a TABLE= data set, you can use TABLE= data sets to create specialized ANOM charts.

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

Table 4.20 Variables Required in a TABLE= Data Set

Variable	Description
group-variable	values of the *group-variable*
LDLU	lower decision limit for rate
LIMITN	nominal number of opportunity units associated with the decision limits
SUBN	number of opportunity units in group
SUBU	response rate for group
_U_	weighted average of group rates
UDLU	upper decision limit for rate

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 whose length is no greater than 48.
- _VAR_. This variable is required if more than one *response* is specified or if the data set contains information for more than one *response*. This variable must be a character variable whose length is no greater than 32.

For an example of a TABLE= data set, see "Saving Decision Limits" on page 106.

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	*group-variable*
Vertical	DATA=	*response*
Vertical	SUMMARY=	group defects per unit variable
Vertical	TABLE=	_SUBU_

Missing Values

An observation read from a DATA=, SUMMARY=, or TABLE= data set is not analyzed if the value of the group variable is missing. For a particular response variable, an observation read from a DATA= data set is not analyzed if the value of the response variable is missing. For a particular response variable, an observation read from a SUMMARY= or TABLE= data set is not analyzed if the values of any of the corresponding summary variables are missing.

Examples: UCHART Statement

This section provides an advanced example of the UCHART statement.

Example 4.3: ANOM u Charts with Angled Axis Labels

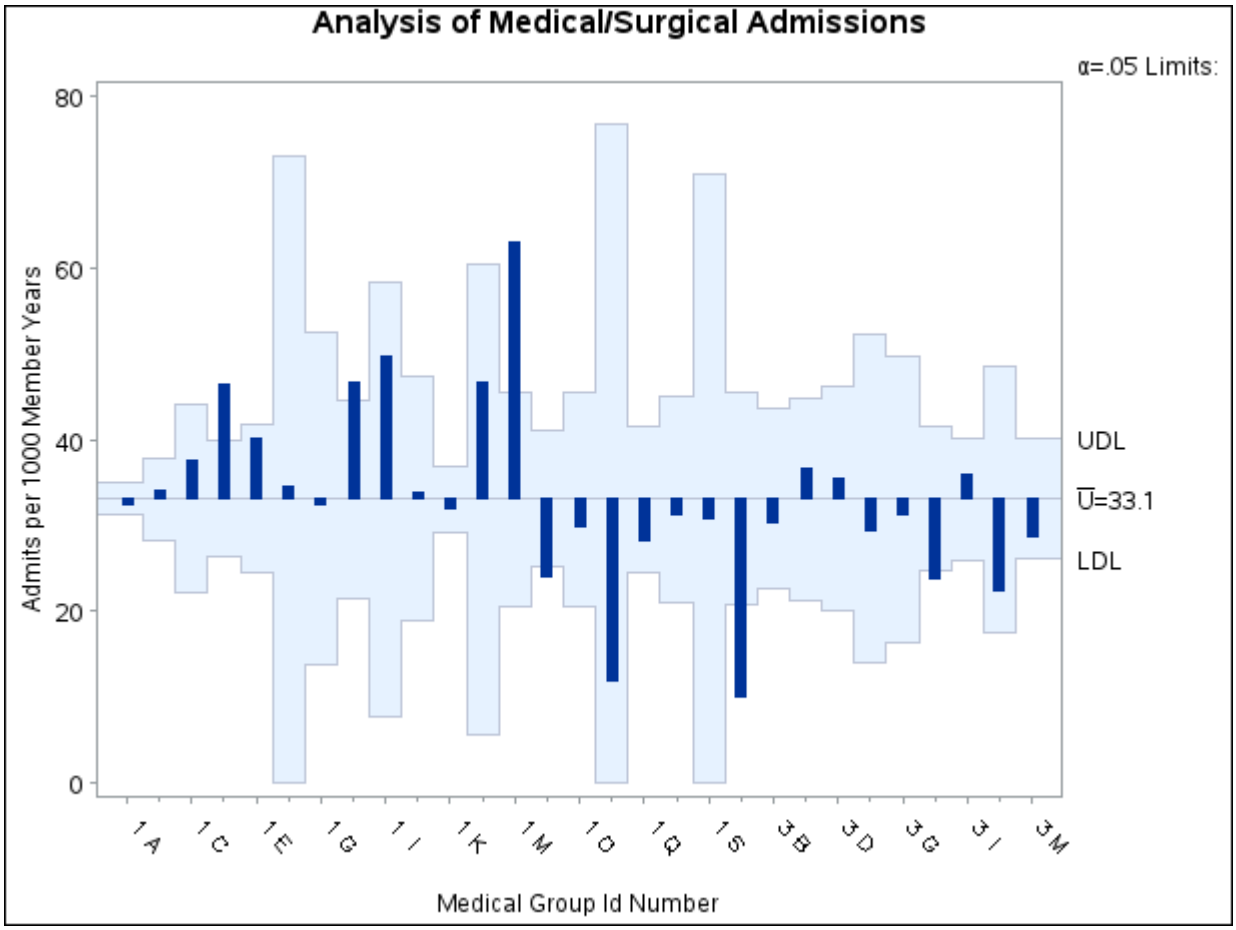
NOTE: See *Creating ANOM Charts with Angled Axis Labels* in the SAS/QC Sample Library.

Consider the example described in "Creating ANOM Charts for Rates from Group Counts" on page 104. In the example, the option TURNHLABELS was used to vertically display the horizontal axis labels. You can also use an AXIS statement to create an angled display of the horizontal or vertical axes labels. The following statements create the u CHART shown in Output 4.3.1:

```
ods graphics off;
title 'Analysis of Medical/Surgical Admissions';
proc anom data=MSAdmits;
   uchart Count*ID / groupn   = KMemberYrs
                     nolegend
                     haxis    = axis1;
axis1 value = (a=-45 h=2.0pct);
label Count = 'Admits per 1000 Member Years';
run;
```

The angle is specified with the A= option in the AXIS1 statement. Valid angle values are between `-90` and `90`. The height of the labels is specified with the H= option in the AXIS1 statement. See Axis and Axis Label Options in Table 4.16 for a list of UCHART statement axis options and *SAS/GRAPH: Reference* for a complete description of the AXIS statement.

Output 4.3.1 ANOM u Chart for C-Sections with Angled Axis Labels



XCHART Statement: ANOM Procedure

Overview: XCHART Statement

The XCHART statement creates an ANOM chart for group (treatment level) means of response values. You can use options in the XCHART statement to

- compute decision limits from the data based on specified parameters, such as the significance level (α)
- tabulate group sample sizes, group means, decision limits, and other information
- save decision limits in an output data set
- save group sample sizes and group means in an output data set
- read decision limits and decision limit parameters from a data set
- display distinct sets of decision limits for different sets of groups
- add block legends and symbol markers to identify special groups
- 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

You have two alternatives for producing ANOM charts with the XCHART statement:

- ODS Graphics output is produced if ODS Graphics is enabled, for example by specifying the ODS GRAPHICS ON statement prior to the PROC statement.
- Otherwise, traditional graphics are produced if SAS/GRAPH® is licensed.

See Chapter 3, "SAS/QC Graphics," for more information about producing these different kinds of graphs.

Getting Started: XCHART Statement

This section introduces the XCHART statement with simple examples that illustrate the most commonly used options. Complete syntax for the XCHART statement is presented in the section "Syntax: XCHART Statement" on page 134, and advanced examples are given in the section "Examples: XCHART Statement" on page 153.

Creating ANOM Charts for Means from Response Values

NOTE: See *Creating ANOM Charts for Means from Response Variables* in the SAS/QC Sample Library.

A manufacturing engineer carries out a study to determine the source of excessive variation in the positioning of labels on shampoo bottles. [2] A labeling machine removes bottles from the line, attaches the labels, and returns the bottles to the line. There are six positions on the machine, and the engineer suspects that one or more of the position heads might be faulty.

A sample of 60 bottles, 10 per position, is run through the machine. For each bottle, the deviation of the label is measured in millimeters, and the machine position is recorded. The following statements create a SAS data set named LabelDeviations, which contains the deviation measurements for the 60 bottles:

```
data LabelDeviations;
   input Position @;
   do i = 1 to 5;
      input Deviation @;
      output;
   end;
   drop i;
   datalines;
1  -0.02386  -0.02853  -0.03001  -0.00428  -0.03623
1  -0.04222  -0.00144  -0.06466   0.00944  -0.00163
2  -0.02014  -0.02725   0.02268  -0.03323   0.03661
2   0.04378   0.05562   0.00977   0.05641   0.01816
3  -0.00728   0.02849  -0.04404  -0.02214  -0.01394
3   0.04855   0.03566   0.02345   0.01339  -0.00203
4   0.06694   0.10729   0.05974   0.06089   0.07551
4   0.03620   0.05614   0.08985   0.04175   0.05298
5   0.03677   0.00361   0.03736   0.01164  -0.00741
5   0.02495  -0.00803   0.03021  -0.00149  -0.04640
6   0.00493  -0.03839  -0.02037  -0.00487  -0.01202
6   0.00710  -0.03075   0.00167  -0.02845  -0.00697
;
```

A partial listing of LabelDeviations is shown in Figure 4.21.

[2] This example is based on a case study described by Hansen (1990).

Figure 4.21 Partial Listing of the Data Set LabelDeviations

The Data Set LabelDeviations

Position	Deviation
1	-0.02386
1	-0.02853
1	-0.03001
1	-0.00428
1	-0.03623
1	-0.04222
1	-0.00144
1	-0.06466
1	0.00944
1	-0.00163
2	-0.02014
2	-0.02725
2	0.02268
2	-0.03323
2	0.03661
2	0.04378
2	0.05562
2	0.00977
2	0.05641
2	0.01816

The data set LabelDeviations is said to be in "strung-out" form, since each observation contains the position and the deviation measurement for a single bottle. The first 10 observations contain the measurements for the first position, the second 10 observations contain the measurements for the second position, and so on. Because the variable Position classifies the observations into groups (treatment levels), it is referred to as the *group-variable*. The input data set must be sorted by the group variable. The variable Deviation contains the deviation measurements and is referred to as the *response variable* (or *response* for short).

The following statements create an ANOM chart for Position:

```
ods graphics on;
title 'Analysis of Label Deviations';
proc anom data=LabelDeviations;
   xchart Deviation*Position / alpha    = 0.05
                               odstitle = title;
   label Deviation = 'Mean Deviation from Center (mm)';
   label Position  = 'Labeler Position';
run;
```

The ODS GRAPHICS ON statement specified before the PROC ANOM statement enables ODS Graphics, so the ANOM chart is created using ODS Graphics instead of traditional graphics. The resulting chart is shown in Figure 4.22.

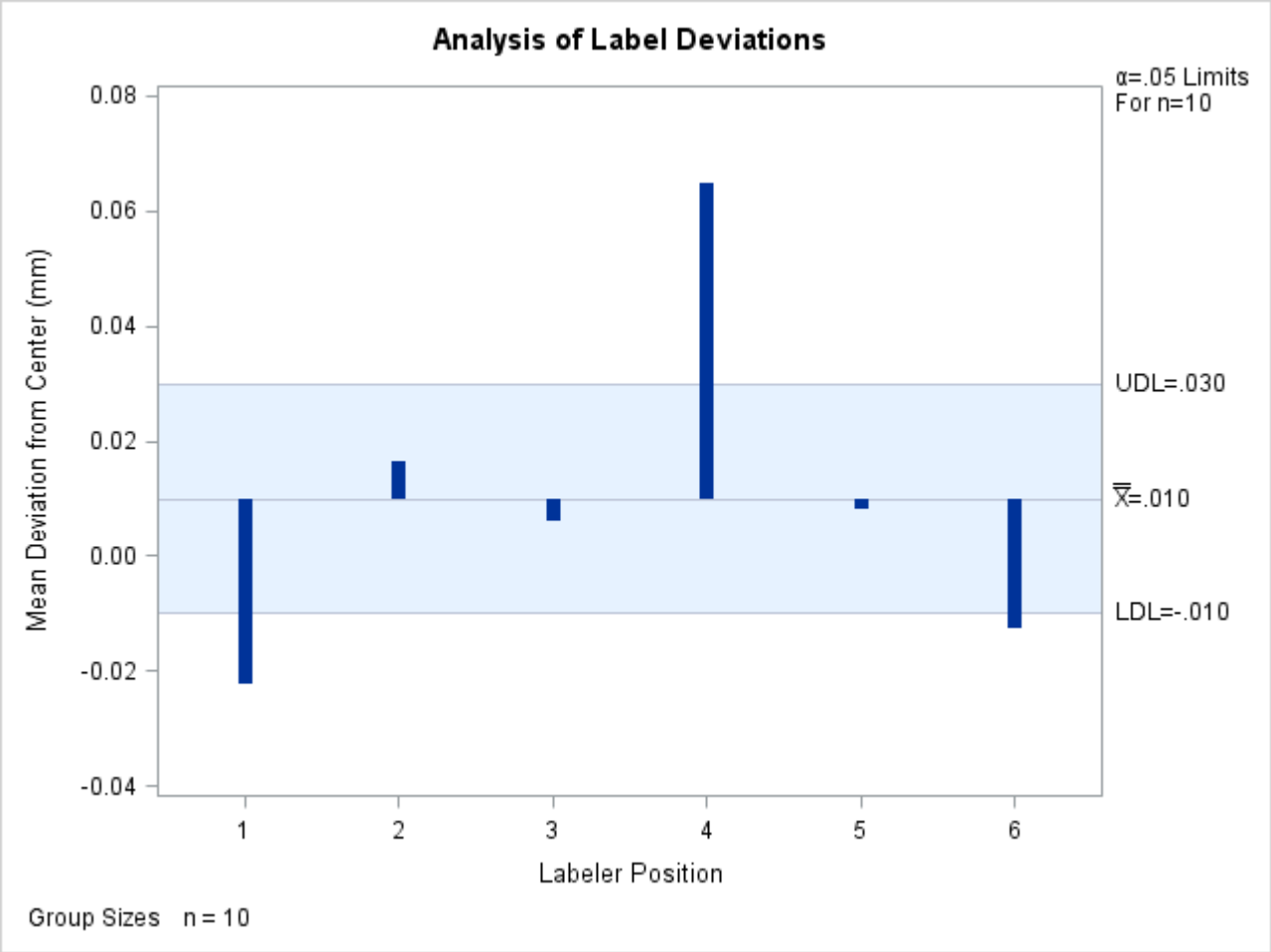
This example illustrates the basic form of the XCHART statement. After the keyword XCHART, you specify the *response* to analyze (in this case, Deviation) followed by an asterisk and the *group-variable* (Position).

Options are specified after the slash (/) in the XCHART statement. A complete list of options is presented in the section "Syntax: XCHART Statement" on page 134.

The input data set is specified with the DATA= option in the PROC ANOM statement when it contains raw measurements for the *response*.

Each point on the ANOM chart represents the average (mean) of the response measurements for a particular sample.

Figure 4.22 ANOM Chart for Means of Labeler Position Data



The average for Position 1 is below the lower decision limit (LDL), and the average for Position 6 is slightly below the lower decision limit. The average for Position 4 exceeds the upper decision limit (UDL). The conclusion is that Positions 1, 4, and 6 are operating differently.

By default, the decision limits shown correspond to a significance level of $\alpha = 0.05$; the formulas for the limits are given in the section "Decision Limits" on page 143. You can also read decision limits from an input data set.

For computational details, see "Constructing ANOM Charts for Means" on page 143. For details on reading raw measurements, see "DATA= Data Set" on page 149.

Creating ANOM Charts for Means from Group Summary Data

NOTE: See *Creating ANOM Charts for Means from Group Summary Data* in the SAS/QC Sample Library.

The previous example illustrates how you can create ANOM charts for means using measurement data. However, in many applications, the data are provided as group summary statistics. This example illustrates how you can use the XCHART statement with data of this type.

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

```
data Labels;
   input Position DeviationX DeviationS;
   DeviationN = 10;
   datalines;
1  -0.02234  0.02281
2   0.01624  0.03348
3   0.00601  0.02885
4   0.06473  0.02149
5   0.00812  0.02592
6  -0.01281  0.01597
;
```

A listing of Labels is shown in Figure 4.23. There is exactly one observation for each group (note that the groups are still indexed by Position). The variable DeviationX contains the group means, the variable DeviationS contains the group standard deviations, and the variable DeviationN contains the group sample sizes (these are all 10).

Figure 4.23 The Summary Data Set Labels

The Data Set Labels

Position	DeviationX	DeviationS	DeviationN
1	-0.02234	0.02281	10
2	0.01624	0.03348	10
3	0.00601	0.02885	10
4	0.06473	0.02149	10
5	0.00812	0.02592	10
6	-0.01281	0.01597	10

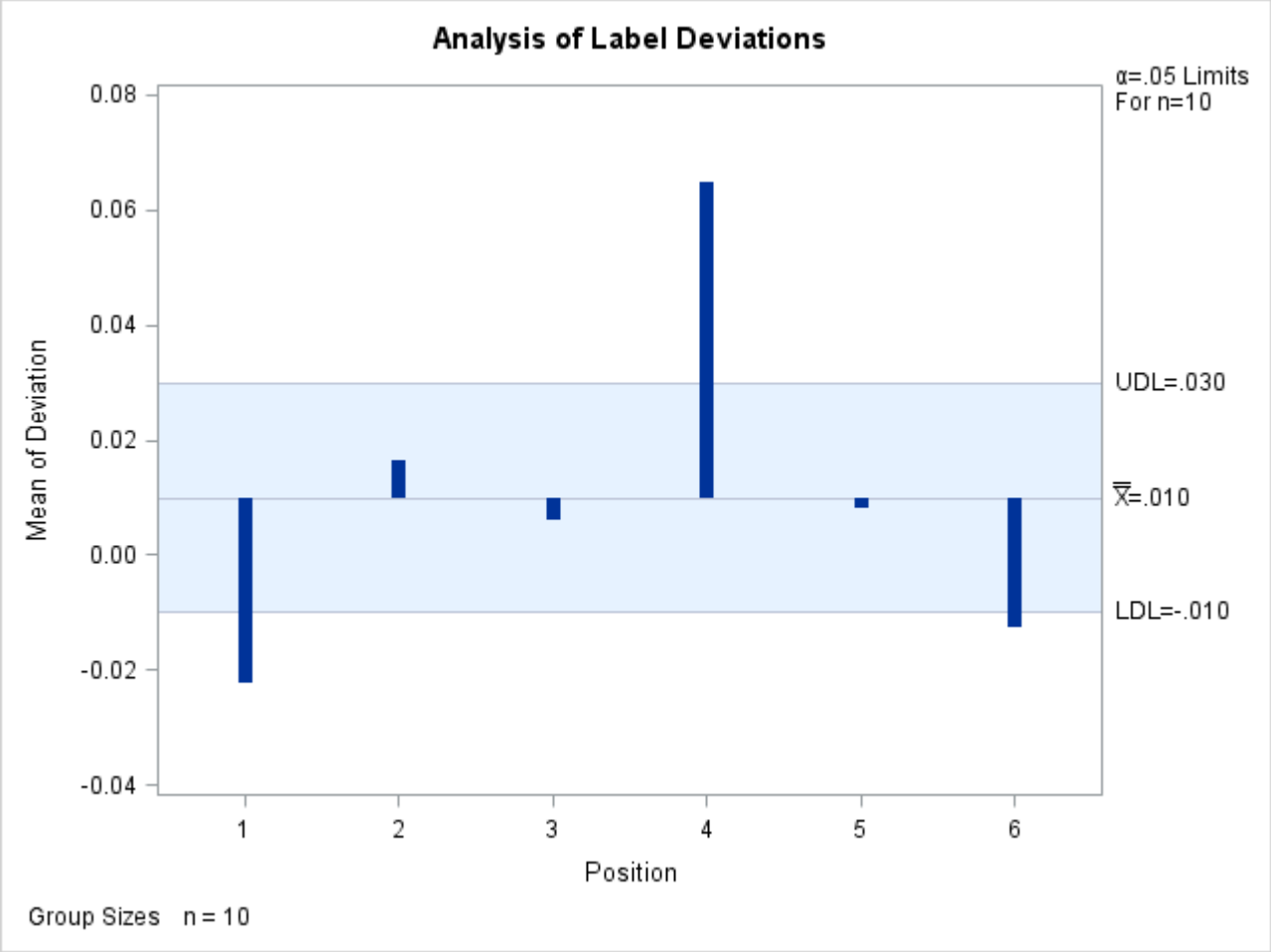
You can read this data set by specifying it as a SUMMARY= data set in the PROC ANOM statement, as follows:

```
title 'Analysis of Label Deviations';
proc anom summary=Labels;
   xchart Deviation*Position / odstitle=title1;
run;
```

The resulting chart is shown in Figure 4.24.

Note that Deviation 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 DeviationX, DeviationS, and DeviationN. The suffix characters *X*, *S*, and *N* indicate *mean*, *standard deviation*, and *sample size*, respectively. Thus, you can specify three group summary variables in a SUMMARY= data set with a single name (Deviation), which is referred to as the *response*. The name Position specified after the asterisk is the name of the *group-variable*.

Figure 4.24 ANOM Chart for Means in Data Set Labels



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

- group variable
- group mean variable
- group standard deviation variable
- group sample size variable

Furthermore, the names of the group mean, standard deviation, and sample size variables must begin with the *response* name specified in the XCHART 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 ANOM statement to rename the variables for the duration of the ANOM procedure step. If a label is associated with the group mean variable, it is used to label the vertical axis.

In summary, the interpretation of *response* depends on the input data set.

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

For more information, see the section "SUMMARY= Data Set" on page 151.

Saving Summary Statistics for Groups

NOTE: See *Saving Summary Statistics for Groups Using ANOM Charts* in the SAS/QC Sample Library.

In this example, the XCHART statement is used to create a data set containing group summary statistics that can be read later by the ANOM procedure (as in the preceding example). The following statements read measurements from the data set LabelDeviations and create a summary data set named LabelSummary:

```
proc anom data=LabelDeviations;
   xchart Deviation*Position / outsummary=LabelSummary
                               nochart;
run;
```

The OUTSUMMARY= 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 4.22.

Figure 4.25 contains a listing of LabelSummary.

Figure 4.25 The Summary Data Set LabelSummary

The Data Set LabelSummary

Position	DeviationX	DeviationS	DeviationN
1	-0.022342	0.022805	10
2	0.016241	0.033478	10
3	0.006011	0.028847	10
4	0.064729	0.021492	10
5	0.008121	0.025920	10
6	-0.012812	0.015974	10

There are four variables in the data set LabelSummary.

- Position identifies the group.
- DeviationX contains the group means.
- DeviationS contains the group standard deviations.
- DeviationN contains the group sizes.

Note that the summary statistic variables are named by adding the suffix characters *X*, *S*, and *N* to the *response* Deviation specified in the XCHART statement. In other words, the variable naming convention for OUTSUMMARY= data sets is the same as that for SUMMARY= data sets.

For more information, see the section "OUTSUMMARY= Data Set" on page 147.

Saving Decision Limits

NOTE: See *Saving Decision Limits Using ANOM Charts for Means* in the SAS/QC Sample Library.

You can save the decision limits for an ANOM chart, together with the parameters used to compute the limits, in a SAS data set.

The following statements read measurements from the data set LabelDeviations (see the section "Creating ANOM Charts for Means from Response Values" on page 127) and save the decision limits displayed in Figure 4.22 in a data set named LabelLimits:

```
proc anom data=LabelDeviations;
   xchart Deviation*Position / outlimits=LabelLimits
                               nochart;
run;
```

The OUTLIMITS= option names the data set containing the decision limits, and the NOCHART option suppresses the display of the chart. The data set LabelLimits is listed in Figure 4.26.

Figure 4.26 The Data Set LabelLimits Containing Decision Limit Information

Decision Limits for Labler Position Deviations

VAR	_GROUP_	_TYPE_	_LIMITN_	_ALPHA_	_LDLX_	_MEAN_	_UDLX_	_MSE_	_DFE_	_LIMITK_
Deviation	Position	ESTIMATE	10	0.05	-.009878975	.009991333	0.029862	.000643646	54	6

The data set LabelLimits contains one observation with the limits for *response* Deviation. The values of _LDLX_ and _UDLX_ are the lower and upper decision limits for the means, and the value of _MEAN_ is the weighted average of the group means, which is represented by the central line.

The values of _MEAN_, _MSE_, _DFE_, _LIMITN_, _LIMITK_, and _ALPHA_ are the parameters used to compute the decision limits as described in the section "Constructing ANOM Charts for Means" on page 143. The value of _MSE_ is the mean square error, and the value of _DFE_ is the associated degrees of freedom. The value of _LIMITN_ is the nominal sample size (n) associated with the decision limits, the value of _LIMITK_ is the number of groups (k), and the value of _ALPHA_ is the value of the significance level (α). The variables _VAR_ and _GROUP_ are bookkeeping variables that save the *response* and *group-variable*. The variable _TYPE_ is a bookkeeping variable that indicates whether the values of _MEAN_ and _MSE_ are estimates computed from the data or standard (known) values specified with procedure options. In most applications, the value of _TYPE_ will be 'ESTIMATE.'

NOTE: See *Saving Summary Statistics & Decision Limits Using ANOM Charts* in the SAS/QC Sample Library.

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

```
proc anom data=LabelDeviations;
   xchart Deviation*Position / outtable=LabelTab
                               nochart;
run;
```

The data set LabelTab is listed in Figure 4.27.

Figure 4.27 The Data Set LabelTab

Summary Statistics and Decision Limits

VAR	Position	_ALPHA_	_LIMITN_	_SUBN_	_LDLX_	_SUBX_	_MEAN_	_UDLX_	_EXLIM_
Deviation	1	0.05	10	10	-.009878975	-0.022342	.009991333	0.029862	LOWER
Deviation	2	0.05	10	10	-.009878975	0.016241	.009991333	0.029862	
Deviation	3	0.05	10	10	-.009878975	0.006011	.009991333	0.029862	
Deviation	4	0.05	10	10	-.009878975	0.064729	.009991333	0.029862	UPPER
Deviation	5	0.05	10	10	-.009878975	0.008121	.009991333	0.029862	
Deviation	6	0.05	10	10	-.009878975	-0.012812	.009991333	0.029862	LOWER

This data set contains one observation for each group sample. The variables _SUBX_ and _SUBN_ contain the group means and sample sizes. The variables _LDLX_ and _UDLX_ contain the lower and upper decision limits, and the variable _MEAN_ contains the central line. The variables _VAR_ and Position contain the *response* name and values of the *group-variable*, respectively. For more information, see the section "OUTTABLE= Data Set" on page 148.

An OUTTABLE= data set can be read later as a TABLE= data set. For example, the following statements read LabelTab and display an ANOM chart (not shown here) identical to the chart in Figure 4.22:

```
title 'Analysis of Label Deviations';
proc anom table=LabelTab;
   xchart Deviation*Position;
label _SUBX_ = 'Mean Deviation from Center (mm)';
run;
```

Because the ANOM procedure simply displays the information in a TABLE= data set, you can use TABLE= data sets to create specialized ANOM charts.

For more information, see the section "TABLE= Data Set" on page 152.

Syntax: XCHART Statement

The basic syntax for the XCHART statement is as follows:

XCHART *response* ∗ *group-variable* **;**

The general form of this syntax is as follows:

XCHART *responses* ∗ *group-variable* < **(***block-variables***)** >
< **=***symbol-variable* | **=***'character'* > < *options* > **;**

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

response

responses

identify one or more responses to be analyzed. The specification of *response* depends on the input data set specified in the PROC ANOM statement.

- If response values (raw data) are read from a DATA= data set, *response* must be the name of the variable containing the values. For an example, see the section "Creating ANOM Charts for Means from Response Values" on page 127.
- If summary data are read from a SUMMARY= data set, *response* must be the common prefix of the summary variables in the SUMMARY= data set. For an example, see the section "Creating ANOM Charts for Means from Group Summary Data" on page 130.
- If summary data and decision limits are read from a TABLE= data set, *response* must be the value of the variable _VAR_ in the TABLE= data set. For an example, see the section "Saving Decision Limits" on page 133.

A *response* is required. If you specify more than one response, enclose the list in parentheses. For example, the following statements request distinct ANOM charts for the means of Weight, Length, and Width:

```
proc anom data=Measures;
   xchart (Weight Length Width)*Day;
run;
```

group-variable

is the variable that identifies groups in the data. The *group-variable* is required. In the preceding XCHART statement, Day is the group variable.

block-variables

are optional variables that group the data into blocks of consecutive groups. The blocks are labeled in a legend, and each *block-variable* provides one level of labels in the legend.

symbol-variable

is an optional variable whose levels (unique values) determine the symbol marker used to plot the means. 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.

options

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

Summary of Options

The following tables list the XCHART statement options by function. Options unique to the ANOM procedure are listed in Table 4.21, and are described in detail in "Dictionary of ANOM Chart Statement Options" on page 180. Options that are common to both the ANOM and SHEWHART procedures are listed in Table 4.22, and are described in detail in "Dictionary of Options: SHEWHART Procedure" on page 1946.

Table 4.21 XCHART Statement Special Options

Option	Description
Options for Specifying Parameters for Decision Limits	
ALPHA=	specifies the probability of a Type I error
DFE=	specifies the degrees of freedom associated with the root mean square error
LIMITK=	specifies number of groups for decision limits
LIMITN=	specifies either nominal sample size for fixed decision limits or varying limits
MEAN=	specifies the mean
MSE=	specifies the mean square error
NOREADLIMITS	computes decision limits for each *response* from the data rather than a LIMITS= data set
READINDEXES=	reads multiple sets of decision limits for each *response* from a LIMITS= data set
TYPE=	identifies parameters as estimates or standard values and specifies value of _TYPE_ in the OUTLIMITS= data set
Options for Displaying Decision Limits	
CINFILL=	specifies color for area inside decision limits
CLIMITS=	specifies color of decision limits, central line, and related labels
LDLLABEL=	specifies label for lower decision limit
LIMLABSUBCHAR=	specifies a substitution character for labels provided as quoted strings; the character is replaced with the value of the decision limit
LLIMITS=	specifies line type for decision limits
NDECIMAL=	specifies number of digits to right of decimal place in default labels for decision limits and central line
NOCTL	suppresses display of central line
NOLDL	suppresses display of lower decision limit
NOLIMITLABEL	suppresses labels for decision limits and central line
NOLIMITS	suppresses display of decision limits
NOLIMITSFRAME	suppresses default frame around decision limit information when multiple sets of decision limits are read from a LIMITS= data set
NOLIMITSLEGEND	suppresses legend for decision limits
NOUDL	suppresses display of upper decision limit
UDLLABEL=	specifies label for upper decision limit
WLIMITS=	specifies width for decision limits and central line
XSYMBOL=	specifies label for central line
Output Data Set Option	
OUTSUMMARY=	creates output data set containing group summary statistics

Table 4.22 XCHART Statement General Options

Option	Description
Options for Plotting and Labeling Points	
ALLLABEL=	labels every point on ANOM chart
CLABEL=	specifies color for labels
CCONNECT=	specifies color for line segments that connect points on chart
CFRAMELAB=	specifies fill color for frame around labeled points
CNEEDLES=	specifies color for needles that connect points to central line
COUT=	specifies color for portions of line segments that connect points outside decision limits
COUTFILL=	specifies color for shading areas between the connected points and decision limits outside the limits
LABELANGLE=	specifies angle at which labels are drawn
LABELFONT=	specifies software font for labels
LABELHEIGHT=	specifies height of labels
NONEEDLES	suppresses vertical needles connecting points to central line
OUTLABEL=	labels points outside decision limits
SYMBOLLEGEND=	specifies LEGEND statement for levels of *symbol-variable*
SYMBOLORDER=	specifies order in which symbols are assigned for levels of *symbol-variable*
TURNALL|TURNOUT	turns point labels so that they are strung out vertically
WNEEDLES=	specifies width of needles
Axis and Axis Label Options	
CAXIS=	specifies color for axis lines and tick marks
CFRAME=	specifies fill colors for frame for plot area
CTEXT=	specifies color for tick mark values and axis labels
DISCRETE	produces horizontal axis for discrete numeric group values
HAXIS=	specifies major tick mark values for horizontal axis
HEIGHT=	specifies height of axis label and axis legend text
HMINOR=	specifies number of minor tick marks between major tick marks on horizontal axis
HOFFSET=	specifies length of offset at both ends of horizontal axis
NOHLABEL	suppresses label for horizontal axis
NOTICKREP	specifies that only the first occurrence of repeated, adjacent group values is to be labeled on horizontal axis
NOVANGLE	requests vertical axis labels that are strung out vertically
SKIPHLABELS=	specifies thinning factor for tick mark labels on horizontal axis
TURNHLABELS	requests horizontal axis labels that are strung out vertically

Table 4.22 *continued*

Option	Description
VAXIS=	specifies major tick mark values for vertical axis of ANOM chart
VFORMAT=	specifies format for vertical axis tick mark labels
VMINOR=	specifies number of minor tick marks between major tick marks on vertical axis
VOFFSET=	specifies length of offset at both ends of vertical axis
VZERO	forces origin to be included in vertical axis for ANOM chart
WAXIS=	specifies width of axis lines
Plot Layout Options	
ALLN	plots means for all groups
BILEVEL	creates ANOM charts using half-screens and half-pages
EXCHART	creates ANOM chart for a response only when a group mean exceeds the decision limits
INTERVAL=	natural time interval between consecutive group positions when time, date, or datetime format is associated with a numeric group variable
MAXPANELS=	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 decision limits
NOCHART	suppresses creation of chart
NOFRAME	suppresses frame for plot area
NOLEGEND	suppresses legend for group sample sizes
NPANELPOS=	specifies number of group positions per panel on each chart
REPEAT	repeats last group position on panel as first group position of next panel
TOTPANELS=	specifies number of pages or screens to be used to display chart
ZEROSTD	displays ANOM chart regardless of whether root mean square error is zero
Reference Line Options	
CHREF=	specifies color for lines requested by HREF= option
CVREF=	specifies color for lines requested by VREF= option
HREF=	specifies position of reference lines perpendicular to horizontal axis on ANOM chart
HREFDATA=	specifies position of reference lines perpendicular to horizontal axis on ANOM chart
HREFLABELS=	specifies labels for HREF= lines
HREFLABPOS=	specifies position of HREFLABELS= labels
LHREF=	specifies line type for HREF= lines
LVREF=	specifies line type for VREF= lines

Table 4.22 *continued*

Option	Description
NOBYREF	specifies that reference line information in a data set applies uniformly to charts created for all BY groups
VREF=	specifies position of reference lines perpendicular to vertical axis on ANOM chart
VREFLABELS=	specifies labels for VREF= lines
VREFLABPOS=	specifies position of VREFLABELS= labels
Grid Options	
CGRID=	specifies color for grid requested with GRID or ENDGRID option
ENDGRID	adds grid after last plotted point
GRID	adds grid to control chart
LENDGRID=	specifies line type for grid requested with the ENDGRID option
LGRID=	specifies line type for grid requested with the GRID option
WGRID=	specifies width of grid lines
Clipping Options	
CCLIP=	specifies color for plot symbol for clipped points
CLIPFACTOR=	determines extent to which extreme points are clipped
CLIPLEGEND=	specifies text for clipping legend
CLIPLEGPOS=	specifies position of clipping legend
CLIPSUBCHAR=	specifies substitution character for CLIPLEGEND= text
CLIPSYMBOL=	specifies plot symbol for clipped points
CLIPSYMBOLHT=	specifies symbol marker height for clipped points
Graphical Enhancement Options	
ANNOTATE=	specifies annotate data set that adds features to ANOM chart
DESCRIPTION=	specifies description of ANOM chart's GRSEG catalog entry
FONT=	specifies software font for labels and legends on chart
NAME=	specifies name of ANOM chart's GRSEG catalog entry
PAGENUM=	specifies the form of the label used in pagination
PAGENUMPOS=	specifies the position of the page number requested with the PAGENUM= option
Options for Producing Graphs Using ODS Styles	
BLOCKVAR=	specifies one or more variables whose values define colors for filling background of *block-variable* legend
CFRAMELAB	draws a frame around labeled points
COUT	draw portions of line segments that connect points outside decision limits in a contrasting color
CSTAROUT	specifies that portions of stars exceeding inner or outer circles are drawn using a different color

Table 4.22 *continued*

Option	Description
OUTFILL	shades areas between decision limits and connected points lying outside the limits
STARFILL=	specifies a variable identifying groups of stars filled with different colors
STARS=	specifies a variable identifying groups of stars whose outlines are drawn with different colors
Options for ODS Graphics	
BLOCKREFTRANSPARENCY=	specifies the wall fill transparency for blocks and phases
INFILLTRANSPARENCY=	specifies the decision limit infill transparency
MARKERS	plots group points with markers
NOBLOCKREF	suppresses block and phase reference lines
NOBLOCKREFFILL	suppresses block and phase wall fills
NOFILLLEGEND	suppresses legend for levels of a STARFILL= variable
NOPHASEREF	suppresses block and phase reference lines
NOPHASEREFFILL	suppresses block and phase wall fills
NOREF	suppresses block and phase reference lines
NOREFFILL	suppresses block and phase wall fills
NOSTARFILLLEGEND	suppresses legend for levels of a STARFILL= variable
NOTRANSPARENCY	disables transparency in ODS Graphics output
ODSFOOTNOTE=	specifies a graph footnote
ODSLEGENDEXPAND	specifies that legend entries contain all levels observed in the data
ODSTITLE=	specifies a graph title
OUTFILLTRANSPARENCY=	specifies decision limit outfill transparency
OVERLAYURL=	specifies URLs to associate with overlay points
PHASEPOS=	specifies vertical position of phase legend
PHASEREFLEVEL=	associates phase and block reference lines with either innermost or the outermost level
PHASEREFTRANSPARENCY=	specifies the wall fill transparency for blocks and phases
REFFILLTRANSPARENCY=	specifies the wall fill transparency for blocks and phases
SIMULATEQCFONT	draws central line labels using a simulated software font
STARTRANSPARENCY=	specifies star fill transparency
URL=	specifies a variable whose values are URLs to be associated with groups
Input Data Set Options	
MISSBREAK	specifies that observations with missing values are not to be processed
Output Data Set Options	
OUTINDEX=	specifies value of _INDEX_ in the OUTLIMITS= data set
OUTLIMITS=	creates output data set containing decision limits
OUTTABLE=	creates output data set containing group summary statistics and decision limits

Table 4.22 *continued*

Option	Description
Tabulation Options	
NOTE: specifying (EXCEPTIONS) after a tabulation option creates a table for exceptional points only.	
TABLE	creates a basic table of group means, group sample sizes, and decision 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
TABLEOUTLIM	augments basic table with columns indicating decision limits exceeded
Block Variable Legend Options	
BLOCKLABELPOS=	specifies position of label for *block-variable* legend
BLOCKLABTYPE=	specifies text size of *block-variable* legend
BLOCKPOS=	specifies vertical position of *block-variable* legend
BLOCKREP	repeats identical consecutive labels in *block-variable* legend
CBLOCKLAB=	specifies fill colors for frames enclosing variable labels in *block-variable* legend
CBLOCKVAR=	specifies one or more variables whose values are colors for filling background of *block-variable* legend
Phase Options	
CPHASELEG=	specifies text color for *phase* legend
NOPHASEFRAME	suppresses default frame for *phase* legend
OUTPHASE=	specifies value of _PHASE_ in the OUTHISTORY= data set
PHASEBREAK	disconnects last point in a *phase* from first point in next *phase*
PHASELABTYPE=	specifies text size of *phase* legend
PHASELEGEND	displays *phase* labels in a legend across top of chart
PHASELIMITS	labels decision limits for each phase, provided they are constant within that phase
PHASEREF	delineates *phases* with vertical reference lines
READPHASES=	specifies *phases* to be read from an input data set
Star Options	
CSTARCIRCLES=	specifies color for STARCIRCLES= circles
CSTARFILL=	specifies color for filling stars
CSTAROUT=	specifies outline color for stars exceeding inner or outer circles
CSTARS=	specifies color for outlines of stars
LSTARCIRCLES=	specifies line types for STARCIRCLES= circles
LSTARS=	specifies line types for outlines of STARVERTICES= stars
STARBDRADIUS=	specifies radius of outer bound circle for vertices of stars

Table 4.22 *continued*

Option	Description
STARCIRCLES=	specifies reference circles for stars
STARINRADIUS=	specifies inner radius of stars
STARLABEL=	specifies vertices to be labeled
STARLEGEND=	specifies style of legend for star vertices
STARLEGENDLAB=	specifies label for STARLEGEND= legend
STAROUTRADIUS=	specifies outer radius of stars
STARSPECS=	specifies method used to standardize vertex variables
STARSTART=	specifies angle for first vertex
STARTYPE=	specifies graphical style of star
STARVERTICES=	superimposes star at each point on ANOM chart
WSTARCIRCLES=	specifies width of STARCIRCLES= circles
WSTARS=	specifies width of STARVERTICES= stars
Overlay Options	
CCOVERLAY=	specifies colors for overlay line segments
COVERLAY=	specifies colors for overlay plots
COVERLAYCLIP=	specifies color for clipped points on overlays
LOVERLAY=	specifies line types for overlay line segments
NOOVERLAYLEGEND	suppresses legend for overlay plots
OVERLAY=	specifies variables to overlay on chart
OVERLAYCLIPSYM=	specifies symbol for clipped points on overlays
OVERLAYCLIPSYMHT=	specifies symbol height for clipped points on overlays
OVERLAYHTML=	specifies links to associate with overlay points
OVERLAYID=	specifies labels for overlay points
OVERLAYLEGLAB=	specifies label for overlay legend
OVERLAYSYM=	specifies symbols for overlays
OVERLAYSYMHT=	specifies symbol heights for overlays
WOVERLAY=	specifies widths of overlay line segments
Options for Interactive ANOM Charts	
HTML=	specifies a variable whose values create links to be associated with groups
HTML_LEGEND=	specifies a variable whose values create links to be associated with symbols in the symbol legend
WEBOUT=	creates an OUTTABLE= data set with additional graphics coordinate data

Details: XCHART Statement

Constructing ANOM Charts for Means

The following notation is used in this section:

X_{ij}	jth response in the ith group
k	number of groups
n_i	sample size of ith group
N	total sample size $= n_1 + \cdots + n_k$
μ_i	expected value of a response in the ith group
σ	standard deviation of a response
$\bar{X}_i$	average response in the ith group
$\overline{\overline{X}}$	weighted average of k group means
s_i^2	sample variance of the responses in the ith group
$\widehat{\sigma^2}$	mean square error (MSE)
ν	degrees of freedom associated with the mean square error
α	significance level
$h(\alpha; k, n, \nu)$	critical value for analysis of means when the sample sizes n_i are equal ($n_i \equiv n$)
$h(\alpha; k, n_1, \ldots, n_k, \nu)$	critical value for analysis of means when the sample sizes n_i are not equal

Plotted Points

Each point on an ANOM chart indicates the value of a group mean ($\bar{X}_i$).

Central Line

By default, the central line on an ANOM chart for means represents the weighted average of the group means, which is computed as

$$\overline{\overline{X}} = \frac{n_1 \bar{X}_1 + \cdots + n_k \bar{X}_k}{n_1 + \cdots + n_k}$$

You can specify a value for $\overline{\overline{X}}$ with the MEAN= option in the XCHART statement or with the variable _MEAN_ in a LIMITS= data set.

Decision Limits

In the analysis of means for continuous data, it is assumed that the responses in the ith group are at least approximately normally distributed with a constant variance:

$$X_{ij} \sim N(\mu_i, \sigma^2), \qquad j = 1, \ldots, n_i$$

When the group sizes are constant ($n_i \equiv n$), then $\nu = N - k = k(n-1)$ and the decision limits are computed as follows:

$$\begin{aligned}
\text{lower decision limit (LDL)} &= \overline{\overline{X}} - h(\alpha; k, n, \nu)\sqrt{\text{MSE}}\sqrt{\frac{k-1}{N}} \\
\text{upper decision limit (UDL)} &= \overline{\overline{X}} + h(\alpha; k, n, \nu)\sqrt{\text{MSE}}\sqrt{\frac{k-1}{N}}
\end{aligned}$$

Here the mean square error (MSE) is computed as follows:

$$\text{MSE} = \widehat{\sigma^2} = \frac{1}{k}\sum_{j=1}^{k} s_j^2$$

For details concerning the function $h(\alpha; k, n, \nu)$, see Nelson (1982a, 1993).

When the group sizes are not constant (the unbalanced case), $\nu = N - k$ and the decision limits for the ith group are computed as follows:

$$\begin{aligned}
\text{lower decision limit (LDL)} &= \overline{\overline{X}} - h(\alpha; k, n_1, \ldots, n_k, \nu)\sqrt{\text{MSE}}\sqrt{\frac{N-n_i}{Nn_i}} \\
\text{upper decision limit (UDL)} &= \overline{\overline{X}} + h(\alpha; k, n_1, \ldots, n_k, \nu)\sqrt{\text{MSE}}\sqrt{\frac{N-n_i}{Nn_i}}
\end{aligned}$$

Here the mean square error (MSE) is computed as follows:

$$\text{MSE} = \widehat{\sigma^2} = \frac{(n_1-1)s_1^2 + \cdots + (n_k-1)s_k^2}{n_1 + \cdots + n_k - k}$$

This requires that ν be positive. A chart is not produced if $\nu > 0$ but MSE is equal to zero (unless you specify the ZEROSTD option). For details concerning the function $h(\alpha; k, n_1, \ldots, n_k, \nu)$, see Nelson (1991).

You can specify parameters for the limits as follows:

- Specify α with the ALPHA= option or with the variable _ALPHA_ in a LIMITS= data set. By default, $\alpha = 0.05$.
- Specify a constant nominal sample size $n_i \equiv n$ for the decision limits in the balanced case with the LIMITN= option or with the variable _LIMITN_ in a LIMITS= data set. By default, n is the observed sample size in the balanced case.
- Specify k with the LIMITK= option or with the variable _LIMITK_ in a LIMITS= data set. By default, k is the number of groups.
- Specify $\overline{\overline{X}}$ with the MEAN= option or with the variable _MEAN_ in a LIMITS= data set. By default, $\overline{\overline{X}}$ is the weighted average of the responses.

- Specify $\widehat{\sigma^2}$ with the MSE= option or with the variable _MSE_ in a LIMITS= data set. By default, $\widehat{\sigma^2}$ is computed as indicated above.
- Specify ν with the DFE= option or with the variable _DFE_ in a LIMITS= data set. By default, ν is determined as indicated above.

Constructing ANOM Charts for Two-Way Layouts

This section provides the computational details for constructing an ANOM chart for the lth factor in an experiment involving two factors ($l = 1$ or 2). It is assumed that there is no interaction effect. See Example 4.5 for an illustration.

The following notation is used in this section:

X_{ijk}	kth response at the ith level of factor 1 and the jth level of factor 2, where $k = 1, 2, \ldots, n_{ij}$
f_l	number of groups (levels) for the lth factor, $l = 1, 2$
n_{ij}	number of replicates in cell (i, j)
N	total sample size $= \sum_{i=1}^{f_1} \sum_{j=1}^{f_2} n_{ij}$
σ^2	variance of a response
$\overline{X}_{ij\cdot}$	average response in cell (i, j)
$\overline{X}_{i\cdot\cdot}$	average response for ith level of factor 1 $= \left(\sum_{j=1}^{f_2} n_{ij} \overline{X}_{ij\cdot}\right) / \left(\sum_{j=1}^{f_2} n_{ij}\right)$
$\overline{X}_{\cdot j\cdot}$	average response for jth level of factor 2 $= \left(\sum_{i=1}^{f_1} n_{ij} \overline{X}_{ij\cdot}\right) / \left(\sum_{i=1}^{f_1} n_{ij}\right)$
$\overline{\overline{X}}$	$\sum_{i=1}^{f_1} \sum_{j=1}^{f_2} n_{ij} \overline{X}_{ij} / N$
s_{ij}^2	sample variance of the responses for the ith level of factor 1 and the jth level of factor 2
$\widehat{\sigma^2}$	mean square error (MSE) in the two-way analysis of variance
ν	degrees of freedom associated with the mean square error in the two-way analysis of variance
α	significance level
$h(\alpha; f_l, n, \nu)$	critical value for analysis of means in a one-way layout for f_l groups (treatment levels) when the sample sizes for each level are constant ($\equiv n$) and ν is the degrees of freedom associated with the mean square error; see the section "Constructing ANOM Charts for Means" on page 143.

Plotted Points

The points on the ANOM chart for factor 1 represent $\overline{X}_{i\cdot\cdot}$, $i = 1, \ldots, f_1$ and the points on the ANOM chart for factor 2 represent $\overline{X}_{\cdot j\cdot}$, $j = 1, \ldots, f_2$.

Central Line

The central line on the ANOM chart for the lth factor is the overall weighted average $\overline{\overline{X}}$. Some authors use the notation $\overline{X}_{\cdots}$ for this average.

Decision Limits

It is assumed that

$$X_{ijk} = \mu + \alpha_i + \beta_j + \epsilon_{ijk}$$

where the quantities ϵ_{ijk} are independent and at least approximately normally distributed with

$$\epsilon_{ijk} \sim N(0, \sigma^2)$$

The correct decision limits for a given factor in a two-way layout are not computed by default when the lth factor is specified as the *group-variable* in the XCHART statement, since the mean square error and degrees of freedom are not adjusted for the two-way structure of the data. Consequently, $\widehat{\sigma^2}$ and ν must be precomputed and provided to the ANOM procedure, as illustrated in Example 4.5.

In the case of a two-way layout with equal group sizes ($n_{ij} \equiv n$), the appropriate decision limits are:

$$\begin{aligned} \text{lower decision limit (LDL)} &= \overline{\overline{X}} - h(\alpha; f_l, n, \nu)\sqrt{\text{MSE}}\sqrt{\frac{f_l - 1}{N}} \\ \text{upper decision limit (UDL)} &= \overline{\overline{X}} + h(\alpha; f_l, n, \nu)\sqrt{\text{MSE}}\sqrt{\frac{f_l - 1}{N}} \end{aligned}$$

where the mean square error (MSE) is computed as in the ANOVA or GLM procedure:

$$\text{MSE} = \widehat{\sigma^2} = \frac{1}{f_1 f_2} \sum_{i=1}^{f_1} \sum_{j=1}^{f_2} s_{ij}^2$$

and the degrees of freedom for error is $\nu = f_1 f_2 (n - 1)$. For details concerning the function $h(\alpha; f_l, n, \nu)$, see Nelson (1982a, 1993).

You can provide the appropriate values of MSE and ν by

- specifying $\widehat{\sigma^2}$ with the MSE= option or with the variable _MSE_ in a LIMITS= data set
- specifying ν with the DFE= option or with the variable _DFE_ in a LIMITS= data set

In addition you can:

- Specify α with the ALPHA= option or with the variable _ALPHA_ in a LIMITS= data set. By default, $\alpha = 0.05$.
- Specify a constant nominal sample size $n_{ij} \equiv n$ for the decision limits in the balanced case with the LIMITN= option or with the variable _LIMITN_ in a LIMITS= data set.
- Specify f_l with the LIMITK= option or with the variable _LIMITK_ in a LIMITS= data set.
- Specify $\overline{\overline{X}}$ with the MEAN= option or with the variable _MEAN_ in a LIMITS= data set.

Output Data Sets

OUTLIMITS= Data Set

The OUTLIMITS= data set saves decision limits and decision limit parameters. The following variables can be saved:

Table 4.23 OUTLIMITS= Data Set

Variable	Description
ALPHA	significance level (α).
DFE	degrees of freedom for mean square error (ν).
GROUP	*group-variable* specified in the XCHART statement
INDEX	optional identifier for the decision limits specified with the OUTINDEX= option
LDLX	lower decision limit for group means
LIMITK	number of groups
LIMITN	group sample size associated with the decision limits
MEAN	weighted average of group means ($\overline{\overline{X}}$)
MSE	mean square error ($\widehat{\sigma^2}$).
TYPE	type (estimate or standard value) of _MEAN_ and _MSE_
UDLX	upper decision limit for group means
VAR	*response* specified in the XCHART statement

Notes:

1. In the unbalanced case, the special missing value V is assigned to the variables _LIMITN_, _LDLX_, and _UDLX_ to indicate that the decision limits vary with the group sample size.
2. Optional BY variables are saved in the OUTLIMITS= data set.

The OUTLIMITS= data set contains one observation for each *response* specified in the XCHART statement. For an example, see the section "Saving Decision Limits" on page 133.

OUTSUMMARY= Data Set

The OUTSUMMARY= data set saves group summary statistics. The following variables can be saved:

- the *group-variable*
- a group mean variable named by *response* suffixed with X
- a group sample size variable named by *response* suffixed with N
- a group standard deviation variable named by *response* suffixed with S

Given a *response* name that contains 32 characters, the procedure first shortens the name to its first 16 characters and its last 15 characters, and then it adds the suffix.

Group summary variables are created for each *response* specified in the XCHART statement. For example, consider the following statements:

```
proc anom data=Steel;
   xchart (Width Diameter)*Lot / outsummary=Summary;
run;
```

The data set Summary contains variables named Lot, WidthX, WidthS, WidthN, DiameterX, DiameterS, and DiameterN. 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 OUTSUMMARY= data set, see the section "Saving Summary Statistics for Groups" on page 132.

OUTTABLE= Data Set

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

Variable	Description
ALPHA	significance level (α)
EXLIM	decision limit exceeded (if any)
group	values of the group variable
LDLX	lower decision limit for group mean
LIMITN	nominal sample size associated with the decision limits
MEAN	central line
SUBN	group sample size
SUBX	group mean
UDLX	upper decision limit for group mean
VAR	*response* specified in the XCHART 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)

NOTE: The variable _EXLIM_ is a character variable of length 8. The variable _PHASE_ is a character variable of length 48. The variable _VAR_ is a character variable whose length is no greater than 32. All other variables are numeric.

For an example, see the section "Saving Decision Limits" on page 133.

ODS Tables

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

Table 4.24 ODS Tables Produced with the XCHART Statement

Table Name	Description	Options
XCHART	ANOM chart summary statistics	TABLE, TABLEALL, TABLEC, TABLEID, TABLEOUT,

ODS Graphics

Before you create ODS Graphics output, ODS Graphics must be enabled (for example, by using the ODS GRAPHICS ON statement). For more information about enabling and disabling ODS Graphics, see the section "Enabling and Disabling ODS Graphics" (Chapter 21, *SAS/STAT User's Guide*). **NOTE:** In SAS/QC 13.1 the ANOM procedure does not support the creation of graphs that are editable with the ODS Graphics Editor.

The appearance of a graph produced with ODS Graphics is determined by the style associated with the ODS destination where the graph is produced. XCHART options used to control the appearance of traditional graphics are ignored for ODS Graphics output. Options for Producing Graphs Using ODS Styles lists options that can be used to control the appearance of graphs produced with ODS Graphics or with traditional graphics using ODS styles. Options for ODS Graphics lists options to be used exclusively with ODS Graphics. Detailed descriptions of these options are provided in "Dictionary of Options: SHEWHART Procedure" on page 1946.

When ODS Graphics is in effect, the XCHART statement assigns a name to the graph it creates. You can use this name to reference the graph when using ODS. The name is listed in Table 4.25.

Table 4.25 ODS Graphics Produced by the XCHART Statement

ODS Graph Name	Plot Description
XChart	ANOM chart for means

See Chapter 3, "SAS/QC Graphics," for more information about ODS Graphics and other methods for producing charts.

Input Data Sets

DATA= Data Set

You can read raw data (response values) from a DATA= data set specified in the PROC ANOM statement. Each *response* specified in the XCHART statement must be a SAS variable in the DATA= data set. This variable provides measurements that must be grouped into group samples indexed by the *group-variable*. The *group-variable*, which is specified in the XCHART 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 *response* and a value for the *group-variable*. If the ith group contains n_i items, there should be n_i consecutive observations for which the value of the *group-variable* is the index of the ith group. For example, if each group contains five items and there are 10 groups, the DATA= data set should contain 50 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 ANOM 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 of a DATA= data set, see the section "Creating ANOM Charts for Means from Response Values" on page 127.

LIMITS= Data Set

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

```
proc anom data=Info limits=Conlims;
   xchart Weight*Batch;
run;
```

The LIMITS= data set can be an OUTLIMITS= data set that was created in a previous run of the ANOM procedure. Such data sets always contain the variables required for a LIMITS= data set; see Table 4.23. 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 minimal combinations of variables:

- the variables _LDLX_, _MEAN_, and _UDLX_, which specify the decision limits directly
- the variables _MEAN_ and _MSE_, with _DFE_ recommended, which are used to calculate the decision limits according to the equations in the section "Decision Limits" on page 143

In addition, note the following:

- The variables _VAR_ and _GROUP_ are always required. These must be character variables whose lengths are no greater than 32.
- _DFE_ is optional. The default is $\nu = N - k$, and in the case of equal group sizes, $\nu = k(n - 1)$.
- _MSE_ is optional if _LDLX_ and _UDLX_ are specified; otherwise it is required.
- _LDLX_ and _UDLX_ must be specified together; otherwise their values are computed.
- _ALPHA_ is optional but is recommended in order to maintain a complete set of decision limit information. The default value is 0.05.

- _LIMITK_ is optional. The default value is k, the number of groups. A group must have at least one non-missing value ($n_i \geq 1$) and there must be at least one group with $n_i \geq 2$. If specified, _LIMITK_ overrides the value of k.
- _LIMITN_ is optional. The default value is the common group size (n), in the balanced case $n_i \equiv n$. If specified, _LIMITN_ overrides the value of n.
- The variable _TYPE_ is optional, but is recommended to maintain a complete set of decision limit information. The variable _TYPE_ must be a character variable of length 8. Valid values are 'ESTIMATE,' 'STANDARD,' 'STDMEAN,' and 'STDRMS.' The default is 'ESTIMATE.'
- The variable _INDEX_ is required if you specify the READINDEX= option; this must be a character variable whose length is no greater than 48.
- BY variables are required if specified with a BY statement.

SUMMARY= Data Set

You can read group summary statistics from a SUMMARY= data set specified in the PROC ANOM statement. This enables you to reuse OUTSUMMARY= data sets that have been created in previous runs of the ANOM procedure or to read output data sets created with SAS summarization procedures, such as PROC MEANS.

A SUMMARY= data set used with the XCHART statement must contain the following:

- the *group-variable*
- a group mean variable for each *response*
- a group sample size variable for each *response*
- a group standard deviation variable for each *response*

The names of the group mean, group range, and group sample size variables must be the *response* name concatenated with the suffix characters *X*, *S*, and *N*, respectively.

For example, consider the following statements:

```
proc anom summary=Summary;
   xchart (Weight Yieldstrength)*Batch;
run;
```

The data set Summary must include the variables Batch, WeightX, WeightS, WeightN, YieldstrengthX, YieldstrengthS, and YieldstrengthN. Note that if you specify a *response* name that contains 32 characters, the names of the summary variables must be formed from the first 16 characters and the last 15 characters of the *response* name, suffixed with the appropriate character.

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

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

- BY variables
- ID variables

By default, the ANOM procedure reads all of the observations in a SUMMARY= 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.

For an example of a SUMMARY= data set, see the section "Creating ANOM Charts for Means from Group Summary Data" on page 130.

TABLE= Data Set

You can read summary statistics and decision limits from a TABLE= data set specified in the PROC ANOM statement. This enables you to reuse an OUTTABLE= data set created in a previous run of the ANOM procedure. Because the ANOM procedure simply displays the information in a TABLE= data set, you can use TABLE= data sets to create specialized ANOM charts.

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

Table 4.26 Variables Required in a TABLE= Data Set

Variable	Description
group-variable	values of the *group-variable*
LDLX	lower decision limit for mean
LIMITN	nominal sample size associated with the decision limits
MEAN	central line
SUBN	group sample size
SUBX	group mean
UDLX	upper decision 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 whose length is no greater than 48.
- _VAR_. This variable is required if more than one *response* is specified or if the data set contains information for more than one *response*. This variable must be a character variable whose length is no greater than 32.

For an example of a TABLE= data set, see the section "Saving Decision Limits" on page 133.

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	*group-variable*
Vertical	DATA=	*response*
Vertical	SUMMARY=	group mean variable
Vertical	TABLE=	_SUBX_

Missing Values

An observation read from a DATA=, SUMMARY=, or TABLE= data set is not analyzed if the value of the group variable is missing. For a particular response variable, an observation read from a DATA= data set is not analyzed if the value of the response variable is missing. Missing values of response variables generally lead to unequal group sample sizes. For a particular response variable, an observation read from a SUMMARY= or TABLE= data set is not analyzed if the values of any of the corresponding summary variables are missing.

Examples: XCHART Statement

This section provides advanced examples of the XCHART statement.

Example 4.4: ANOM Charts with Unequal Group Sizes

NOTE: See *ANOM Charts with Unequal Group Sizes* in the SAS/QC Sample Library.

Consider the example described in "Creating ANOM Charts for Means from Response Values" on page 127. Suppose that four of the 10 measurements were missing for the third and fourth labeler positions. The following statements create a SAS data set named LabelDev2, which contains the resulting deviation measurements:

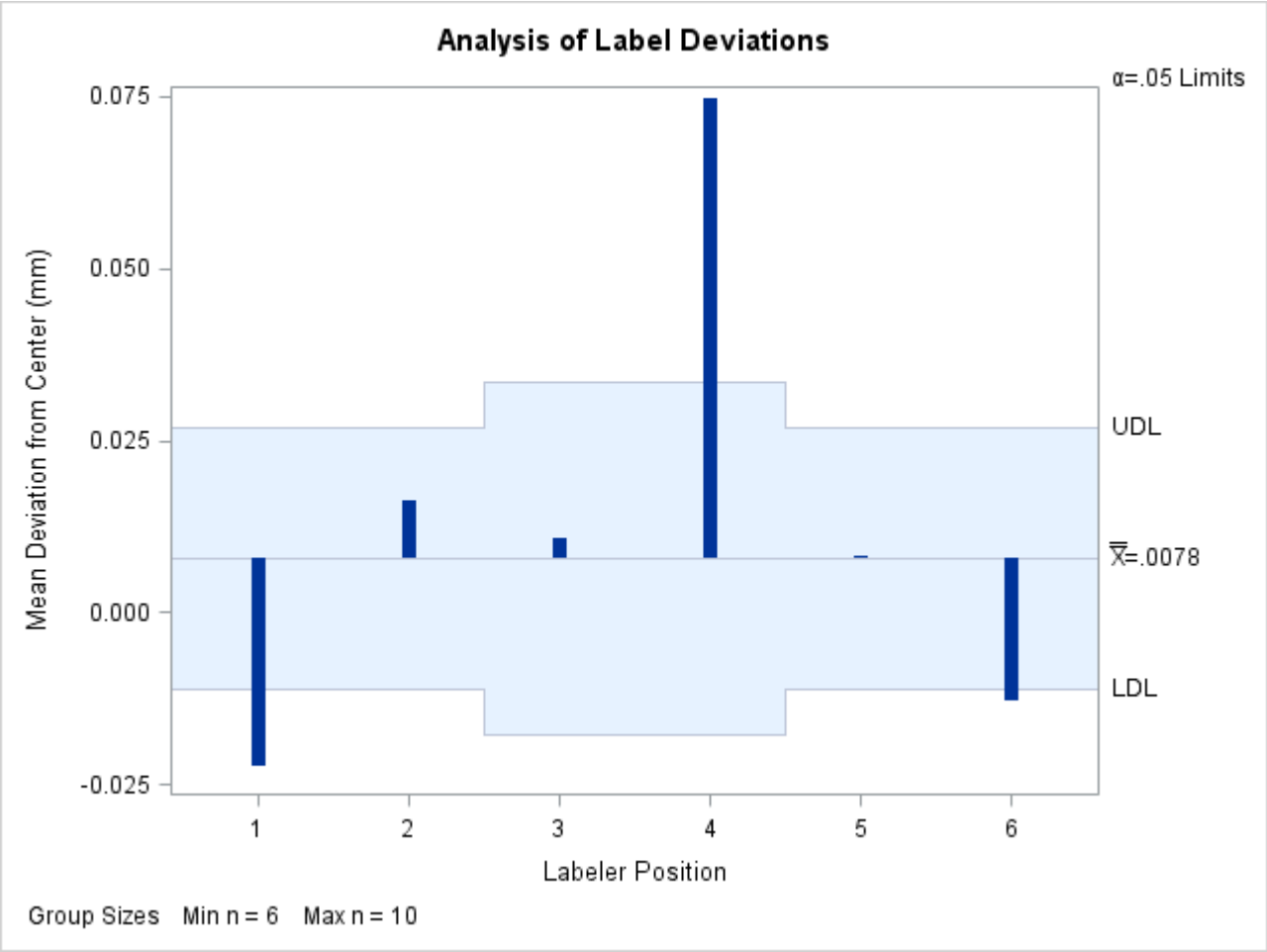
```
data LabelDev2;
   input Position @;
   do i = 1 to 5;
      input Deviation @;
      output;
   end;
   drop i;
   datalines;
1  -0.0239  -0.0285  -0.0300  -0.0043  -0.0362
1  -0.0422  -0.0014  -0.0647   0.0094  -0.0016
2  -0.0201  -0.0273   0.0227  -0.0332   0.0366
2   0.0438   0.0556   0.0098   0.0564   0.0182
3  -0.0073   0.0285    .        .      -0.0139
3    .       0.0357   0.0235    .      -0.0020
4   0.0669   0.1073    .        .       0.0755
4    .       0.0561   0.0899    .       0.0530
```

```
5   0.0368   0.0036   0.0374   0.0116  -0.0074
5   0.0250  -0.0080   0.0302  -0.0015  -0.0464
6   0.0049  -0.0384  -0.0204  -0.0049  -0.0120
6   0.0071  -0.0308   0.0017  -0.0285  -0.0070
;
```

The following statements create the ANOM chart shown in Output 4.4.1:

```
ods graphics on;
title 'Analysis of Label Deviations';
proc anom data=LabelDev2;
   xchart Deviation*Position / odstitle=title;
   label Deviation = 'Mean Deviation from Center (mm)';
   label Position  = 'Labeler Position';
run;
```

Output 4.4.1 ANOM Chart with Unequal Group Sizes



Note that the decision limits are automatically adjusted for the varying group sizes. The legend reports the minimum and maximum group sizes.

Example 4.5: ANOM for a Two-Way Classification

NOTE: See *ANOM for a Two-Way Classification* in the SAS/QC Sample Library.

A chemical engineer is interested in the effects of two factors, position and depth, on the concentration of a cleaning solution; refer to Ramig (1983) for details concerning the use of ANOM in a two-way classification such as this. The engineer is interested in the following questions:

1. Are there significant group or interaction effects due to position or depth?
2. Assuming a main effect is significant, which levels are significantly different from the overall mean and in which direction?

There are five positions and three depths. The engineer runs a two-way factorial experiment with two replications per cell. The following statements create a data set named Cleaning, which provides the concentration measurements for the $5 \times 3 \times 2 = 30$ observations.

```
data Cleaning;
   do position = 1 to 5;
      do depth = 1 to 3;
         do rep = 1 to 2;
            input concentration @@;
            output;
         end;
      end;
   end;
datalines;
15 16 15 14 19 5
15 16 14 14 0  8
19 15 16 16 11 8
18 16 19 15 8  14
15 12 19 15 8  11
;
```

In order to test for main effects and an interaction effect, the following statements use the GLM procedure:

```
ods graphics off;
proc glm data=Cleaning;
   class position depth;
   model concentration = position depth position*depth;
run;
```

The results are shown in Output 4.5.1:

Output 4.5.1 GLM Results

Analysis of Label Deviations

The GLM Procedure

Dependent Variable: concentration

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	14	390.4666667	27.8904762	2.21	0.0694
Error	15	189.0000000	12.6000000		
Corrected Total	29	579.4666667			

R-Square	Coeff Var	Root MSE	concentration Mean
0.673838	26.22893	3.549648	13.53333

Source	DF	Type I SS	Mean Square	F Value	Pr > F
position	4	50.4666667	12.6166667	1.00	0.4374
depth	2	281.6666667	140.8333333	11.18	0.0011
position*depth	8	58.3333333	7.2916667	0.58	0.7802

The results in Output 4.5.1 show no significant interaction effect[3] and a significant main effect due to depth. Since no interaction effect is present, you can use analysis of means to evaluate the effect of each factor as if two separate experiments had been run to determine the effect of each factor. In other words, the analysis of means is done twice, once for each factor. However, each analysis must be based on the mean square error ($\widehat{\sigma^2} = 12.6$) and the degrees of freedom for error ($\nu = 15$) from the two-way analysis of variance. These values must be specified since the ANOM procedure assumes a one-way layout by default for computing the decision limits.

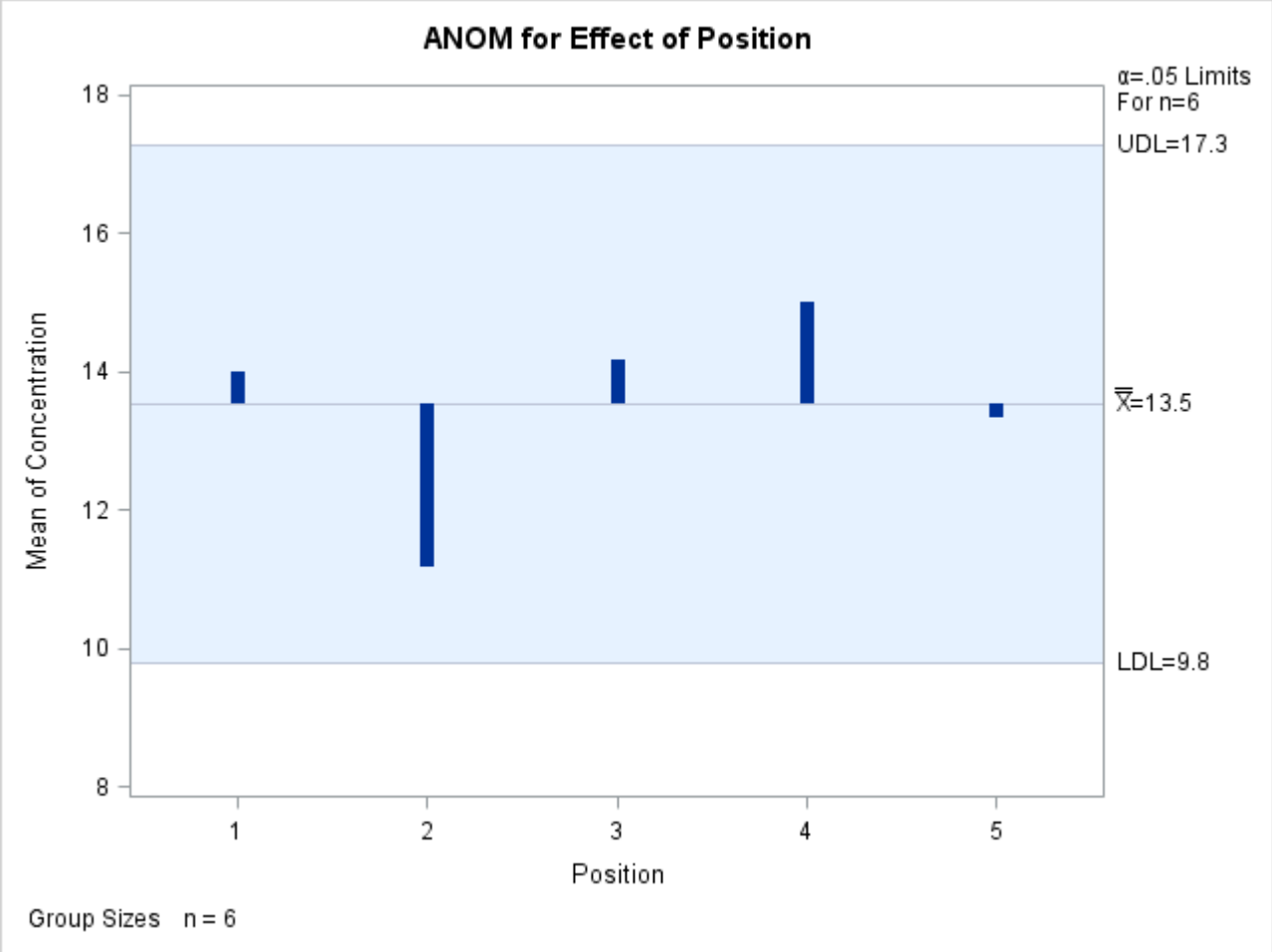
The following statements create the ANOM chart for the effect of position shown in Output 4.5.2:

```
ods graphics on;
title "ANOM for Effect of Position";
proc anom data=Cleaning;
   xchart concentration * position /
      mse      = 12.6
      dfe      = 15
      outtable = posmain
      odstitle = title;
   label position      = 'Position'
         concentration = 'Mean of Concentration';
run;
```

The MSE= and DFE= options are used to specify $\widehat{\sigma^2}$ and ν respectively. See the section "Constructing ANOM Charts for Two-Way Layouts" on page 145 for how the specified values are used to compute the decision limits. The OUTTABLE= option stores the output data set PosMain, which can be used to create a combined chart for the two factors.

[3] See Example 4.7 for an example that discusses the use of ANOM for the cell means when an interaction effect is present.

Output 4.5.2 ANOM for Effect of Position



Each point on the ANOM chart represents the average response for a particular level of position. In this case, all of the points are between the upper decision limit (UDL) and the lower decision limit (LDL). This is not surprising considering the fact that the main effect of Position was not significant in the ANOVA.

In order to create the ANOM chart for the effect of depth, the response must be sorted by depth.

```
proc sort data=Cleaning out=Cleaning2;
   by depth;
run;
```

Note that for the previous chart, the measurements were sorted by Position in the original data set.

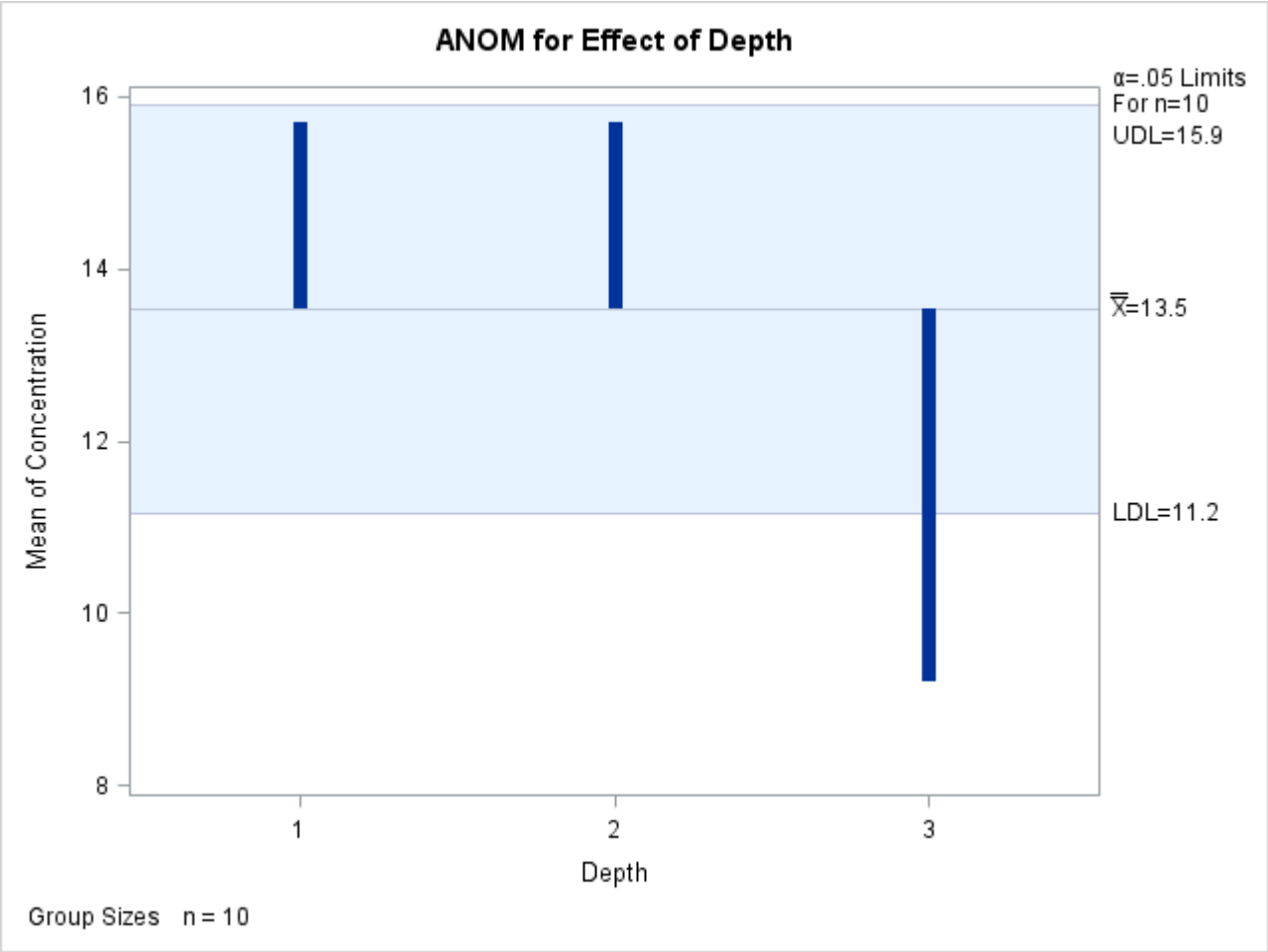
The following statements create the chart for depth:

```
title "ANOM for Effect of Depth";
proc anom data=Cleaning2;
   xchart concentration * depth /
      mse      = 12.6
      dfe      = 15
      outtable = depmain
      odstitle = title;
```

```
    label depth           = 'Depth'
          concentration   = 'Mean of Concentration';
run;
```

This produces the chart shown in Output 4.5.3: The OUTTABLE= option stores the output data set depmain, which can be used to create a combined chart for the two factors.

Output 4.5.3 ANOM for Effect of Depth



Since the average concentration for Depth 3 is less than the lower decision limit, you can conclude that the average response for Depth 3 is significantly less than the overall mean.

Example 4.6: ANOM Charts Using LIMITS= Data Set

NOTE: See *ANOM Charts Using LIMITS= Data Set* in the SAS/QC Sample Library.

In Example 4.5, statistics from a two-way ANOVA were passed to the ANOM procedure using options in order to compute the decision limits for the factor effects. This example shows how you can pass the statistics in a LIMITS= data set using the variables _MSE_ and _DFE_.

The GLM output in Output 4.5.1 provides the statistics. The following statements save the results from PROC GLM in the data sets MyFit, MyMeans, and MyOverAll:

```
ods select FitStatistics ModelANOVA OverAllANOVA;
ods output FitStatistics = MyFit
           ModelANOVA    = MyLimits
           OverAllANOVA  = MyOverAll;

proc glm data=Cleaning;
   class position depth;
   model concentration = position depth position*depth;
run;
```

The results of PROC GLM are identical to the results in Output 4.5.1.

The following statements create a LIMITS= data set to be used to create an ANOM chart for the effect of Position:

```
data ANOMParms;
   keep _var_ _group_ _alpha_ _mean_;
      length _var_ _group_ $ 14;
   set MyFit (rename=(Dependent=_var_ DepMean  =_mean_));
      _group_ = 'position';
      _alpha_ = 0.05;
run;

data ANOMParms;
   merge ANOMParms
         MyLimits (where=(source='position')
                   keep = source DF);
   _limitk_ = DF+1;
   drop source DF;
   merge MyOverAll (where=(source='Error')
            keep = source df ms
            rename=( df = _dfe_ ms = _mse_));
   drop source;
   merge MyOverAll (where=(source='Corrected Total')
                    keep = source DF);
   _limitn_ = (DF+1)/_limitk_;
   drop source DF;
run;
```

The data set ANOMParms contains a complete set of parameters, as shown in Output 4.6.1. Note these are the same values specified in the options for Example 4.5.

Output 4.6.1 Data Set ANOMParms

Parameters for ANOM for Effect of Position

var	_group_	_mean_	_alpha_	_limitk_	_dfe_	_mse_	_limitn_
concentration	position	13.53333	0.05	5	15	12.6000000	6

The following statements read the parameters in ANOMParms to create an ANOM chart for the effect of position:

```
ods graphics on;
title "ANOM for Effect of Position";
proc anom data=Cleaning limits=ANOMParms;
   xchart concentration * position /
      outtable = postable
      odstitle = title;
   label position      = 'Position'
         concentration = 'Mean of Concentration';
run;
```

The chart produced is identical to the one in Output 4.5.2. Note that the procedure creates a TABLE= input data set postable. You can use postable to create a combined chart for the two factors position and depth.

You can create a LIMITS= data set ANOMParmsB for the factor depth by using the above code and substituting 'depth' for the _group_ variable. You can use the OUTTABLE= statement to store the TABLE= input data set for depth as deptable. The resulting data set ANOMParmsB is shown in Output 4.6.2:

Output 4.6.2 Data Set ANOMParmsB

ANOM for Effect of Position

Obs	_var_	_group_	_mean_	_alpha_	_limitk_	_dfe_	_mse_	_limitn_
1	concentration	depth	13.53333	0.05	3	15	12.6000000	10

Example 4.7: ANOM for Cell Means in Presence of Interaction

NOTE: See *ANOM for Cell Means in the Presence of Interaction* in the SAS/QC Sample Library.

This example illustrates the use of analysis of means in an experiment with two factors where an interaction effect is present. The following data set CleaningInteract is a modified version of the data set Cleaning, which includes an interaction effect for position and depth.

Consider the following data set CleaningInteract:

```
data CleaningInteract;
   do position = 1 to 5;
      do depth = 1 to 3;
         do rep = 1 to 2;
            input concentration @@;
            output;
         end;
      end;
   end;
datalines;
15 16 15 14 19  5
15 16 14 14  0  1
19 15 16 16 11  8
18 16 24 23  8 14
15 12 23 24  8 11
;
```

The following statements use PROC GLM to test for an interaction:

```
ods graphics off;
proc glm data=CleaningInteract;
   class position depth;
   model concentration = position depth position*depth;
run;
```

The analysis of variance results in Output 4.7.1 indicate a significant interaction between position and depth.

Output 4.7.1 GLM Results

ANOM for Effect of Position

The GLM Procedure

Dependent Variable: concentration

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	14	885.666667	63.261905	6.66	0.0004
Error	15	142.500000	9.500000		
Corrected Total	29	1028.166667			

R-Square	Coeff Var	Root MSE	concentration Mean
0.861404	21.75676	3.082207	14.16667

Source	DF	Type I SS	Mean Square	F Value	Pr > F
position	4	169.0000000	42.2500000	4.45	0.0144
depth	2	515.4666667	257.7333333	27.13	<.0001
position*depth	8	201.2000000	25.1500000	2.65	0.0496

Since an interaction effect is present, an appropriate way to analyze the data is to create an ANOM chart for the cell means.

In order to create the chart you first need to compute the cell means and a new *group* variable which designates the cells. The following statements use PROC MEANS for this purpose.

```
proc means data=CleaningInteract n mean std;
   class position depth;
   var concentration;
   types position*depth;
   output out=cellmeans mean=concentrationX std=concentrationS;
run;
data cellmeans; set cellmeans;
   rename _FREQ_ = concentrationN;
   pos = put(position, z1.);
   dep = put(depth, z1.);
   cell  = cat('P',pos, 'D', dep);
   drop _TYPE_ pos dep;
run;
```

The cell means are stored in the data set cellmeans shown in Output 4.7.2:

Output 4.7.2 Data Set cellmeans

ANOM for Effect of Position

The MEANS Procedure

Analysis Variable : concentration					
position	depth	N Obs	N	Mean	Std Dev
1	1	2	2	15.5000000	0.7071068
	2	2	2	14.5000000	0.7071068
	3	2	2	12.0000000	9.8994949
2	1	2	2	15.5000000	0.7071068
	2	2	2	14.0000000	0
	3	2	2	0.5000000	0.7071068
3	1	2	2	17.0000000	2.8284271
	2	2	2	16.0000000	0
	3	2	2	9.5000000	2.1213203
4	1	2	2	17.0000000	1.4142136
	2	2	2	23.5000000	0.7071068
	3	2	2	11.0000000	4.2426407
5	1	2	2	13.5000000	2.1213203
	2	2	2	23.5000000	0.7071068
	3	2	2	9.5000000	2.1213203

ANOM for Effect of Position

position	depth	concentrationN	concentrationX	concentrationS	cell
1	1	2	15.5	0.70711	P1D1
1	2	2	14.5	0.70711	P1D2
1	3	2	12.0	9.89949	P1D3
2	1	2	15.5	0.70711	P2D1
2	2	2	14.0	0.00000	P2D2
2	3	2	0.5	0.70711	P2D3
3	1	2	17.0	2.82843	P3D1
3	2	2	16.0	0.00000	P3D2
3	3	2	9.5	2.12132	P3D3
4	1	2	17.0	1.41421	P4D1
4	2	2	23.5	0.70711	P4D2
4	3	2	11.0	4.24264	P4D3
5	1	2	13.5	2.12132	P5D1
5	2	2	23.5	0.70711	P5D2
5	3	2	9.5	2.12132	P5D3

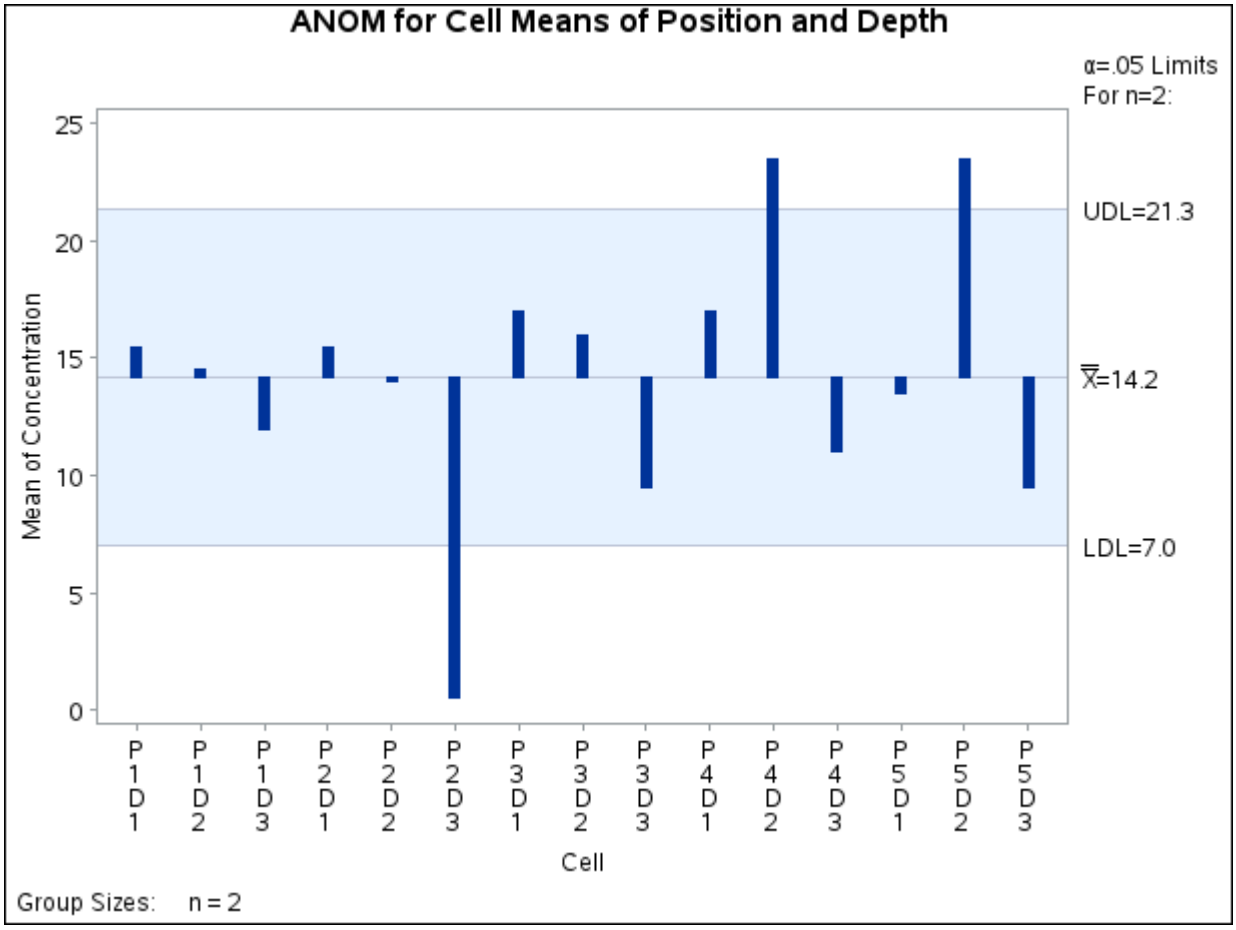
The data set cellmeans has the structure of a SUMMARY= input data set for the ANOM procedure. For details concerning a SUMMARY= data set, see the section "Creating ANOM Charts for Means from Group Summary Data" on page 130.

The following statements use cellmeans to create the ANOM chart for the cell means using SUMMARY= option:

```
ods graphics off;
title "ANOM for Cell Means of Position and Depth";
proc ANOM summary = cellmeans;
   xchart concentration * cell / turnhlabels;
   label concentrationX = 'Mean of Concentration';
   label cell           = 'Cell';
run;
```

The chart is shown in Output 4.7.3:

Output 4.7.3 ANOM for Cell Means of Position and Depth



The chart shows that the cell means for P2D3, P4D2, and P5D2 are significantly different from the average concentration level.

INSET Statement: ANOM Procedure

Overview: INSET Statement

The INSET statement enables you to enhance an ANOM chart by adding a box or table (referred to as an *inset*) of summary statistics directly to the graph. An inset can display statistics calculated by the ANOM procedure or arbitrary values provided in a SAS data set.

Note that an INSET statement by itself does not produce a display but must be used in conjunction with a chart statement.

You can use options in the INSET statement to

- specify the position of the inset
- specify a header for the inset table
- specify graphical enhancements, such as background colors, text colors, text height, text font, and drop shadows

Getting Started: INSET Statement

This section introduces the INSET statement with examples that illustrate commonly used options. Complete syntax for the INSET statement is presented in the section "Syntax: INSET Statement" on page 169.

Displaying Summary Statistics on an ANOM Chart

NOTE: See *Displaying Summary Statistics on an ANOM Chart* in the SAS/QC Sample Library.

A manufacturing engineer carries out a study to determine the source of excessive variation in the positioning of labels on shampoo bottles. [4] A labeling machine removes bottles from the line, attaches the labels, and returns the bottles to the line. There are six positions on the machine, and the engineer suspects that one or more of the position heads might be faulty.

A sample of 60 bottles, 10 per position, is run through the machine. For each bottle, the deviation of each label is measured in millimeters, and the machine position is recorded. The following statements create a SAS data set named LabelDeviations, which contains the deviation measurements for the 60 bottles:

[4] This example is based on a case study described by Hansen (1990).

```
data LabelDeviations;
   input Position @;
   do i = 1 to 5;
      input Deviation @;
      output;
   end;
   drop i;
   datalines;
1  -0.02386  -0.02853  -0.03001  -0.00428  -0.03623
1  -0.04222  -0.00144  -0.06466   0.00944  -0.00163
2  -0.02014  -0.02725   0.02268  -0.03323   0.03661
2   0.04378   0.05562   0.00977   0.05641   0.01816
3  -0.00728   0.02849  -0.04404  -0.02214  -0.01394
3   0.04855   0.03566   0.02345   0.01339  -0.00203
4   0.06694   0.10729   0.05974   0.06089   0.07551
4   0.03620   0.05614   0.08985   0.04175   0.05298
5   0.03677   0.00361   0.03736   0.01164  -0.00741
5   0.02495  -0.00803   0.03021  -0.00149  -0.04640
6   0.00493  -0.03839  -0.02037  -0.00487  -0.01202
6   0.00710  -0.03075   0.00167  -0.02845  -0.00697
;
```

The following statements generate an ANOM chart from the LabelDeviations data. An INSET statement is used to display the mean square error (MSE) and the number of groups outside of the decision limits (NOUT) on the chart:

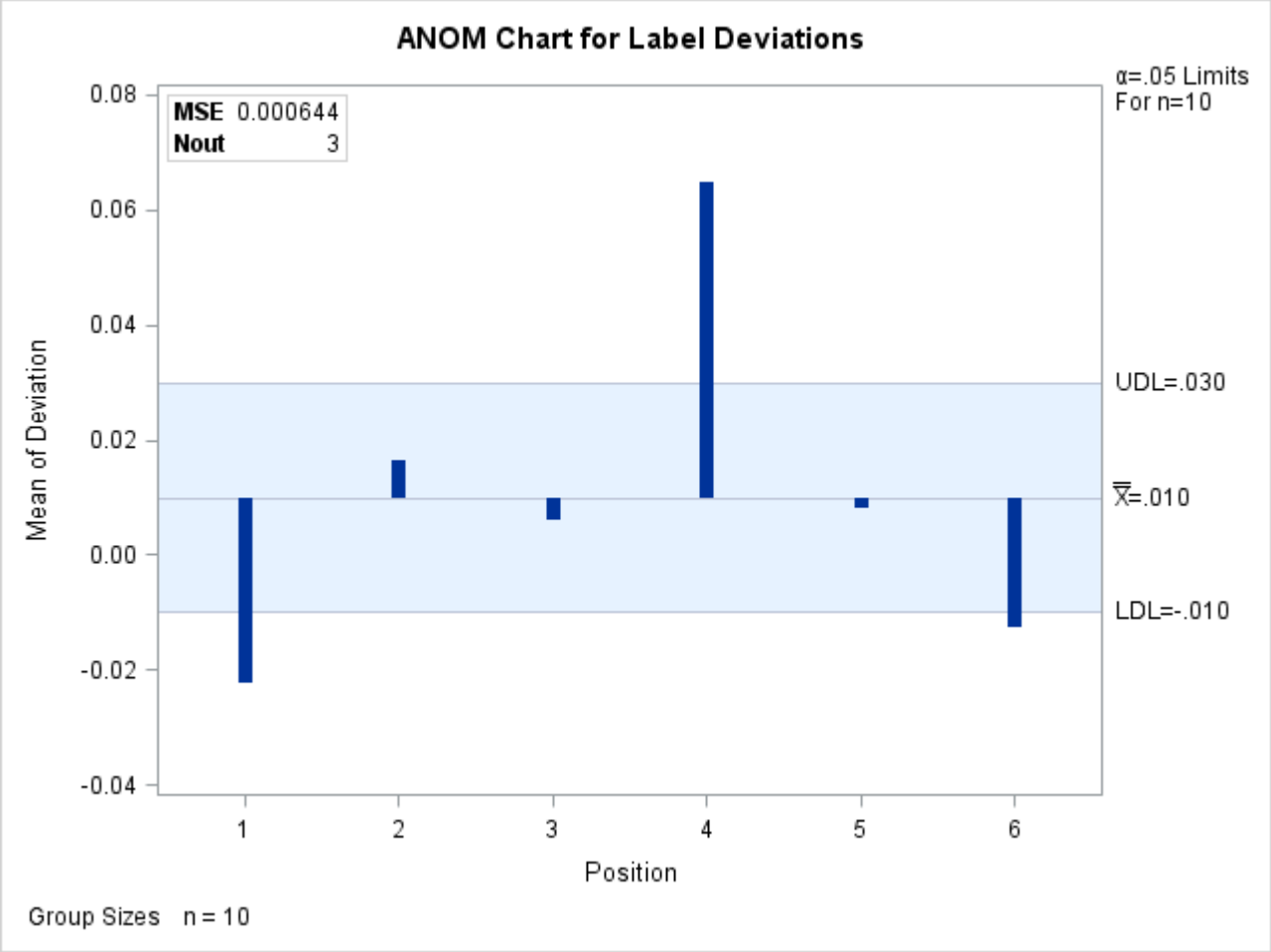
```
ods graphics on;
title 'ANOM Chart for Label Deviations';
proc anom data=LabelDeviations;
   xchart Deviation*Position / odstitle = title;
   inset mse nout / height = 3;
run;
```

The ODS GRAPHICS ON statement specified before the PROC ANOM statement enables ODS Graphics, so the chart is created using ODS Graphics instead of traditional graphics. The resulting ANOM chart is displayed in Figure 4.28.

The INSET statement immediately follows the chart statement that creates the graphical display (in this case, the XCHART statement). Specify the keywords for inset statistics (such as ALPHA) immediately after the word INSET. The inset statistics appear in the order in which you specify the keywords. The HEIGHT= option on the INSET statement specifies the text height used to display the statistics in the inset.

A complete list of keywords that you can use with the INSET statement is provided in "Summary of INSET Keywords" on page 170. Note that the set of keywords available for a particular display may depend on both the chart statement that precedes the INSET statement and the options that you specify in the chart statement.

Figure 4.28 An ANOM Chart with an Inset



The following examples illustrate options commonly used for enhancing the appearance of an inset.

Formatting Values and Customizing Labels

NOTE: See *Formatting and Positioning the Inset on an ANOM Chart* in the SAS/QC Sample Library.

By default, each inset statistic is identified with an appropriate label, and each numeric value is printed using an appropriate format. However, you may want to provide your own labels and formats. For example, in Figure 4.28 the default format used for the MSE prints an excessive number of decimal places. In the inset produced by the following statements, the unwanted decimal places are eliminated and the default MSE label is replaced by one specified by the user:

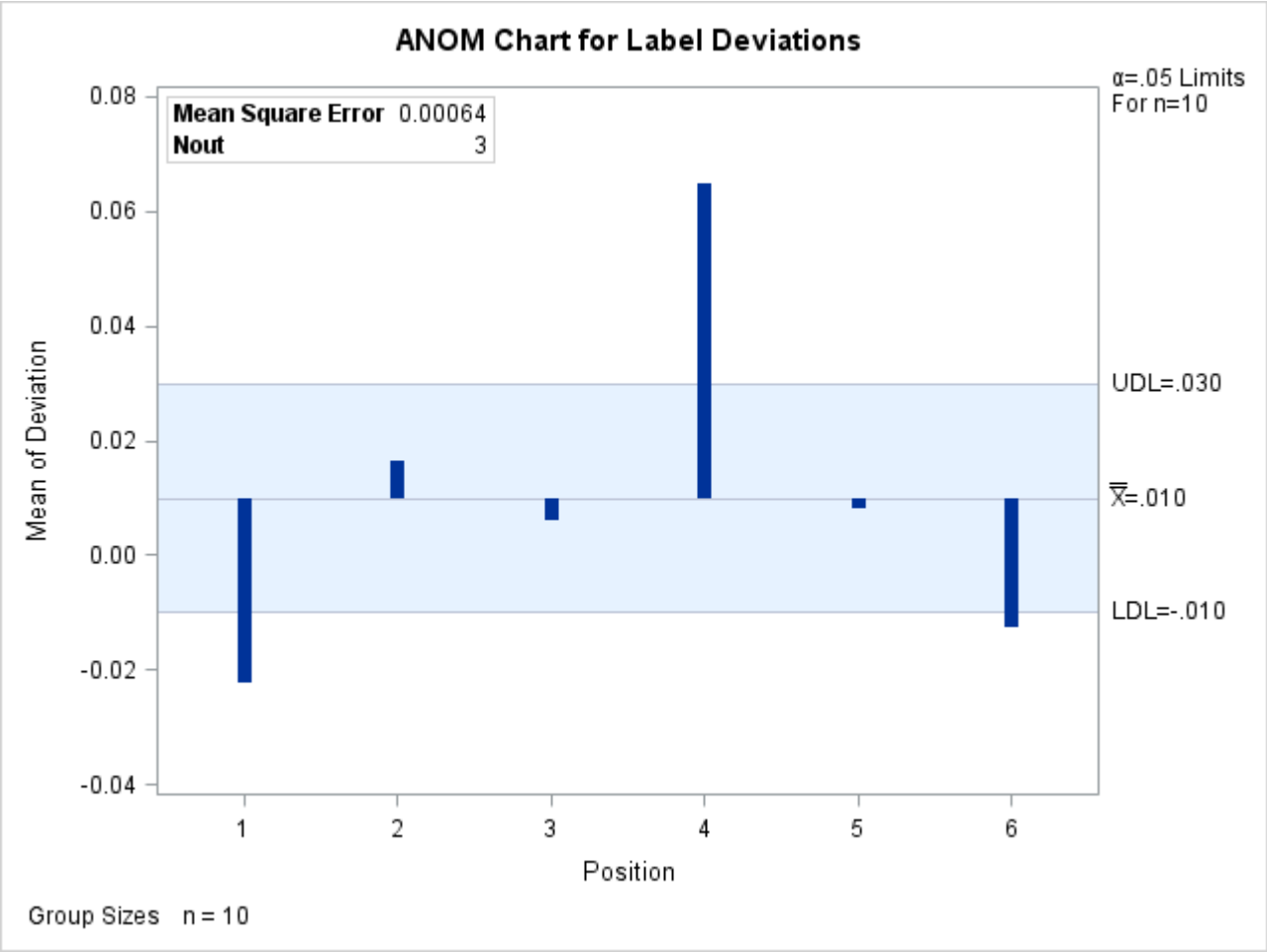
```
title 'ANOM Chart for Label Deviations';
proc anom data=LabelDeviations;
   xchart Deviation*Position / odstitle=title;
   inset mse='Mean Square Error' (7.5) nout;
run;
```

The resulting ANOM chart is displayed in Figure 4.29. You can provide your own label by specifying the keyword for that statistic followed by an equal sign (=) and the label in quotes. The label can have up to 24 characters.

The format 7.5 specified in parentheses after the MSE keyword displays the statistic with a field width of seven and five decimal places. In general, you can specify any numeric SAS format in parentheses after an inset keyword. You can also specify a format to be used for all the statistics in the INSET statement with the FORMAT= option (see Figure 4.30). For more information about SAS formats, refer to *SAS Formats and Informats: Reference*.

Note that if you specify both a label and a format for a statistic, the label must appear before the format.

Figure 4.29 Formatting Values and Customizing Labels in an Inset



Adding a Header and Positioning the Inset

NOTE: See *Adding a Header and Positioning the Inset on an ANOM Chart* in the SAS/QC Sample Library.

In the previous examples, the insets are displayed in the upper left corners of the plots, the default position for insets added to ANOM charts. You can control the inset position with the POSITION= option. In addition, you can display a header at the top of the inset with the HEADER= option. The following statements create a data set to be used with the INSET DATA= keyword and the chart shown in Figure 4.30:

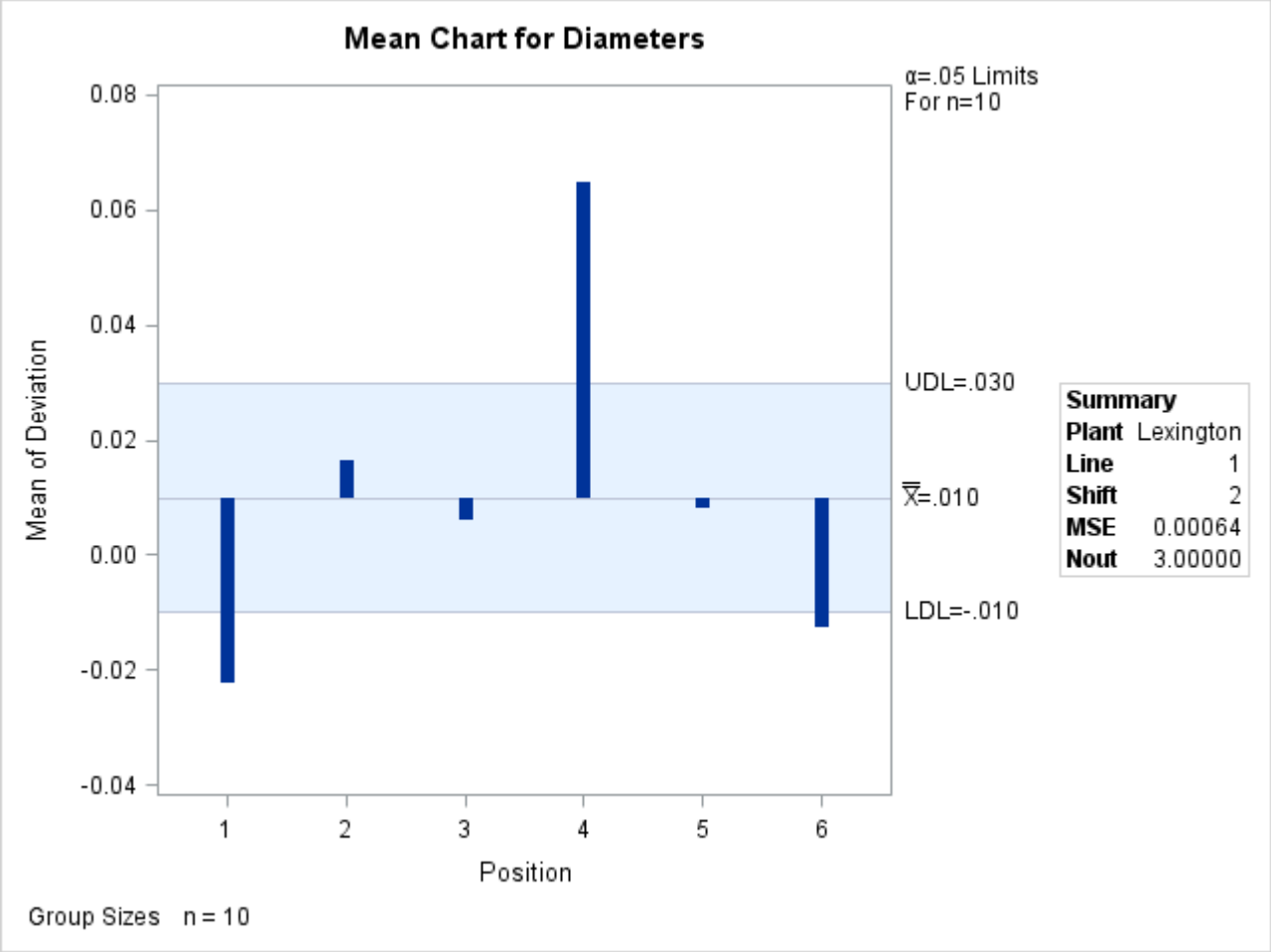
```
data Location;
   length _LABEL_ $ 10 _VALUE_ $ 12;
   input _LABEL_ _VALUE_ &;
   datalines;
Plant     Lexington
Line      1
Shift     2
;

title 'Mean Chart for Diameters';
proc anom data=LabelDeviations;
   xchart Deviation*Position / odstitle=title;
   inset data=Location mse nout /
      format   = 8.5
      position = rm
      cshadow  = black
      height   = 3
      header   = 'Summary';
run;
```

The header (in this case, *Summary*) can be up to 40 characters. POSITION=RM is specified to position the inset in the right margin. For more information about positioning, see "Details: INSET Statement" on page 175. The CSHADOW= option is used to display a drop shadow on this inset. The *options*, such as HEADER=, POSITION=, and CSHADOW=, are specified after the slash (/) in the INSET statement. For more details on INSET statement options, see "Dictionary of Options" on page 173.

Note that the contents of the data set Location appear before other statistics in the inset. The position of the DATA= keyword in the keyword list determines the position of the data set's contents in the inset. The FORMAT= option applies format 8.5 to the statistics listed in the INSET statement. Note that the format does not apply to the values from the Location data set. You can associate a format with the _VALUE_ variable in the data set to format those values.

Figure 4.30 Adding a Header and Repositioning the Inset



Syntax: INSET Statement

The syntax for the INSET statement is as follows:

INSET *keyword-list* < / *options* > **;**

You can use any number of INSET statements in the ANOM procedure. However, when ODS Graphics is enabled, at most two insets are displayed inside the plot area and at most two are displayed in the chart margins. Each INSET statement produces a separate inset and must follow one of the chart statements. The inset appears on every panel (page) produced by the last chart statement preceding it. The statistics are displayed in the order in which they are specified. The following statements produce an ANOM boxchart with two insets and an ANOM chart for means with one inset.

```
proc anom data=LabelDeviations;
   boxchart Deviation*Position;
      inset alpha mse dfe;
      inset ldl mean udl;
   xchart Deviation*Position;
      inset ngroups nmin nmax;
run;
```

The statistics displayed in an inset are computed for a specific response variable using observations for the current BY group. For example, in the following statements, there are two response variables (Weight and Diameter) and a BY variable (Location). If there are three different locations (levels of Location), then a total of six ANOM charts are produced. The statistics in each inset are computed for a particular variable and location. The labels in the inset are the same for each ANOM chart.

```
proc anom data=Axles;
   by Location;
   xchart (Weight Diameter)*Batch;
   inset alpha mse dfe;
run;
```

The components of the INSET statement are described as follows.

keyword-list

can include any of the *keywords* listed in "Summary of INSET Keywords" on page 170. By default, inset statistics are identified with appropriate labels, and numeric values are printed using appropriate formats. However, you can provide customized labels and formats. You provide the customized label by specifying the *keyword* for that statistic followed by an equal sign (=) and the label in quotes. Labels can have up to 24 characters. You provide the numeric format in parentheses after the *keyword*. Note that if you specify both a label and a format for a statistic, the label must appear before the format. For an example, see "Formatting Values and Customizing Labels" on page 166.

options

appear after the slash (/) and control the appearance of the inset. For example, the following INSET statement uses two appearance *options* (POSITION= and CTEXT=):

```
inset n nmin nmax / position=ne ctext=yellow;
```

The POSITION= option determines the location of the inset, and the CTEXT= option specifies the color of the text of the inset.

See "Summary of Options" on page 171 for a list of all available *options*, and "Dictionary of Options" on page 173 for detailed descriptions. Note the difference between *keywords* and *options*; *keywords* specify the information to be displayed in an inset, whereas *options* control the appearance of the inset.

Summary of INSET Keywords

All keywords available with the ANOM procedure's INSET statement request a single statistic in an inset, except for the DATA= keyword. The DATA= keyword specifies a SAS data set containing (label, value) pairs to be displayed in an inset. The data set must contain the variables _LABEL_ and _VALUE_. _LABEL_ is a character variable whose values provide labels for inset entries. _VALUE_ can be character or numeric, and provides values displayed in the inset. The label and value from each observation in the DATA= data set occupy one line in the inset. Figure 4.30 shows an inset containing entries from a DATA= data set.

Table 4.27 Summary Statistics

Keyword	Description
ALPHA	significance level
DATA=	(label, value) pairs from *SAS-data-set*
DFE	degrees of freedom
LDL	lower decision limit
MEAN	weighted average of group means
MSE	mean square error
N	nominal group size
NGROUPS	number of groups
NHIGH	number of groups above upper decision limit
NLOW	number of groups below lower decision limit
NMAX	maximum group size
NMIN	minimum group size
NOBS	total number of observations
NOUT	total number of groups outside decision limits
RMSE	root mean square error
UDL	upper decision limit

You can use the keywords in Table 4.28 only when producing ODS Graphics output. The labels for the statistics use Greek letters.

Table 4.28 Keywords Specific to ODS Graphics Output

Keyword	Description
UALPHA	probability of Type 1 error
UMU	weighted average of group means

Summary of Options

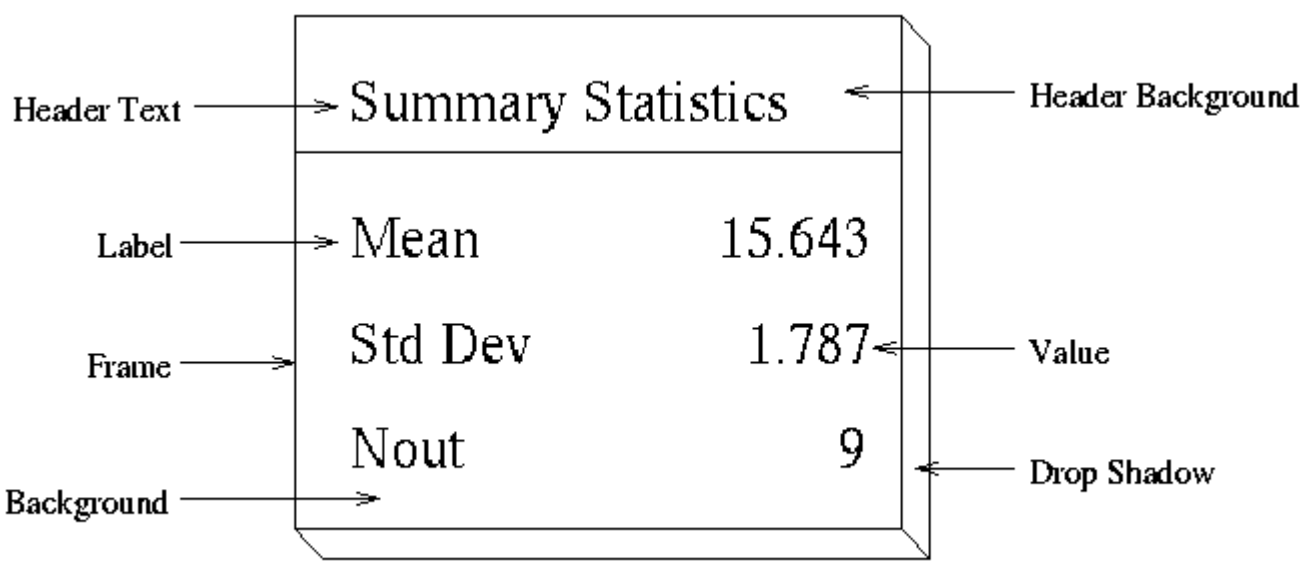
The following table lists the INSET statement options. For complete descriptions, see "Dictionary of Options" on page 173.

Table 4.29 INSET Options

Option	Description
CFILL=	specifies color of inset background
CFILLH=	specifies color of header background
CFRAME=	specifies color of frame
CHEADER=	specifies color of header text
CSHADOW=	specifies color of drop shadow
CTEXT=	specifies color of inset text
DATA	specifies data units for POSITION=(x, y) coordinates
FONT=	specifies font of text
FORMAT=	specifies format of values in inset
HEADER=	specifies header text
HEIGHT=	specifies height of inset text
NOFRAME	suppresses frame around inset
POSITION=	specifies position of inset
REFPOINT=	specifies reference point of inset positioned with POSITION=(x, y) coordinates

The following sections provide detailed descriptions of options for the INSET statement. Terms used in this section are illustrated in Figure 4.31.

Figure 4.31 The Inset



Dictionary of Options

General Options

You can specify the following options whether you use ODS Graphics or traditional graphics:

DATA

specifies that data coordinates are to be used in positioning the inset with the POSITION= option. The DATA option is available only when you specify POSITION= (x, y), and it must be placed immediately after the coordinates (x, y). For details, see the entry for the POSITION= option or "Positioning the Inset Using Coordinates" on page 177. See Figure 4.33 for an example.

FONT=*font*

specifies the font used for text in the inset. By default, the font associated with the GraphLabelText style element is used for the inset header and that associated with the GraphValueText style element is used for text in the body of the inset.

FORMAT=*format*

specifies a format for all the values displayed in an inset. If you specify a format for a particular statistic, then this format overrides the format you specified with the FORMAT= option.

HEADER='*string*'

specifies the header text. The *string* cannot exceed 40 characters. If you do not specify the HEADER= option, no header line appears in the inset.

HEIGHT=*height*

HEIGHT=SMALL

specifies the height of the text in the inset. By default, the GraphLabelText style element determines the size of inset header text and the GraphValueText style element determines the size of text in the body of the inset.

When you produce traditional graphics, you can specify the *height* in screen percent units to be used for text in both the header and the body of the inset.

When you produce ODS Graphics output, you can specify HEIGHT=SMALL to reduce the height of text in the inset. The GraphValueText size is used for the inset header and the GraphDataText size is used in the inset body.

NOFRAME

suppresses the frame drawn around the text.

POSITION=*position*

POS=*position*

determines the position of the inset. The *position* can be a compass point keyword, a margin keyword, or a pair of coordinates (x, y). You can specify coordinates in axis percent units or axis data units. For more information, see "Details: INSET Statement" on page 175. By default, POSITION=NW, which positions the inset in the upper left (northwest) corner of the display.

NOTE: You cannot specify coordinates with the POSITION= option when producing ODS Graphics output.

REFPOINT=BR | BL | TR | TL

RP=BR | BL | TR | TL

specifies the reference point for an inset that is positioned by a pair of coordinates with the POSITION= option. Use the REFPOINT= option with POSITION= coordinates. The REFPOINT= option specifies which corner of the inset frame you want positioned at coordinates (x, y). The keywords BL, BR, TL, and TR represent bottom left, bottom right, top left, and top right, respectively. See Figure 4.34 for an example. The default is REFPOINT=BL.

If you specify the position of the inset as a compass point or margin keyword, the REFPOINT= option is ignored. For more information, see "Positioning the Inset Using Coordinates" on page 177.

Options for Traditional Graphics

You can specify the following options only when traditional graphics are produced. The ANOM procedure produces traditional graphics when ODS Graphics is disabled and SAS/GRAPH is licensed.

CFILL=*color* | BLANK

specifies the color of the background (including the header background if you do not specify the CFILLH= option).

If you do not specify the CFILL= option, then by default, the background is empty. This means that items that overlap the inset (such as needles representing group data or decision limits) show through the inset. If you specify any value for the CFILL= option, then overlapping items no longer show through the inset. Specify CFILL=BLANK to leave the background uncolored and also to prevent items from showing through the inset.

CFILLH=*color*

specifies the color of the header background. By default, if you do not specify a CFILLH= color, the CFILL= color is used.

CFRAME=*color*

specifies the color of the frame. By default, the frame is the same color as the axis of the plot.

CHEADER=*color*

specifies the color of the header text. By default, if you do not specify a CHEADER= color, the CTEXT= color is used.

CSHADOW=*color*

CS=*color*

specifies the color of the drop shadow. See Figure 4.30 for an example. By default, if you do not specify the CSHADOW= option, a drop shadow is not displayed.

CTEXT=*color*

CT=*color*

specifies the color of the text. By default, the inset text color is the same as the other text on the plot.

Details: INSET Statement

This section provides details on three different methods of positioning the inset using the POSITION= option. With the POSITION= option, you can specify

- compass points
- keywords for margin positions
- coordinates in data units or percent axis units

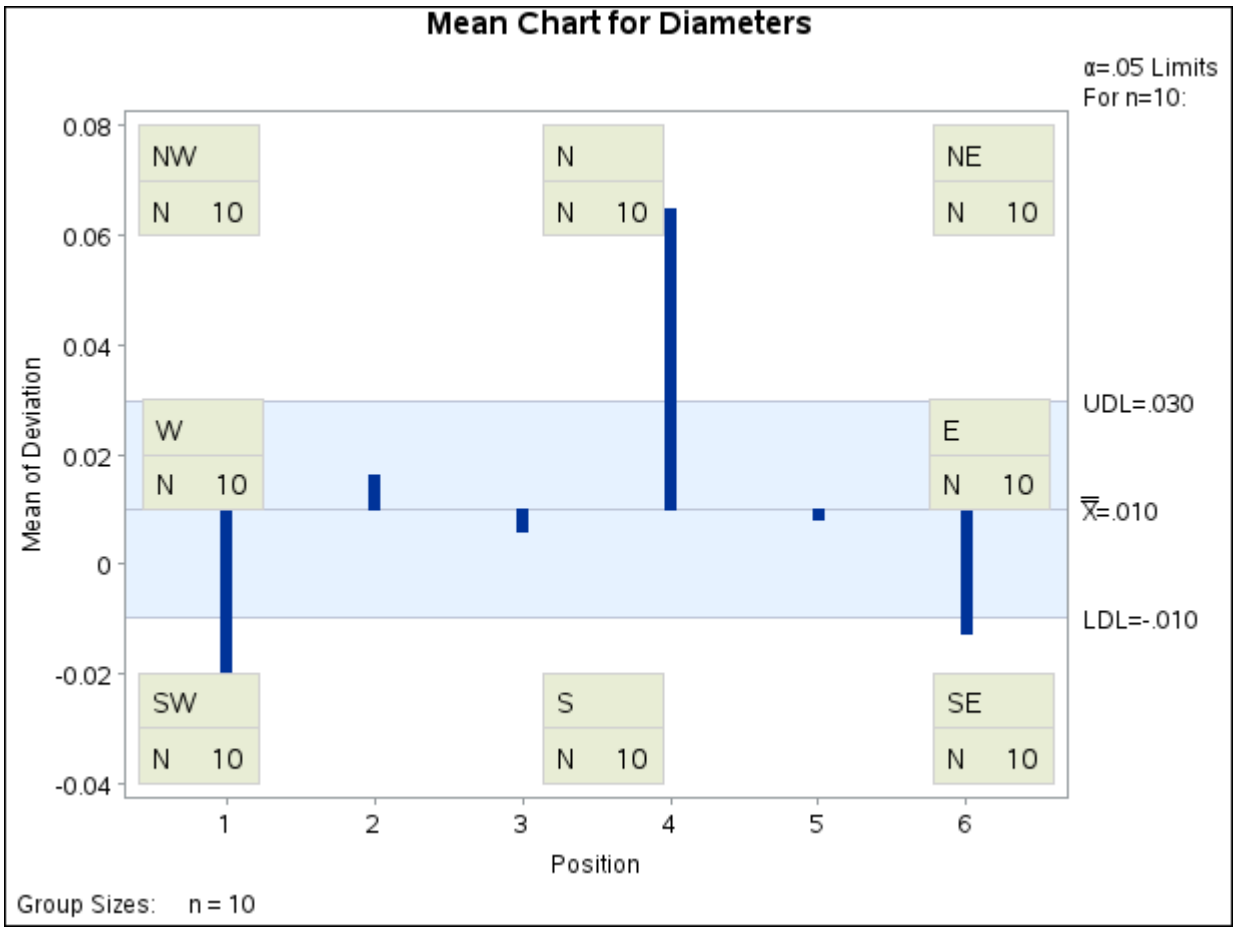
Positioning the Inset Using Compass Points

NOTE: See *Positioning the Inset on an ANOM Chart Using Compass Points* in the SAS/QC Sample Library.

You can specify the eight compass points N, NE, E, SE, S, SW, W, and NW as keywords for the POSITION= option. The following statements create the display in Figure 4.32, which demonstrates all eight compass positions. The default is NW.

```
ods graphics off;
title 'Mean Chart for Diameters';
proc anom data=LabelDeviations;
   xchart Deviation*Position;
   inset n / height=3 cfill=ywh header='NW' pos=nw;
   inset n / height=3 cfill=ywh header='N ' pos=n ;
   inset n / height=3 cfill=ywh header='NE' pos=ne;
   inset n / height=3 cfill=ywh header='E ' pos=e ;
   inset n / height=3 cfill=ywh header='SE' pos=se;
   inset n / height=3 cfill=ywh header='S ' pos=s ;
   inset n / height=3 cfill=ywh header='SW' pos=sw;
   inset n / height=3 cfill=ywh header='W ' pos=w ;
run;
```

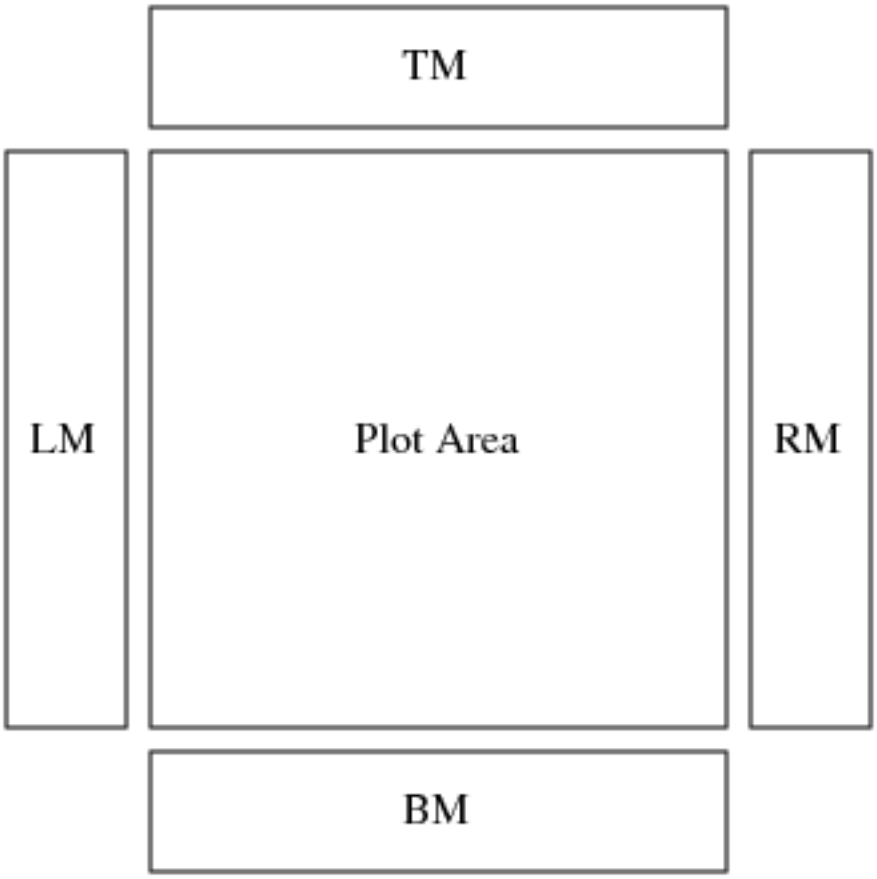
Figure 4.32 Insets Positioned Using Compass Points



Positioning the Inset in the Margins

Using the INSET statement you can also position an inset in one of the four margins surrounding the plot area using the margin keywords LM, RM, TM, or BM, as illustrated in Figure 4.7.4.

Output 4.7.4 Positioning Insets in the Margins



For an example of an inset placed in the right margin, see Figure 4.30. Margin positions are recommended if a large number of statistics are listed in the INSET statement. If you attempt to display a lengthy inset in the interior of the plot, it is likely that the inset will collide with the data display.

Positioning the Inset Using Coordinates

NOTE: See *Positioning the Inset Using Coordinates on an ANOM Chart* in the SAS/QC Sample Library.

When you produce traditional graphics, you can also specify the position of the inset with coordinates: POSITION= (x, y). The coordinates can be given in axis percent units (the default) or in axis data units.

Data Unit Coordinates

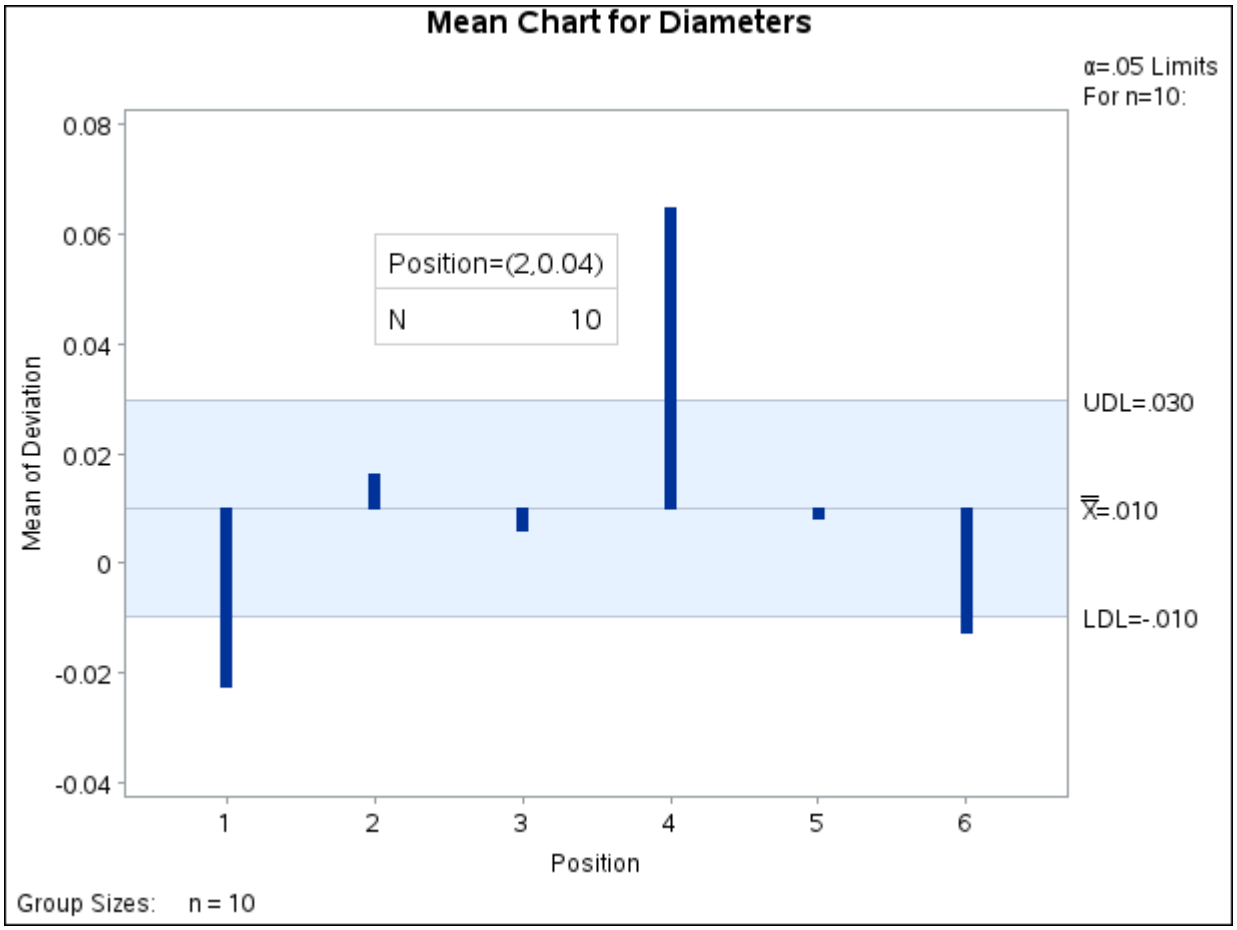
If you specify the DATA option immediately following the coordinates, the inset is positioned using axis data units. For example, the following statements place the bottom left corner of the inset at 2 on the horizontal axis and 0.04 on the vertical axis:

```
ods graphics off;
title 'Mean Chart for Diameters';
proc anom data=LabelDeviations;
   xchart Deviation*Position;
   inset n /
      header   = 'Position=(2,0.04)'
```

```
        height   = 3
        position = (2,0.04) data;
run;
```

The ANOM chart is displayed in Figure 4.33. By default, the specified coordinates determine the position of the bottom left corner of the inset. You can change this reference point with the REFPOINT= option, as in the next example.

Figure 4.33 Inset Positioned Using Data Unit Coordinates



Axis Percent Unit Coordinates

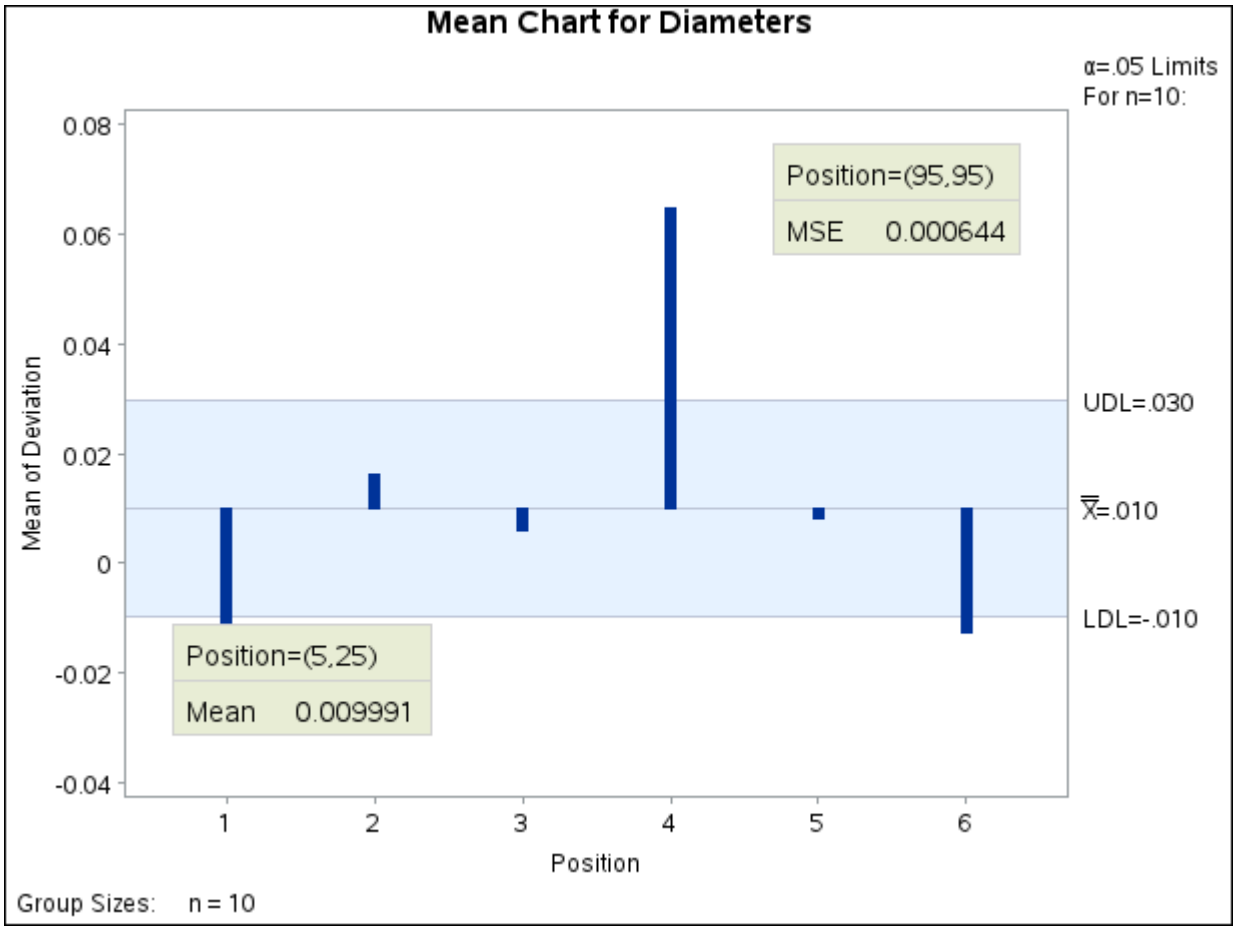
If you do not use the DATA option, the inset is positioned using axis percent units. The coordinates of the bottom left corner of the display are $(0, 0)$, while the upper right corner is $(100, 100)$. For example, the following statements create an ANOM chart with two insets, both positioned using coordinates in axis percent units:

```
title 'Mean Chart for Diameters';
proc anom data=LabelDeviations;
   xchart Deviation*Position;
   inset mean / position = (5,25)
                header   = 'Position=(5,25)'
                height   = 3
                cfill   = ywh
```

```
                 refpoint = tl;
    inset mse / position = (95,95)
                header   = 'Position=(95,95)'
                height   = 3
                cfill    = ywh
                refpoint = tr;
  run;
```

The display is shown in Figure 4.34. Notice that the REFPOINT= option is used to determine which corner of the inset is to be placed at the coordinates specified with the POSITION= option. The first inset has REFPOINT=TL, so the top left corner of the inset is positioned 5% of the way across the horizontal axis and 25% of the way up the vertical axis. The second inset has REFPOINT=TR, so the top right corner of the inset is positioned 95% of the way across the horizontal axis and 95% of the way up the vertical axis. Note also that coordinates in axis percent units must be *between* 0 and 100.

Figure 4.34 Inset Positioned Using Axis Percent Unit Coordinates



Dictionary of ANOM Chart Statement Options

This section provides detailed descriptions of options that you can specify in the following chart statements:

- BOXCHART
- PCHART
- UCHART
- XCHART

Options that are common to the ANOM and SHEWHART procedures are listed in the "Summary of Options" subsection in the sections for each chart statement. They are described in detail in "Dictionary of Options: SHEWHART Procedure" on page 1946

Options are specified after the slash (/) in a chart statement. For example, to place the label "Mean" on the center line of an ANOM chart, you can use the XSYMBOL= option as follows:

```
proc anom data=Measures;
   xchart Length*Sample / xsymbol='Mean';
run;
```

The options described in this section are listed alphabetically. For tables of options organized by function, see the "Summary of Options" sections in the sections for the various chart statements. Unless indicated otherwise, the options listed here are available with every chart statement.

ALPHA=*value*
specifies the probability of a Type I error.

CINFILL=*color* | **EMPTY** | **NONE**
specifies the color for the area inside the decision limits. By default, this area filled with an appropriate color from the ODS style. You can specify the keyword EMPTY or NONE to leave the area between the decision limits unfilled. If you specify a color, it is ignored when ODS Graphics is enabled.

CLIMITS=*color*
specifies the color for the decision limits, the central line, and related labels in traditional graphics. This option is ignored when ODS Graphics is enabled.

DFE=*n*
specifies the degrees of freedom *n* associated with the root mean square error.

GROUPN=*value*
GROUPN=*variable*
specifies the group sizes as a constant *value* or as the values of a variable in the DATA= data set. The GROUPN= option is available only in the PCHART and UCHART statements. You must specify GROUPN= in a PCHART or UCHART statement when your input data set is a DATA= data set.

If you specify multiple *responses* in a chart statement, the GROUPN= option is used with all of the *responses* listed.

LDLLABEL='*label*'
specifies a label for the lower decision limit in the ANOM chart. The label can be of length 16 or less. Enclose the label in quotes. The default label is of the form LDL=value if the decision limit has a fixed value; otherwise, the default label is LDL. A related option is UDLLABEL=.

LIMITK=*k*
specifies the number of groups for computing decision limits.

LIMITN=*n*
specifies either a nominal sample size for fixed decision limits or varying limits.

LIMLABSUBCHAR='*c*'
specifies a substitution character *c* for labels provided as quoted strings with the LDLLABEL=, UDLLABEL=, PSYMBOL=, USYMBOL=, and XSYMBOL= options. The substitution character must appear in the label. When the label is displayed on the chart, the character is replaced with the value of the corresponding decision limit or center line, provided that this value is constant across groups. Otherwise, the default label for a varying decision limit or center line is displayed.

LLIMITS=*linetype*
specifies the line type for decision limits in traditional graphics. This option is ignored when ODS Graphics is enabled.

MEAN=*value*
specifies the (known) mean of the response. This value is used for each response specified in the chart statement.

MSE=*value*
specifies the mean square error.

NDECIMAL=
specifies the number of digits to the right of decimal place in default labels for decision limits and central line

NOCTL
suppresses display of the central line.

NOLDL
suppresses display of the lower decision limit.

NOLIMITLABEL
suppresses labels for the decision limits and central line.

NOLIMIT0
suppresses display of the lower decision limit if it is 0.

NOLIMIT1
suppresses display of the upper decision limit if it is 1 (100%).

NOLIMITS
suppresses display of the decision limits.

NOLIMITSFRAME
suppresses the default frame around decision limit information when multiple sets of decision limits are read from a LIMITS= data set.

NOLIMITSLEGEND
suppresses the decision limits legend.

NONEEDLES
suppresses the needles connecting points to the center line.

NOREADLIMITS
specifies that the decision limits for each response listed in the chart statement are *not* to be read from the LIMITS= data set specified in the PROC ANOM statement. There are two basic methods of displaying decision limits: calculating decision limits from the data and reading decision limits from a LIMITS= data set. If you specify a LIMITS= data set but want the decision limits to be calculated from the data, specify the NOREADLIMITS option.

NOUDL
suppresses display of the upper decision limit.

OUTSUMMARY=*SAS-data-set*

OUT=*SAS-data-set*

OUTHISTORY=*SAS-data-set*
creates an output data set that contains group summary statistics. You can use an OUTSUMMARY= data set as a SUMMARY= input data set in a subsequent run of the procedure. You cannot request an OUTSUMMARY= data set if the input data set is a TABLE= data set. See "Output Data Sets" in the section for the chart statement in which you are interested.

P=
specifies the weighted average of group proportions.

PSYMBOL='*label***'**
specifies the label for the central line on an ANOM *p* chart.

READINDEXES=
reads multiple sets of decision limits for each *response* from a LIMITS= data set.

TYPE=
identifies parameters as estimates or standard values and specifies value of _TYPE_ in the OUTLIMITS= data set.

U=
specifies the weighted average of group rates.

UDLLABEL=
specifies the label for the upper decision limit.

USYMBOL='*label***'**
specifies the label for the central line on an ANOM *u* chart.

WLIMITS=
specifies the width for the decision limits and central line in traditional graphics. This option is ignored when ODS Graphics is enabled.

XSYMBOL='*label*'
specifies the label for the central line on an ANOM chart or ANOM boxchart.

References

Fritzsch, K. and Hsu, J. C. (1997), "On Analysis of Means," in S. Panchapakesan and N. Balakrishnan, eds., *Advances in Statistical Decision Theory and Methodology*, Boston: Birkhäuser.

Halperin, M., Greenhouse, S. W., Cornfield, J., and Zalokar, J. (1955), "Tables of Percentage Points for the Studentized Maximum Absolute Deviate in Normal Samples," *Journal of the American Statistical Association*, 50, 185–195.

Hansen, E. (1990), "Making the 'Complex' Simple," in *Problem-Driven Case Studies in Quality Improvement, Second Annual Symposium*, Madison, WI: Center for Quality and Productivity Improvement.

Laplace, P. S. (1827), "Mémoire sur le flux et reflux lunaire atmosphérique," *Connaissance des Temps pour l'An 1830*, 3–18.

Nelson, L. S. (1983), "Exact Critical Values for Use with the Analysis of Means," *Journal of Quality Technology*, 15, 40–44.

Nelson, P. R. (1981), "Numerical Evaluation of an Equicorrelated Multivariate Non-central t Distribution," *Communications in Statistics—Simulation and Computation*, B10, 41–50.

Nelson, P. R. (1982a), "Exact Critical Points for the Analysis of Means," *Communications in Statistics—Theory and Methods*, 11, 699–709.

Nelson, P. R. (1982b), "Multivariate Normal and t Distributions with $\rho_{jk} = \alpha_j \alpha_k$," *Communications in Statistics—Simulation and Computation*, 11, 239–248.

Nelson, P. R. (1991), "Numerical Evaluation of Multivariate Normal Integrals with Correlations $\rho_{lj} = -\alpha_l \alpha_j$," in A. Öztürk and E. C. van der Meulen, eds., *Frontiers of Statistical Scientific Theory and Industrial Applications: Proceedings of the ICOSCO I Conference*, 97–114, Columbus, OH: American Sciences Press.

Nelson, P. R. (1993), "Additional Uses for the Analysis of Means and Extended Tables of Critical Values," *Technometrics*, 35, 61–71.

Nelson, P. R., Coffin, M., and Copeland, K. A. F. (2003), *Introductory Statistics for Engineering Experimentation*, San Diego: Academic Press/Elsevier.

Ott, E. R. (1967), "Analysis of Means: A Graphical Procedure," *Industrial Quality Control*, 24, 101–109, reprinted in *Journal of Quality Technology*, 15 (1983), 10–18.

Ott, E. R. (1975), *Process Quality Control: Troubleshooting and Interpretation of Data*, New York: McGraw-Hill.

Ramig, P. F. (1983), "Application of the Analysis of Means," *Journal of Quality Technology*, 15, 19–25.

Rodriguez, R. N. (1996), "Health Care Applications of Statistical Process Control: Examples Using the SAS System," in *Proceedings of the Twenty-First Annual SAS Users Group International Conference*, Cary, NC: SAS Institute Inc.

Soong, W. C. and Hsu, J. C. (1997), "Using Complex Integration to Compute Multivariate Normal Probabilities," *Journal of Computational and Graphical Statistics*, 6, 397–415.

Subject Index

Syntax Index