SAS/STAT® 9.22 User's Guide
The TTEST Procedure
(Book Excerpt)

Chapter 93

The TTEST Procedure

Contents

Overview: TTEST Procedure

The TTEST procedure performs *t* tests and computes confidence limits for one sample, paired observations, two independent samples, and the AB/BA crossover design. Two-sided, TOST (two one-sided test) equivalence, and upper and lower one-sided hypotheses are supported for means, mean differences, and mean ratios for either normal or lognormal data.

Table 93.1 summarizes the designs, analysis criteria, hypotheses, and distributional assumptions supported in the TTEST procedure, along with the syntax used to specify them.

Table 93.1   Features Supported in the TTEST Procedure

Feature	Syntax
Design	
One-sample	VAR statement
Paired	PAIRED statement
Two-independent-sample	CLASS statement, VAR statement
AB/BA crossover	VAR / CROSSOVER
Analysis Criterion	
Mean difference	PROC TTEST TEST=DIFF
Mean ratio	PROC TTEST TEST=RATIO
Hypothesis	
Two-sided	PROC TTEST SIDES=2
Equivalence	PROC TTEST TOST (*< lower* , *> upper*)
Lower one-sided	PROC TTEST SIDES=L
Upper one-sided	PROC TTEST SIDES=U
Distribution	
Normal	PROC TTEST DIST=NORMAL
Lognormal	PROC TTEST DIST=LOGNORMAL

FREQ and WEIGHT statements are available. Data can be input in the form of observations or, in certain cases, summary statistics. Output includes summary statistics; confidence limits for means, standard deviations, and coefficients of variation; hypothesis tests; and a variety of graphical displays, including histograms, densities, box plots, confidence intervals, Q-Q plots, profiles, and agreement plots.

PROC TTEST now uses ODS Graphics to create graphs as part of its output. For general information about ODS Graphics, see Chapter 21, "Statistical Graphics Using ODS." For specific information

about the statistical graphics available with the TTEST procedure, see the PLOTS option in the PROC TTEST statement and the section "ODS Graphics" on page 7918.

Getting Started: TTEST Procedure

One-Sample *t* Test

A one-sample *t* test can be used to compare a sample mean to a given value. This example, taken from Huntsberger and Billingsley (1989, p. 290), tests whether the mean length of a certain type of court case is more than 80 days by using 20 randomly chosen cases. The data are read by the following DATA step:

```
data time;
   input time @@;
   datalines;
 43   90   84   87   116    95   86    99    93   92
121   71   66   98    79   102   60   112   105   98
;
run;
```

The only variable in the data set, time, is assumed to be normally distributed. The trailing at signs (@@) indicate that there is more than one observation on a line. The following statements invoke PROC TTEST for a one-sample *t* test:

```
ods graphics on;

proc ttest h0=80 plots(showh0) sides=u alpha=0.1;
   var time;
run;

ods graphics off;
```

The ODS GRAPHICS statement requests graphical output. The VAR statement indicates that the time variable is being studied, while the H0= option specifies that the mean of the time variable should be compared to the null value 80 rather than the default of 0. The PLOTS(SHOWH0) option requests that this null value be displayed on all relevant graphs. The SIDES=U option reflects the focus of the research question, namely whether the mean court case length is *greater than* 80 days, rather than *different than* 80 days (in which case you would use the default SIDES=2 option). The ALPHA=0.1 option requests 90% confidence intervals rather than the default 95% confidence intervals. The output is displayed in Figure 93.1.

Figure 93.1 One-Sample *t* Test Results

```
                        The TTEST Procedure

                         Variable:  time

    N         Mean       Std Dev      Std Err      Minimum      Maximum

    20      89.8500      19.1456       4.2811      43.0000        121.0


         Mean         90% CL Mean           Std Dev        90% CL Std Dev

      89.8500      84.1659  Infty          19.1456       15.2002  26.2374


                          DF     t Value      Pr > t

                          19       2.30       0.0164
```

Summary statistics appear at the top of the output. The sample size (N), mean, standard deviation, and standard error are displayed with the minimum and maximum values of the time variable. The 90% confidence limits for the mean and standard deviation are shown next. Due to the SIDES=U option, the interval for the mean is an upper one-sided interval with a finite lower bound (84.1659 days). The limits for the standard deviation are the equal-tailed variety, per the default CI=EQUAL option in the PROC TTEST statement. At the bottom of the output are the degrees of freedom, *t* statistic value, and *p*-value for the *t* test. At the 10% α level, this test indicates that the mean length of the court cases is significantly greater than from 80 days ($t = 2.30$, $p = 0.0164$).

The summary panel in Figure 93.2 shows a histogram with overlaid normal and kernel densities, a box plot, the 90% confidence interval for the mean, and the null value of 80 days.

Figure 93.2 Summary Panel
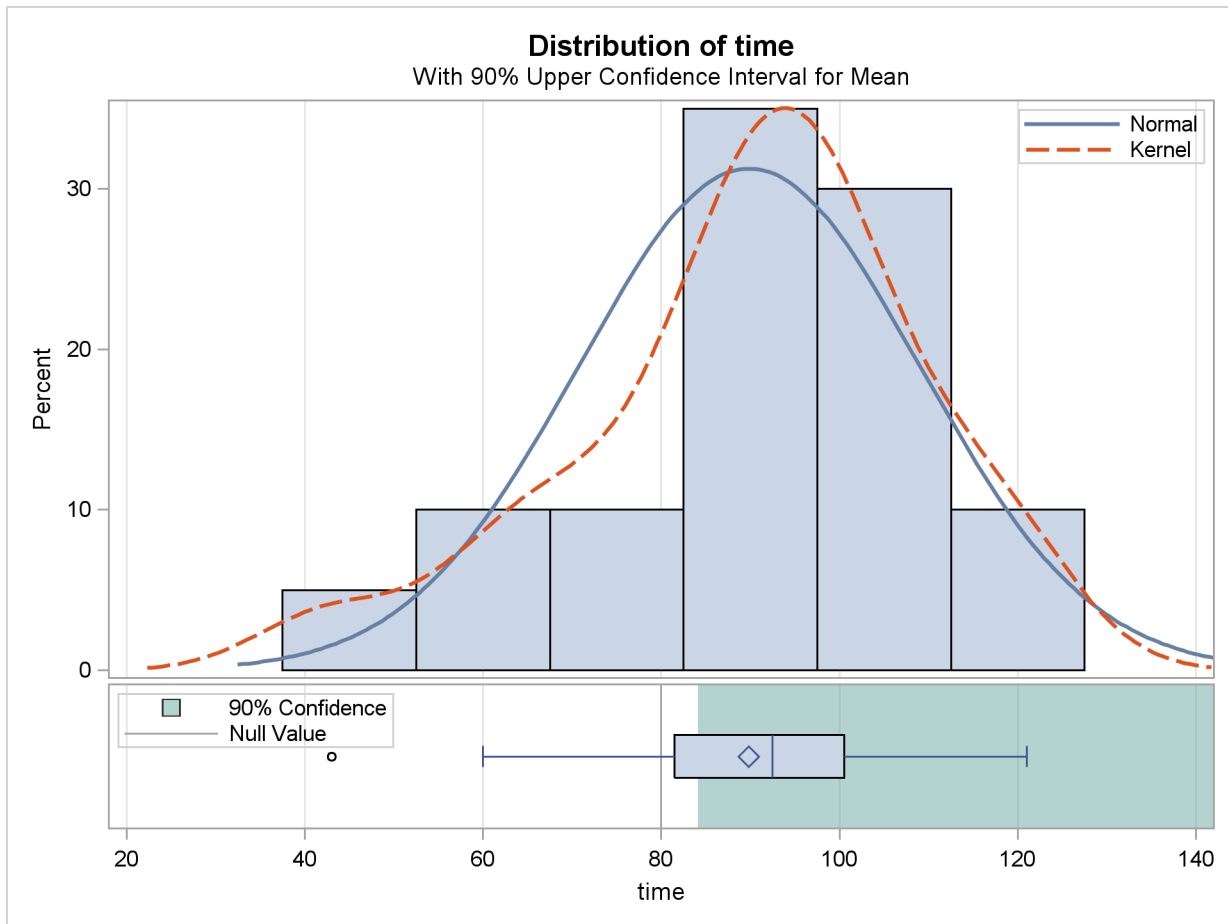
The confidence interval excludes the null value, consistent with the rejection of the null hypothesis at $\alpha = 0.1$.

The Q-Q plot in Figure 93.3 assesses the normality assumption.

Figure 93.3 Q-Q Plot
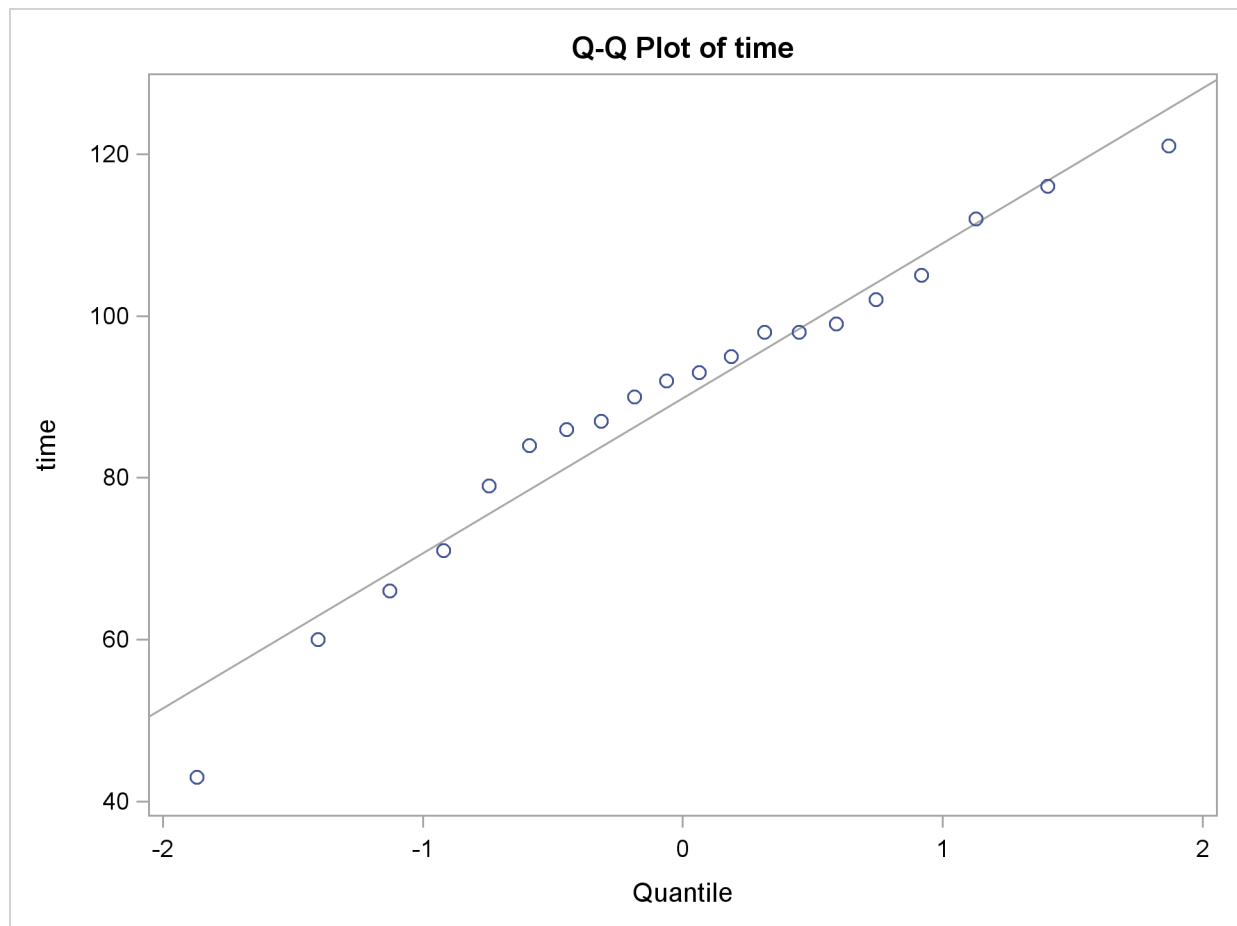
The curvilinear shape of the Q-Q plot suggests a possible slight deviation from normality. You could use the UNIVARIATE procedure with the NORMAL option to numerically check the normality assumptions.

Comparing Group Means

If you want to compare values obtained from two different groups, and if the groups are independent of each other and the data are normally or lognormally distributed in each group, then a group t test can be used. Examples of such group comparisons include the following:

- test scores for two third-grade classes, where one of the classes receives tutoring

- fuel efficiency readings of two automobile nameplates, where each nameplate uses the same fuel

- sunburn scores for two sunblock lotions, each applied to a different group of people

- political attitude scores of males and females

In the following example, the golf scores for males and females in a physical education class are compared. The sample sizes from each population are equal, but this is not required for further analysis. The scores are thought to be approximately normally distributed within gender. The data are read by the following statements:

```
data scores;
   input Gender $ Score @@;
   datalines;
f 75   f 76   f 80   f 77   f 80   f 77   f 73
m 82   m 80   m 85   m 85   m 78   m 87   m 82
;
run;
```

The dollar sign ($) following Gender in the INPUT statement indicates that Gender is a character variable. The trailing at signs (@@) enable the procedure to read more than one observation per line.

You can use a group *t* test to determine whether the mean golf score for the men in the class differs significantly from the mean score for the women. If you also suspect that the distributions of the golf scores of males and females have unequal variances, then you might want to specify the COCHRAN option in order to use the Cochran approximation (in addition to the Satterthwaite approximation, which is included by default). The following statements invoke PROC TTEST for the case of unequal variances, along with both types of confidence limits for the pooled standard deviation.

```
ods graphics on;

proc ttest cochran ci=equal umpu;
   class Gender;
   var Score;
run;

ods graphics off;
```

The ODS GRAPHICS statement requests graphical output. The CLASS statement contains the variable that distinguishes the groups being compared, and the VAR statement specifies the response variable to be used in calculations. The COCHRAN option produces *p*-values for the unequal variance situation by using the Cochran and Cox (1950) approximation. Equal-tailed and uniformly most powerful unbiased (UMPU) confidence intervals for σ are requested by the CI= option. Output from these statements is displayed in Figure 93.4 through Figure 93.7.

Figure 93.4 Simple Statistics

```
                        The TTEST Procedure

                         Variable:  Score

   Gender        N       Mean     Std Dev     Std Err     Minimum     Maximum

   f             7    76.8571      2.5448      0.9619     73.0000     80.0000
   m             7    82.7143      3.1472      1.1895     78.0000     87.0000
   Diff (1-2)        -5.8571      2.8619      1.5298
```

Simple statistics for the two populations being compared, as well as for the difference of the means between the populations, are displayed in Figure 93.4. The Gender column indicates the population corresponding to the statistics in that row. The sample size (N), mean, standard deviation, standard error, and minimum and maximum values are displayed.

Confidence limits for means and standard deviations are shown in Figure 93.5.

Figure 93.5 Simple Statistics

```
   Gender        Method                Mean      95% CL Mean        Std Dev

   f                               76.8571     74.5036  79.2107      2.5448
   m                               82.7143     79.8036  85.6249      3.1472
   Diff (1-2)    Pooled            -5.8571     -9.1902  -2.5241      2.8619
   Diff (1-2)    Satterthwaite     -5.8571     -9.2064  -2.5078

                                                              95% UMPU CL
    Gender        Method         95% CL Std Dev               Std Dev

    f                             1.6399    5.6039        1.5634    5.2219
    m                             2.0280    6.9303        1.9335    6.4579
    Diff (1-2)    Pooled          2.0522    4.7242        2.0019    4.5727
    Diff (1-2)    Satterthwaite
```

For the mean differences, both pooled (assuming equal variances for males and females) and Satterthwaite (assuming unequal variances) 95% intervals are shown. The confidence limits for the standard deviations are of the equal-tailed variety.

The test statistics, associated degrees of freedom, and *p*-values are displayed in Figure 93.6.

Figure 93.6 *t* Tests

```
       Method            Variances        DF     t Value     Pr > |t|

       Pooled            Equal            12      -3.83       0.0024
       Satterthwaite     Unequal      11.496      -3.83       0.0026
       Cochran           Unequal          6       -3.83       0.0087
```

The Method column denotes which *t* test is being used for that row, and the Variances column indicates what assumption about variances is being made. The pooled test assumes that the two populations have equal variances and uses degrees of freedom $n_1 + n_2 - 2$, where n_1 and n_2 are the sample sizes for the two populations. The remaining two tests do not assume that the populations have equal variances. The Satterthwaite test uses the Satterthwaite approximation for degrees of freedom, while the Cochran test uses the Cochran and Cox approximation for the *p*-value. All three tests result in highly significant *p*-values, supporting the conclusion of a significant difference between males' and females' golf scores.

The "Equality of Variances" test in Figure 93.7 reveals insufficient evidence of unequal variances (the Folded F statistic $F' = 1.53$, with $p = 0.6189$).

Figure 93.7 Tests of Equality of Variances

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	6	6	1.53	0.6189

The summary panel in Figure 93.8 shows comparative histograms, normal and kernel densities, and box plots, comparing the distribution of golf scores between genders.

Figure 93.8 Summary Panel
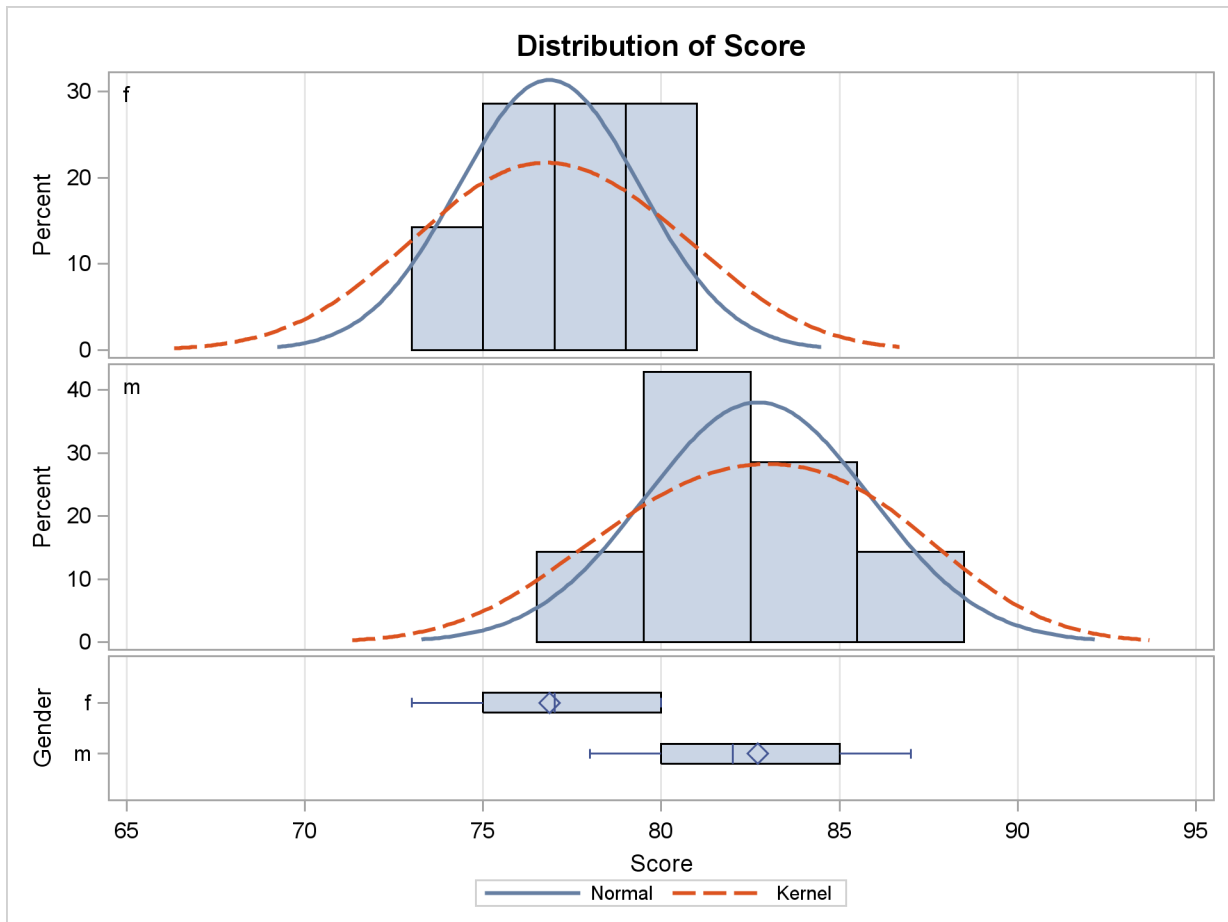
The Q-Q plots in Output 93.9 assess the normality assumption for each gender.

Figure 93.9 Q-Q Plot
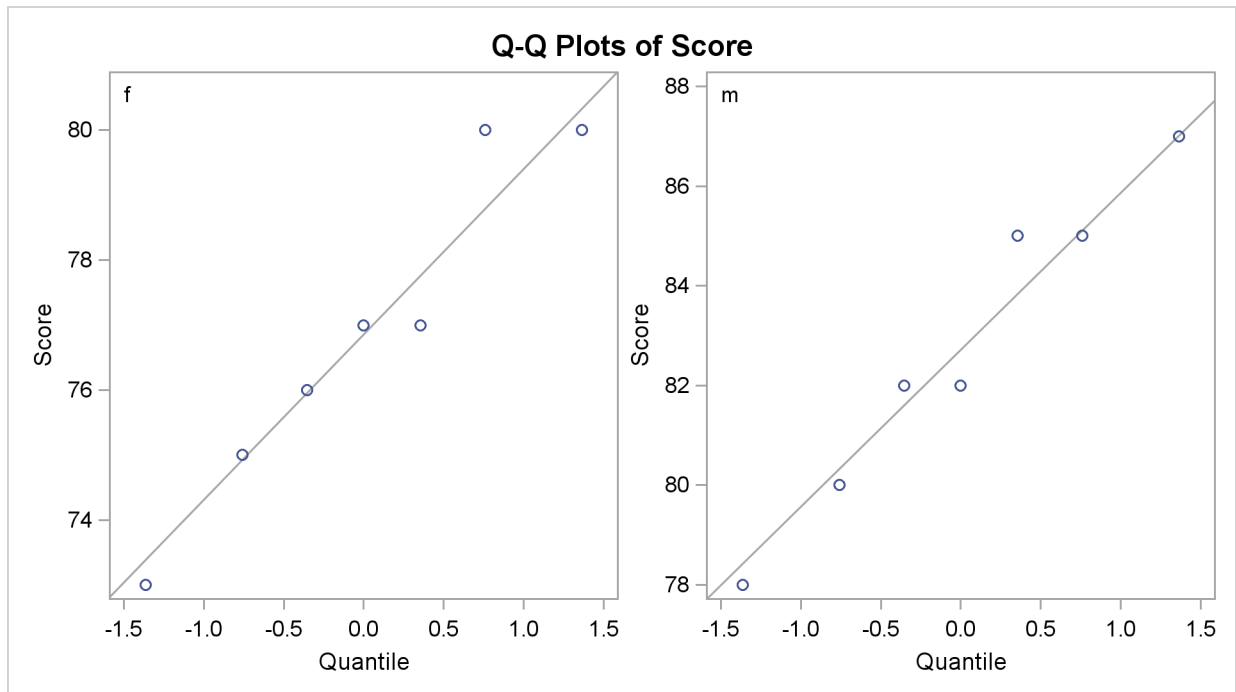
The plots for both males and females show no obvious deviations from normality. You can check the assumption of normality more rigorously by using PROC UNIVARIATE with the NORMAL option; if the assumption of normality is not reasonable, you should analyze the data with the nonparametric Wilcoxon rank sum test by using PROC NPAR1WAY.

Syntax: TTEST Procedure

The following statements are available in PROC TTEST:

> **PROC TTEST** < *options* > ;
>> **CLASS** *variable* ;
>> **PAIRED** *variables* ;
>> **BY** *variables* ;
>> **VAR** *variables* < / *options* > ;
>> **FREQ** *variable* ;
>> **WEIGHT** *variable* ;

No statement can be used more than once. There is no restriction on the order of the statements after the PROC TTEST statement.

PROC TTEST Statement

PROC TTEST *< options >* ;

The PROC TTEST statement invokes the procedure. Table 93.2 summarizes the options in the PROC TTEST statement by function. The options are then described fully in alphabetical order.

Table 93.2 PROC TTEST Statement Options

Option	Description
Basic Options	
DATA=	Specifies input data set
ORDER=	Determines sort order of CLASS variable or CROSSOVER treatment variables
Analysis Options	
ALPHA=	Specifies 1− confidence level
DIST=	Specifies distributional assumption (normal or lognormal)
H0=	Specifies null value
SIDES=	Specifies number of sides and direction
TEST=	Specifies test criterion (difference or ratio)
TOST	Requests equivalence test and specifies bounds
Displayed Output	
CI=	Requests confidence interval for standard deviation or CV
COCHRAN	Requests Cochran t test
PLOTS	Produces ODS statistical graphics
Output Ordering	
BYVAR	Groups results by PAIRED or VAR variables
NOBYVAR	Groups results by tables

The following options can appear in the PROC TTEST statement.

ALPHA=p

specifies that confidence intervals (except test-based mean confidence intervals when the TOST option is used) are to be $100(1 − p)\%$ confidence intervals, where $0 < p < 1$. When the TOST option is used, the test-based mean confidence intervals are $100(1 − 2p)\%$ confidence intervals. By default, PROC TTEST uses ALPHA=0.05. If p is 0 or less, or 1 or more, an error message is printed.

BYVAR

groups the results by the PAIRED or VAR variables. The BYVAR option is enabled by default. Note that this represents a change from previous releases for how the results are grouped with

respect to variables and tables. Prior to SAS 9.2, multiple variables were included in each table, similar to the new NOBYVAR option.

CI=EQUAL | UMPU | NONE

CL=EQUAL | UMPU | NONE

specifies whether a confidence interval is displayed for σ and, if so, what kind. The CI=EQUAL option specifies an equal-tailed confidence interval, and it is the default. The CI=UMPU option specifies an interval based on the uniformly most powerful unbiased test of $H_0 : \sigma = \sigma_0$. The CI=NONE option requests that no confidence interval be displayed for σ. The values EQUAL and UMPU together request that both types of confidence intervals be displayed. If the value NONE is specified with one or both of the values EQUAL and UMPU, NONE takes precedence. For more information, see the section "Two-Independent-Sample Design" on page 7907.

COCHRAN

requests the Cochran and Cox (1950) approximation of the probability level for the unequal variances situation. For more information, see the section "Two-Independent-Sample Design" on page 7907.

DATA=_SAS-data-set_

names the SAS data set for the procedure to use. By default, PROC TTEST uses the most recently created SAS data set. The input data set can contain summary statistics of the observations instead of the observations themselves. The number, mean, and standard deviation of the observations are required for each BY group (one sample and paired differences) or for each class within each BY group (two samples). For more information about the DATA= option, see the section "Input Data Set of Statistics" on page 7901.

DIST=LOGNORMAL | NORMAL

specifies the underlying distribution assumed for the data. The default is NORMAL, unless TEST=RATIO is specified, in which case the default is LOGNORMAL.

H0=_m_

requests tests against a null value of _m_, unless the TOST option is used, in which case _m_ is merely used to derive the lower and upper equivalence bounds. For the crossover design, the value _m_ applies for both treatment and period tests. By default, PROC TTEST uses H0=0 when TEST=DIFF (or DIST=NORMAL for a one-sample design) and H0=1 when TEST=RATIO (or DIST=LOGNORMAL for a one-sample design).

NOBYVAR

includes all PAIRED or VAR variables together in each output table. If the NOBYVAR option is not specified, then the BYVAR option is enabled, grouping the results by the PAIRED and VAR variables.

ORDER=DATA | FORMATTED | FREQ | INTERNAL | MIXED

specifies the order in which to sort the levels of the classification variables (which are specified in the CLASS statement) and treatment variables (which are specified in the CROSSOVER option in the VAR statement). The default is ORDER=MIXED, which corresponds to the ordering in releases previous to SAS 9.2.

This option applies to the levels for all classification or treatment variables, except when you use the ORDER=FORMATTED option with numeric classification or treatment variables that have no explicit format. With this option, the levels of such variables are ordered by their internal value.

The ORDER= option can take the following values:

Value of ORDER=	Levels Sorted By
DATA	Order of appearance in the input data set.
FORMATTED	External formatted value, except for numeric variables with no explicit format, which are sorted by their unformatted (internal) value.
FREQ	Descending frequency count; levels with the most observations come first in the order. In the event of a tie, ORDER=MIXED is used.
INTERNAL	Unformatted value.
MIXED	Same as ORDER=FORMATTED if the unformatted variable is character-valued; same as ORDER=INTERNAL otherwise (the unformatted variable is numeric-valued).

For FORMATTED and INTERNAL, the sort order is machine-dependent.

For more information about sorting order, see the chapter on the SORT procedure in the *Base SAS Procedures Guide* and the discussion of BY-group processing in *SAS Language Reference: Concepts*.

PLOTS < **(**global-plot-options**)** > < = plot-request < (options) > >

PLOTS < **(**global-plot-options**)** > < = **(**plot-request < (options) > < ... plot-request < (options) > >**)** >
controls the plots produced through ODS Graphics. When you specify only one plot request, you can omit the parentheses around the plot request. Here are some examples:

```
plots=none
plots=(histogram boxplot interval qq profiles agreement)
plots(unpack)=summary
plots(showh0)=interval(type=pergroup)
plots=(summary(unpack) interval(type=period))
```

You must enable ODS Graphics before requesting plots, for example, like this:

```
ods graphics on;

proc ttest plots=all;
   var oxygen;
run;

ods graphics off;
```

For general information about ODS Graphics, see Chapter 21, "Statistical Graphics Using ODS." If you have enabled ODS Graphics but do not specify the PLOTS option, then PROC TTEST produces a default set of plots. (**NOTE:** The graphical results are unavailable if your input data set contains summary statistics rather than observation values.)

For a one-sample design, the default plots are the following:

- summary plot (histogram with overlaid normal and kernel densities, box plot, and confidence interval band)
- Q-Q plot

For a two-independent-sample design, the default plots are the following:

- summary plot (comparative histograms with overlaid densities and box plots)
- Q-Q plot

For a paired design, the default plots are the following:

- summary plot (histogram, densities, box plot, and confidence interval) of the difference or ratio
- Q-Q plot of the difference or ratio
- profiles plot
- agreement plot

For a crossover design, the default plots are the following:

- comparative histograms with overlaid densities by treatment and period
- comparative box plots by treatment and period
- Q-Q plots by treatment and period
- profiles over treatment plot
- agreement of treatments plot

For more detailed descriptions of plots, see the section "Interpreting Graphs" on page 7919.

The global plot options include the following:

ONLY
> suppresses the default plots. Only plots specifically requested are displayed.

SHOWH0
SHOWNULL
> shows the null value (as specified by the H0= option in the PROC TTEST statement) in all relevant plots. For one-sample and paired designs, the null value can appear in SUMMARY, BOX, and INTERVAL. For two-independent-sample and crossover designs, the null value can appear only in INTERVAL.

UNPACKPANEL

UNPACK

> suppresses paneling. By default, multiple plots can appear in some output panels. Specify UNPACKPANEL to get each plot in a separate panel. You can specify PLOTS(UNPACKPANEL) to unpack the default plots. You can also specify UNPACK as a suboption with SUMMARY.

The plot requests include the following:

ALL

> produces all appropriate plots. You can specify other options with ALL; for example, to request all plots and specify that intervals should be for the period difference in a crossover design, specify PLOTS=(ALL INTERVAL(TYPE=PERIOD)).

AGREEMENT

AGREEMENTPLOT

> requests an agreement plot. This plot is produced by default for paired and crossover designs, the only designs for which the AGREEMENT option is valid.
>
> For paired designs, the second response in each pair is plotted against the first response. See the section "Agreement Plots for Paired Designs" on page 7919 for further details.
>
> For crossover designs, the AGREEMENT plot request has the following options:

> **TYPE=PERIOD**
>
> > plots the response in the second period against the response in the first period. See the section "Period Agreement Plots for Crossover Designs" on page 7919 for further details.

> **TYPE=TREATMENT**
>
> > plots the response associated with the second treatment against the response associated with the first treatment. This is the default TYPE= option for crossover designs. See the section "Treatment Agreement Plots for Crossover Designs" on page 7920 for further details.

BOX

BOXPLOT

> requests a box plot or comparative box plots. This plot is produced by default for crossover designs. For other designs, a box plot appears as part of the SUMMARY plot by default.
>
> For one-sample and paired designs, a confidence interval for the mean is shown as a band in the background, along with the equivalence bounds if the TOST option is used in the PROC TTEST statement.
>
> For a two-independent-sample design, comparative box plots (one for each class) are shown. For a crossover design, comparative box plots for all four combinations of the two treatments and two periods are shown.
>
> See the section "Box Plots" on page 7920 for further details.

HISTOGRAM

HIST

HISTDENS

> requests a histogram or comparative histograms with overlaid normal and kernel densities. This plot is produced by default for crossover designs. For other designs, it appears as part of the SUMMARY plot by default.
>
> For one-sample and paired designs, the histogram and densities are based on the test criterion (which is the mean difference or ratio for a paired design). For a two-independent-sample design, comparative histograms (one for each class) are shown. For a crossover design, histograms for all four combinations of the two treatments and two periods are shown.
>
> See the section "Histograms" on page 7920 for further details.

INTERVAL

INTERVALPLOT

> requests plots of confidence interval for means.
>
> For a two-independent-sample design, the INTERVAL plot request has the following options:

> **TYPE=PERGROUP**
>
> > shows two separate two-sided confidence intervals, one for each class. This option cannot be used along with the SHOWH0 global plot option.

> **TYPE=TEST**
>
> > shows pooled and Satterthwaite confidence intervals. This is the default TYPE= option for two-independent-sample designs.

> For a crossover design, The INTERVAL plot request has the following options:

> **TYPE=PERGROUP**
>
> > shows four separate two-sided intervals, one for each treatment-by-period combination. This option cannot be used along with the SHOWH0 global plot option.

> **TYPE=PERIOD**
>
> > shows pooled and Satterthwaite confidence intervals for the period difference or ratio. This option is invalid if the IGNOREPERIOD option is used in the VAR statement.

> **TYPE=TREATMENT**
>
> > shows pooled and Satterthwaite confidence intervals for the treatment difference or ratio. This is the default TYPE= option for crossover designs.

> See the section "Confidence Intervals" on page 7920 for further details.

NONE

> suppresses all plots.

PROFILES

PROFILESPLOT

> requests a profiles plot. This plot is produced by default for paired and crossover designs, the only designs for which the PROFILES option is valid.
>
> For paired designs, a line is drawn for each observation from left to right connecting the first response to the second response. See the section "Profiles for Paired Designs" on page 7921 for further details.
>
> For crossover designs, the PROFILES plot request has the following options:

TYPE=PERIOD

> > shows response profiles over period, connecting the first period on the left to the second period on the right for each subject. See the section "Profiles over Period for Crossover Designs" on page 7921 for further details.

TYPE=TREATMENT

> > shows response profiles over treatment values, connecting the first treatment on the left to the second treatment on the right for each observation. This is the default TYPE= option for crossover designs. See the section "Profiles over Treatment for Crossover Designs" on page 7921 for further details.

QQ

QQPLOT

> requests a normal quantile-quantile (Q-Q) plot. This plot is produced by default for all designs.
>
> For two-sample designs, separate plots are shown for each class in a single panel. For crossover design, separate plots are shown for each treatment-by-period combination in a single panel.
>
> See the section "Q-Q Plots" on page 7921 for further details.

SUMMARY

SUMMARYPLOT

> requests HISTOGRAM and BOX plots together in a single panel, sharing common X axes. This plot is produced by default for one-sample, paired, and two-independent-sample designs, the only designs for which the SUMMARY option is valid. See the documentation for the BOX and HISTOGRAM plot requests for details. The SUMMARY plot request has the following option:

UNPACK

> > plots histograms with overlaid densities in one panel and box plots (along with confidence interval bands, if one-sample or paired design) in another. Note that specifying PLOTS(ONLY)=SUMMARY(UNPACK) is exactly the same as specifying PLOTS(ONLY)=(BOX HISTOGRAM).

SIDES=2 | L | U

SIDED=2 | L | U

SIDE=2 | L | U

specifies the number of sides (or tails) and direction of the statistical tests and test-based confidence intervals. The values are interpreted as follows:

SIDES=2 (the default) specifies two-sided tests and confidence intervals for means.

SIDES=L specifies lower one-sided tests, in which the alternative hypothesis indicates a mean less than the null value, and lower one-sided confidence intervals between minus infinity and the upper confidence limit.

SIDES=U specifies upper one-sided tests, in which the alternative hypothesis indicates a mean greater than the null value, and upper one-sided confidence intervals between the lower confidence limit and infinity.

TEST=DIFF | RATIO

specifies the test criterion. Use TEST=DIFF to test the difference of means and TEST=RATIO to test the ratio of means. The default is DIFF, unless DIST=LOGNORMAL is specified, in which case the default is RATIO. This option is ignored for one-sample designs.

TOST (< *lower* ,> *upper*)

requests Schuirman's TOST equivalence test. The *upper* equivalence bound must be specified. If TEST=DIFF, then the default value for the *lower* equivalence bound is $2m-upper$, where m is the value of the H0= option. If TEST=RATIO, then the default value for *lower* is $m/upper$.

BY Statement

BY *variables* **;**

You can specify a BY statement with PROC TTEST to obtain separate analyses on 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 TTEST 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*.

CLASS Statement

CLASS *variable* **;**

A CLASS statement giving the name of the classification (or grouping) variable must accompany the PROC TTEST statement in the two-independent-sample case. It should be omitted for the one-sample, paired, and AB/BA crossover designs. If it is used without the VAR statement, all numeric variables in the input data set (except those that appear in the CLASS, BY, FREQ, or WEIGHT statement) are included in the analysis.

The classification variable must have two, and only two, levels. PROC TTEST divides the observations into the two groups for the *t* test by using the levels of this variable. You can use either a numeric or a character variable in the CLASS statement.

Class levels are determined from the formatted values of the CLASS variable. Thus, you can use formats to define group levels. See the discussions of the FORMAT procedure, the FORMAT statement, formats, and informats in *SAS Language Reference: Dictionary*.

FREQ Statement

FREQ *variable* **;**

The *variable* in the FREQ statement identifies a variable that contains the frequency of occurrence of each observation. PROC TTEST treats each observation as if it appears *n* times, where *n* is the value of the FREQ variable for the observation. If the value is not an integer, only the integer portion is used. If the frequency value is less than 1 or is missing, the observation is not used in the analysis. When the FREQ statement is not specified, each observation is assigned a frequency of 1. The FREQ statement cannot be used if the DATA= data set contains statistics instead of the original observations.

PAIRED Statement

PAIRED *PairLists* **;**

The *PairLists* in the PAIRED statement identifies the variables to be compared in paired comparisons. You can use one or more *PairLists*. Variables or lists of variables are separated by an asterisk (*) or a colon (:). The asterisk requests comparisons between each variable on the left with each variable on the right. The colon requests comparisons between the first variable on the left and the first on the

right, the second on the left and the second on the right, and so forth. The number of variables on the left must equal the number on the right when the colon is used. The differences are calculated by taking the variable on the left minus the variable on the right for both the asterisk and colon. A pair formed by a variable with itself is ignored. Use the PAIRED statement only for paired comparisons. The CLASS and VAR statements cannot be used with the PAIRED statement.

Examples of the use of the asterisk and the colon are shown in Table 93.3.

Table 93.3 PAIRED Statement in the TTEST Procedure

These PAIRED statements...	yield these comparisons
`PAIRED A*B;`	A-B
`PAIRED A*B C*D;`	A-B and C-D
`PAIRED (A B)*(C D);`	A-C, A-D, B-C, and B-D
`PAIRED (A B)*(C B);`	A-C, A-B, and B-C
`PAIRED (A1-A2)*(B1-B2);`	A1-B1, A1-B2, A2-B1, and A2-B2
`PAIRED (A1-A2):(B1-B2);`	A1-B1 and A2-B2

VAR Statement

VAR *variables* < / *options* > ;

The VAR statement names the variables to be used in the analyses. One-sample comparisons are conducted when the VAR statement is used without the CROSSOVER option or CLASS statement. Two-independent-sample comparisons are conducted when the VAR statement is used with a CLASS statement.

An AB/BA crossover analysis is conducted when the CROSSOVER option is used in the VAR statement. In this case, you must specify an even number of variables. Each set of two variables represents the responses in the first and second periods of the AB/BA crossover design. For example, if you use the CROSSOVER option and specify `VAR x1 x2 x3 x4`, then you will get two analyses. One analysis will have x1 as the period 1 response and x2 as the period 2 response. The other analysis will have x3 as the period 1 response and x4 as the period 2 response.

The VAR statement cannot be used with the PAIRED statement. If the VAR statement is omitted, all numeric variables in the input data set (except a numeric variable appearing in the BY, CLASS, FREQ, or WEIGHT statement) are included in the analysis.

You can specify the following options after a slash (/):

CROSSOVER= (*variable1 variable2* **)**
> specifies the variables representing the treatment applied in each of the two periods in an AB/BA crossover design. The treatment variables must have two, and only two, levels. For any given observation, the levels for the two variables must be different, due to the restrictions of the AB/BA crossover design. You can use either numeric or character variables.

Treatment levels are determined from the formatted values of the variables. Thus, you can use formats to define the treatment levels. See the discussions of the FORMAT procedure, the FORMAT statement, formats, and informats in *SAS Language Reference: Dictionary*.

IGNOREPERIOD

indicates that the period effect shall be ignored—that is, assumed to be equal to 0 (if TEST=DIFF) or 1 (if TEST=RATIO). This assumption increases the degrees of freedom for the test of the treatment difference by one and is usually more powerful, but it risks incorrect results if there is actually a period effect.

WEIGHT Statement

> **WEIGHT** *variable* **;**

The WEIGHT statement weights each observation in the input data set by the value of the WEIGHT variable. The values of the WEIGHT variable can be nonintegral, and they are not truncated. Observations with negative, zero, or missing values for the WEIGHT variable are not used in the analyses. Each observation is assigned a weight of 1 when the WEIGHT statement is not used. The WEIGHT statement cannot be used with an input data set of summary statistics.

Details: TTEST Procedure

Input Data Set of Statistics

PROC TTEST accepts data containing either observation values or summary statistics. Observation values are supported for all analyses, whereas summary statistics are supported only for a subset of analyses. If the analysis involves the paired design, the AB/BA crossover design, or the lognormal distributional assumption (DIST=LOGNORMAL), then observation values must be used. The graphical results are unavailable if your input data set contains summary statistics rather than raw observed values.

PROC TTEST assumes that the DATA= data set contains statistics if it contains a character variable with name _TYPE_ or _STAT_. The TTEST procedure expects this character variable to contain the names of statistics. If both _TYPE_ and _STAT_ variables exist and are of type character, PROC TTEST expects _TYPE_ to contain the names of statistics including 'N', 'MEAN', and 'STD' for each BY group (or for each class within each BY group for two-sample *t* tests). If no 'N', 'MEAN', or 'STD' statistics exist, an error message is printed.

FREQ, WEIGHT, and PAIRED statements cannot be used with input data sets of statistics. BY, CLASS, and VAR statements are the same regardless of data set type. For paired comparisons, see

the _DIF_ values for the _TYPE_=T observations in output produced by the OUTSTATS= option in the PROC COMPARE statement (see the *Base SAS Procedures Guide*).

Missing Values

An observation is omitted from the calculations if it has a missing value for either the CLASS variable, a CROSSOVER variable, a PAIRED variable, the variable to be tested (in a one-sample or two-independent-sample design), or either of the two response variables (in a crossover design). If more than one variable or pair of variables is listed in the VAR statement, a missing value in one variable or pair does not eliminate the observation from the analysis of other nonmissing variables or variable pairs.

Computational Methods

This section describes the computational formulas for the estimates, confidence limits, and tests for each analysis in the TTEST procedure. The first subsection defines some common notation. The second subsection discusses the distinction between arithmetic and geometric means. The third subsection explains the concept of the coefficient of variation. The following four subsections address the four supported designs (one-sample, paired, two-independent-sample, and AB/BA crossover). The content in each of those subsections is divided into separate discussions according to different values of the DIST= and TEST= options in the PROC TTEST statement. The last subsection describes TOST equivalence analyses.

Common Notation

Table 93.4 displays notation for some of the commonly used symbols.

Table 93.4 Common Notation

Symbol	Description
μ	Population value of (arithmetic) mean
μ_0	Null value of test (value of H0= option in PROC TTEST statement)
σ^2	Population variance
σ	Population value of standard deviation
γ	Population value of geometric mean
CV	Population value of coefficient of variation (ratio of population standard deviation and population arithmetic mean)
α	Value of ALPHA= option in PROC TTEST statement
$t_{p,\nu}$	pth percentile of t distribution with ν degrees of freedom (d.f.)

Table 93.4 *continued*

Symbol	Description
F_{p, v_1, v_2}	pth percentile of F distribution with v_1 numerator d.f. and v_2 denominator d.f.
$\chi^2_{p,v}$	pth percentile of chi-square distribution with v d.f.

Arithmetic and Geometric Means

The *arithmetic mean* (more commonly called simply the *mean*) of the distribution of a random variable X is its expected value, $E(X)$. The arithmetic mean is the natural parameter of interest for a normal distribution because the distribution of the difference of normal random variables has a known normal distribution, and the arithmetic mean of a normal difference is equal to the difference of the individual arithmetic means. (No such convenient property holds for geometric means with normal data, with either differences or ratios.)

The usual estimate of an arithmetic mean is the sum of the values divided by the number of values:

$$\text{arithmetic mean} = \frac{1}{n} \sum_{i=1}^{n} y_i$$

The *geometric mean* of the distribution of a random variable X is $\exp(E(\log(X))$, the exponentiation of the mean of the natural logarithm. The geometric mean is the natural parameter of interest for a lognormal distribution because the distribution of a ratio of lognormal random variables has a known lognormal distribution, and the geometric mean of a lognormal ratio is equal to the ratio of the individual geometric means. (No such convenient property holds for arithmetic means with lognormal data, with either differences or ratios.)

The usual estimate of a geometric mean is the product of the values raised to the power $1/n$, where n is the number of values:

$$\text{geometric mean} = \left(\prod_{i=1}^{n} y_i \right)^{\frac{1}{n}}$$

Coefficient of Variation

The *coefficient of variation* (abbreviated "CV") of the distribution of a random variable X is the ratio of the standard deviation to the (arithmetic) mean, or $\sqrt{\text{Var}(X)}/E(X)$. Conceptually, it is a measure of the variability of X expressed in units corresponding to the mean of X.

For lognormal data, the CV is the natural measure of variability (rather than the standard deviation) because the CV is invariant to multiplication of a lognormal variable by a constant. For a two-independent-sample design, the assumption of equal CVs on a lognormal scale is analogous to the assumption of equal variances on the normal scale. When the CVs of two independent samples of lognormal data are assumed equal, the pooled estimate of variability is used.

One-Sample Design

Define the following notation:

$$n^\star = \text{number of observations in data set}$$
$$y_i = \text{value of } i\text{th observation, } i \in \{1, \ldots, n^\star\}$$
$$f_i = \text{frequency of } i\text{th observation, } i \in \{1, \ldots, n^\star\}$$
$$w_i = \text{weight of } i\text{th observation, } i \in \{1, \ldots, n^\star\}$$
$$n = \text{sample size} = \sum_i^{n^\star} f_i$$

Normal Data (DIST=NORMAL)

The mean estimate $\bar{y}$, standard deviation estimate s, and standard error SE are computed as follows:

$$\bar{y} = \frac{\sum_i^{n^\star} f_i w_i y_i}{\sum_i^{n^\star} f_i w_i}$$

$$s = \left(\frac{\sum_i^{n^\star} f_i w_i (y_i - \bar{y})^2}{n - 1} \right)^{\frac{1}{2}}$$

$$\text{SE} = \frac{s}{\sqrt{\sum_i^{n^\star} f_i w_i}}$$

The $100(1 - \alpha)\%$ confidence interval for the mean μ is

$$\left(\bar{y} - t_{1-\frac{\alpha}{2}, n-1} \text{SE} \;,\; \bar{y} + t_{1-\frac{\alpha}{2}, n-1} \text{SE} \right), \quad \text{SIDES=2}$$
$$\left(-\infty \;,\; \bar{y} + t_{1-\alpha, n-1} \text{SE} \right), \quad \text{SIDES=L}$$
$$\left(\bar{y} - t_{1-\alpha, n-1} \text{SE} \;,\; \infty \right), \quad \text{SIDES=U}$$

The t value for the test is computed as

$$t = \frac{\bar{y} - \mu_0}{\text{SE}}$$

The p-value of the test is computed as

$$p\text{-value} = \begin{cases} P\left(t^2 > F_{1-\alpha, 1, n-1} \right) \;, & \text{2-sided} \\ P\left(t < t_{\alpha, n-1} \right) \;, & \text{lower 1-sided} \\ P\left(t > t_{1-\alpha, n-1} \right) \;, & \text{upper 1-sided} \end{cases}$$

The equal-tailed confidence interval for the standard deviation (CI=EQUAL) is based on the acceptance region of the test of $H_0 : \sigma = \sigma_0$ that places an equal amount of area ($\frac{\alpha}{2}$) in each tail of the chi-square distribution:

$$\left\{ \chi^2_{\frac{\alpha}{2}, n-1} \leq \frac{(n-1)s^2}{\sigma_0^2} \leq \chi^2_{\frac{1-\alpha}{2}, n-1} \right\}$$

The acceptance region can be algebraically manipulated to give the following $100(1-\alpha)\%$ confidence interval for σ^2:

$$\left(\frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2},n-1}} \ , \ \frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2},n-1}} \right)$$

Taking the square root of each side yields the $100(1-\alpha)\%$ CI=EQUAL confidence interval for σ:

$$\left(\left(\frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2},n-1}} \right)^{\frac{1}{2}} \ , \ \left(\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2},n-1}} \right)^{\frac{1}{2}} \right)$$

The other confidence interval for the standard deviation (CI=UMPU) is derived from the uniformly most powerful unbiased test of $H_0 \colon \sigma = \sigma_0$ (Lehmann 1986). This test has acceptance region

$$\left\{ c_1 \le \frac{(n-1)s^2}{\sigma_0^2} \le c_2 \right\}$$

where the critical values c_1 and c_2 satisfy

$$\int_{c_1}^{c_2} f_{n-1}(y)dy = 1 - \alpha$$

and

$$\int_{c_1}^{c_2} y f_{n-1}(y)dy = (n-1)(1-\alpha)$$

where $f_\nu(y)$ is the p.d.f. of the chi-square distribution with ν degrees of freedom. This acceptance region can be algebraically manipulated to arrive at

$$P\left\{ \frac{(n-1)s^2}{c_2} \le \sigma^2 \le \frac{(n-1)s^2}{c_1} \right\} = 1 - \alpha$$

where c_1 and c_2 solve the preceding two integrals. To find the area in each tail of the chi-square distribution to which these two critical values correspond, solve $c_1 = \chi^2_{1-\alpha_2,n-1}$ and $c_2 = \chi^2_{\alpha_1,n-1}$ for α_1 and α_2; the resulting α_1 and α_2 sum to α. Hence, a $100(1-\alpha)\%$ confidence interval for σ^2 is given by

$$\left(\frac{(n-1)s^2}{\chi^2_{1-\alpha_2,n-1}} \ , \ \frac{(n-1)s^2}{\chi^2_{\alpha_1,n-1}} \right)$$

Taking the square root of each side yields the $100(1-\alpha)\%$ CI=UMPU confidence interval for σ:

$$\left(\left(\frac{(n-1)s^2}{\chi^2_{1-\alpha_2,n-1}} \right)^{\frac{1}{2}} \ , \ \left(\frac{(n-1)s^2}{\chi^2_{\alpha_1,n-1}} \right)^{\frac{1}{2}} \right)$$

Lognormal Data (DIST=LOGNORMAL)

The DIST=LOGNORMAL analysis is handled by log-transforming the data and null value, performing a DIST=NORMAL analysis, and then transforming the results back to the original scale. This simple technique is based on the properties of the lognormal distribution as discussed in Johnson, Kotz, and Balakrishnan (1994, Chapter 14).

Taking the natural logarithms of the observation values and the null value, define

$$z_i = \log(y_i) \ , \quad i \in \{1, \ldots, n^\star\}$$
$$\gamma_0 = \log(\mu_0)$$

First a DIST=NORMAL analysis is performed on z_i in place of y_i. The geometric mean estimate $\hat{\gamma}$ and CV estimate $\widehat{CV}$ of the original lognormal data are computed as follows:

$$\hat{\gamma} = \exp(\bar{y})$$
$$\widehat{CV} = \left(\exp(s^2) - 1\right)^{\frac{1}{2}}$$

The t value and p-value remain the same. The confidence limits for the geometric mean and CV on the original lognormal scale are computed from the confidence limits for the arithmetic mean and standard deviation in the DIST=NORMAL analysis on the log-transformed data, in the same way that $\hat{\gamma}$ is derived from $\bar{y}$ and $\widehat{CV}$ is derived from s.

Paired Design

Define the following notation:

$$n^\star = \text{number of observations in data set}$$
$$y_{1i} = \text{value of } i\text{th observation for first PAIRED variable}, \quad i \in \{1, \ldots, n^\star\}$$
$$y_{2i} = \text{value of } i\text{th observation for second PAIRED variable}, \quad i \in \{1, \ldots, n^\star\}$$
$$f_i = \text{frequency of } i\text{th observation}, \quad i \in \{1, \ldots, n^\star\}$$
$$w_i = \text{weight of } i\text{th observation}, \quad i \in \{1, \ldots, n^\star\}$$
$$n = \text{sample size} = \sum_i^{n^\star} f_i$$

Normal Difference (DIST=NORMAL TEST=DIFF)

The analysis is the same as the analysis for the one-sample design in the section "Normal Data (DIST=NORMAL)" on page 7904 based on the differences

$$d_i = y_{1i} - y_{2i} \ , \quad i \in \{1, \ldots, n^\star\}$$

Lognormal Ratio (DIST=LOGNORMAL TEST=RATIO)

The analysis is the same as the analysis for the one-sample design in the section "Lognormal Data (DIST=LOGNORMAL)" on page 7906 based on the ratios

$$r_i = y_{1i}/y_{2i} \ , \quad i \in \{1, \ldots, n^\star\}$$

Normal Ratio (DIST=NORMAL TEST=RATIO)

The hypothesis $H_0: \mu_1/\mu_2 = \mu_0$, where μ_1 and μ_2 are the means of the first and second PAIRED variables, respectively, can be rewritten as $H_0: \mu_1 - \mu_0\mu_2 = 0$. The t value and p-value are computed in the same way as in the one-sample design in the section "Normal Data (DIST=NORMAL)" on page 7904 based on the transformed values

$$z_i = y_{1i} - \mu_0 y_{2i} \ , \quad i \in \{1, \ldots, n^\star\}$$

Estimates and confidence limits are not computed for this situation.

Two-Independent-Sample Design

Define the following notation:

$$
\begin{aligned}
n_1^\star &= \text{number of observations at first class level} \\
n_2^\star &= \text{number of observations at second class level} \\
y_{1i} &= \text{value of } i\text{th observation at first class level}, \ i \in \{1, \ldots, n_1^\star\} \\
y_{2i} &= \text{value of } i\text{th observation at second class level}, \ i \in \{1, \ldots, n_2^\star\} \\
f_{1i} &= \text{frequency of } i\text{th observation at first class level}, \ i \in \{1, \ldots, n_1^\star\} \\
f_{2i} &= \text{frequency of } i\text{th observation at second class level}, \ i \in \{1, \ldots, n_2^\star\} \\
w_{1i} &= \text{weight of } i\text{th observation at first class level}, \ i \in \{1, \ldots, n_1^\star\} \\
w_{2i} &= \text{weight of } i\text{th observation at second class level}, \ i \in \{1, \ldots, n_2^\star\}
\end{aligned}
$$

$$n_1 = \text{sample size for first class level} = \sum_i^{n_1^\star} f_{1i}$$

$$n_2 = \text{sample size for second class level} = \sum_i^{n_2^\star} f_{2i}$$

Normal Difference (DIST=NORMAL TEST=DIFF)

Observations at the first class level are assumed to be distributed as $N(\mu_1, \sigma_1^2)$, and observations at the second class level are assumed to be distributed as $N(\mu_2, \sigma_2^2)$, where μ_1, μ_2, σ_1, and σ_2 are unknown.

The within-class-level mean estimates ($\bar{y}_1$ and $\bar{y}_2$), standard deviation estimates (s_1 and s_2), standard errors (SE_1 and SE_2), and confidence limits for means and standard deviations are computed in the same way as for the one-sample design in the section "Normal Data (DIST=NORMAL)" on page 7904.

The mean difference $\mu_1 - \mu_2 = \mu_d$ is estimated by

$$\bar{y}_d = \bar{y}_1 - \bar{y}_2$$

Under the assumption of equal variances ($\sigma_1^2 = \sigma_2^2$), the pooled estimate of the common standard deviation is

$$s_p = \left(\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \right)^{\frac{1}{2}}$$

The pooled standard error (the estimated standard deviation of $\bar{y}_d$ assuming equal variances) is

$$SE_p = s_p \left(\frac{1}{\sum_{i=1}^{n_1^\star} f_{1i} w_{1i}} + \frac{1}{\sum_{i=1}^{n_2^\star} f_{2i} w_{2i}} \right)^{\frac{1}{2}}$$

The pooled $100(1 - \alpha)\%$ confidence interval for the mean difference μ_d is

$$\left(\bar{y}_d - t_{1-\frac{\alpha}{2}, n_1+n_2-2} SE_p \ , \ \bar{y}_d + t_{1-\frac{\alpha}{2}, n_1+n_2-2} SE_p \right), \quad \text{SIDES=2}$$

$$\left(-\infty \ , \ \bar{y}_d + t_{1-\alpha, n_1+n_2-2} SE_p \right), \quad \text{SIDES=L}$$

$$\left(\bar{y}_d - t_{1-\alpha, n_1+n_2-2} SE_p \ , \ \infty \right), \quad \text{SIDES=U}$$

The t value for the pooled test is computed as

$$t_p = \frac{\bar{y}_d - \mu_0}{SE_p}$$

The p-value of the test is computed as

$$p\text{-value} = \begin{cases} P\left(t_p^2 > F_{1-\alpha, 1, n_1+n_2-2} \right) \ , & \text{2-sided} \\ P\left(t_p < t_{\alpha, n_1+n_2-2} \right) \ , & \text{lower 1-sided} \\ P\left(t_p > t_{1-\alpha, n_1+n_2-2} \right) \ , & \text{upper 1-sided} \end{cases}$$

Under the assumption of unequal variances (the Behrens-Fisher problem), the unpooled standard error is computed as

$$SE_u = \left(\frac{s_1^2}{\sum_{i=1}^{n_1^\star} f_{1i} w_{1i}} + \frac{s_2^2}{\sum_{i=1}^{n_2^\star} f_{2i} w_{2i}} \right)^{\frac{1}{2}}$$

Satterthwaite's (1946) approximation for the degrees of freedom, extended to accommodate weights, is computed as

$$df_u = \frac{SE_u^4}{\dfrac{s_1^4}{(n_1-1)\left(\sum_{i=1}^{n_1^\star} f_{1i} w_{1i} \right)^2} + \dfrac{s_2^4}{(n_2-1)\left(\sum_{i=1}^{n_2^\star} f_{2i} w_{2i} \right)^2}}$$

The unpooled Satterthwaite $100(1 - \alpha)\%$ confidence interval for the mean difference μ_d is

$$\left(\bar{y}_d - t_{1-\frac{\alpha}{2}, df_u} SE_u \ , \ \bar{y}_d + t_{1-\frac{\alpha}{2}, df_u} SE_u \right), \quad \text{SIDES=2}$$

$$\left(-\infty \ , \ \bar{y}_d + t_{1-\alpha, df_u} SE_u \right), \quad \text{SIDES=L}$$

$$\left(\bar{y}_d - t_{1-\alpha, df_u} SE_u \ , \ \infty \right), \quad \text{SIDES=U}$$

The t value for the unpooled Satterthwaite test is computed as

$$t_u = \frac{\bar{y}_d - \mu_0}{\text{SE}_u}$$

The p-value of the unpooled Satterthwaite test is computed as

$$p\text{-value} = \begin{cases} P\left(t_u^2 > F_{1-\alpha,1,\text{df}_u}\right) & , \quad \text{2-sided} \\ P\left(t_u < t_{\alpha,\text{df}_u}\right) & , \qquad \text{lower 1-sided} \\ P\left(t_u > t_{1-\alpha,\text{df}_u}\right) & , \qquad \text{upper 1-sided} \end{cases}$$

When the COCHRAN option is specified in the PROC TTEST statement, the Cochran and Cox (1950) approximation of the p-value of the t_u statistic is the value of p such that

$$t_u = \frac{\left(\frac{s_1^2}{\sum_{i=1}^{n_1^\star} f_{1i}\,w_{1i}}\right) t_1 + \left(\frac{s_2^2}{\sum_{i=1}^{n_2^\star} f_{2i}\,w_{2i}}\right) t_2}{\left(\frac{s_1^2}{\sum_{i=1}^{n_1^\star} f_{1i}\,w_{1i}}\right) + \left(\frac{s_2^2}{\sum_{i=1}^{n_2^\star} f_{2i}\,w_{2i}}\right)}$$

where t_1 and t_2 are the critical values of the t distribution corresponding to a significance level of p and sample sizes of n_1 and n_2, respectively. The number of degrees of freedom is undefined when $n_1 \neq n_2$. In general, the Cochran and Cox test tends to be conservative (Lee and Gurland 1975).

The $100(1 - \alpha)\%$ CI=EQUAL and CI=UMPU confidence intervals for the common population standard deviation σ assuming equal variances are computed as discussed in the section "Normal Data (DIST=NORMAL)" on page 7904 for the one-sample design, except replacing s^2 by s_p^2 and $(n-1)$ by $(n_1 + n_2 - 1)$.

The folded form of the F statistic, F', tests the hypothesis that the variances are equal (Steel and Torrie 1980), where

$$F' = \frac{\max(s_1^2, s_2^2)}{\min(s_1^2, s_2^2)}$$

A test of F' is a two-tailed F test because you do not specify which variance you expect to be larger. The p-value gives the probability of a greater F value under the null hypothesis that $\sigma_1^2 = \sigma_2^2$. Note that this test is not very robust to violations of the assumption that the data are normally distributed, and thus it is not recommended without confidence in the normality assumption.

Lognormal Ratio (DIST=LOGNORMAL TEST=RATIO)

The DIST=LOGNORMAL analysis is handled by log-transforming the data and null value, performing a DIST=NORMAL analysis, and then transforming the results back to the original scale. See the section "Normal Data (DIST=NORMAL)" on page 7904 for the one-sample design for details on how the DIST=NORMAL computations for means and standard deviations are transformed into the DIST=LOGNORMAL results for geometric means and CVs. As mentioned in the section "Coefficient of Variation" on page 7903, the assumption of equal CVs on the lognormal scale is analogous to the assumption of equal variances on the normal scale.

Normal Ratio (DIST=NORMAL TEST=RATIO)

The distributional assumptions, equality of variances test, and within-class-level mean estimates ($\bar{y}_1$ and $\bar{y}_2$), standard deviation estimates (s_1 and s_2), standard errors (SE_1 and SE_2), and confidence limits for means and standard deviations are the same as in the section "Normal Difference (DIST=NORMAL TEST=DIFF)" on page 7907 for the two-independent-sample design.

The mean ratio $\mu_1/\mu_2 = \mu_r$ is estimated by

$$\hat{\mu}_r = \bar{y}_1/\bar{y}_2$$

No estimates or confidence intervals for the ratio of standard deviations are computed.

Under the assumption of equal variances ($\sigma_1^2 = \sigma_2^2$), the pooled confidence interval for the mean ratio is the Fieller (1954) confidence interval, extended to accommodate weights. Let

$$a_p = \frac{s_p^2 t_{1-\frac{\alpha}{2}, n_1+n_2-2}^2}{\sum_{i=1}^{n_2^\star} f_{2i} w_{2i}} - \bar{y}_2^2$$

$$b_p = \bar{y}_1 \bar{y}_2$$

$$c_p = \frac{s_p^2 t_{1-\frac{\alpha}{2}, n_1+n_2-2}^2}{\sum_{i=1}^{n_1^\star} f_{1i} w_{1i}} - \bar{y}_1^2$$

where s_p is the pooled standard deviation defined in the section "Normal Difference (DIST=NORMAL TEST=DIFF)" on page 7907 for the two-independent-sample design. If $a_p \geq 0$ (which occurs when $\bar{y}_2$ is too close to zero), then the pooled two-sided $100(1-\alpha)\%$ Fieller confidence interval for μ_r does not exist. If $a < 0$, then the interval is

$$\left(-\frac{b_p}{a_p} + \frac{(b_p^2 - a_p c_p)^{\frac{1}{2}}}{a_p} \quad , -\frac{b_p}{a_p} - \frac{(b_p^2 - a_p c_p)^{\frac{1}{2}}}{a_p} \right)$$

For the one-sided intervals, let

$$a_p^\star = \frac{s_p^2 t_{1-\alpha, n_1+n_2-2}^2}{\sum_{i=1}^{n_2^\star} f_{2i} w_{2i}} - \bar{y}_2^2$$

$$c_p^\star = \frac{s_p^2 t_{1-\alpha, n_1+n_2-2}^2}{\sum_{i=1}^{n_1^\star} f_{1i} w_{1i}} - \bar{y}_1^2$$

which differ from a_p and c_p only in the use of α in place of $\alpha/2$. If $a_p^\star \geq 0$, then the pooled one-sided $100(1-\alpha)\%$ Fieller confidence intervals for μ_r do not exist. If $a_p^\star < 0$, then the intervals are

$$\left(-\infty \quad , -\frac{b_p}{a_p^\star} - \frac{(b_p^2 - a_p^\star c_p^\star)^{\frac{1}{2}}}{a_p^\star} \right), \quad \text{SIDES=L}$$

$$\left(-\frac{b_p}{a_p^\star} + \frac{(b_p^2 - a_p^\star c_p^\star)^{\frac{1}{2}}}{a_p^\star} \quad , \quad \infty \right), \quad \text{SIDES=U}$$

The pooled t test assuming equal variances is the Sasabuchi (1988a, 1988b) test. The hypothesis $H_0\colon \mu_r = \mu_0$ is rewritten as $H_0\colon \mu_1 - \mu_0\mu_2 = 0$, and the pooled t test in the section "Normal Difference (DIST=NORMAL TEST=DIFF)" on page 7907 for the two-independent-sample design is conducted on the original y_{1i} values ($i \in \{1, \ldots, n_1^{\star}\}$) and transformed values of y_{2i}

$$y_{2i}^{\star} = \mu_0 y_{2i} \ , \quad i \in \{1, \ldots, n_2^{\star}\}$$

with a null difference of 0. The t value for the Sasabuchi pooled test is computed as

$$t_p = \frac{\bar{y}_1 - \mu_0 \bar{y}_2}{s_p \left(\dfrac{1}{\sum_{i=1}^{n_1^{\star}} f_{1i} w_{1i}} + \dfrac{\mu_0^2}{\sum_{i=1}^{n_2^{\star}} f_{2i} w_{2i}} \right)^{\frac{1}{2}}}$$

The p-value of the test is computed as

$$p\text{-value} = \begin{cases} P\left(t_p^2 > F_{1-\alpha,1,n_1+n_2-2}\right) \ , & \text{2-sided} \\ P\left(t_p < t_{\alpha,n_1+n_2-2}\right) \ , & \text{lower 1-sided} \\ P\left(t_p > t_{1-\alpha,n_1+n_2-2}\right) \ , & \text{upper 1-sided} \end{cases}$$

Under the assumption of unequal variances, the unpooled Satterthwaite-based confidence interval for the mean ratio μ_r is computed according to the method in Dilba, Schaarschmidt, and Hothorn (2006), extended to accommodate weights. The degrees of freedom are computed as

$$\mathrm{df}_u = \frac{\left(\dfrac{s_1^2}{\sum_{i=1}^{n_1^{\star}} f_{1i} w_{1i}} + \dfrac{\hat{\mu}_r^2 s_2^2}{\sum_{i=1}^{n_2^{\star}} f_{2i} w_{2i}} \right)^2}{\dfrac{s_1^4}{(n_1-1)\left(\sum_{i=1}^{n_1^{\star}} f_{1i} w_{1i}\right)^2} + \dfrac{\hat{\mu}_r^4 s_2^4}{(n_2-1)\left(\sum_{i=1}^{n_2^{\star}} f_{2i} w_{2i}\right)^2}}$$

Note that the estimate $\hat{\mu}_r = \bar{y}_1/\bar{y}_2$ is used in df_u. Let

$$a_u = \frac{s_2^2 t_{1-\frac{\alpha}{2},\mathrm{df}_u}^2}{\sum_{i=1}^{n_2^{\star}} f_{2i} w_{2i}} - \bar{y}_2^2$$

$$b_u = \bar{y}_1 \bar{y}_2$$

$$c_u = \frac{s_1^2 t_{1-\frac{\alpha}{2},\mathrm{df}_u}^2}{\sum_{i=1}^{n_1^{\star}} f_{1i} w_{1i}} - \bar{y}_1^2$$

where s_1 and s_2 are the within-class-level standard deviations defined in the section "Normal Difference (DIST=NORMAL TEST=DIFF)" on page 7907 for the two-independent-sample design. If $a_u \geq 0$ (which occurs when $\bar{y}_2$ is too close to zero), then the unpooled Satterthwaite-based two-sided $100(1-\alpha)\%$ confidence interval for μ_r does not exist. If $a_u < 0$, then the interval is

$$\left(-\frac{b_u}{a_u} + \frac{(b_u^2 - a_u c_u)^{\frac{1}{2}}}{a_u} \ , \ -\frac{b_u}{a_u} - \frac{(b_u^2 - a_u c_u)^{\frac{1}{2}}}{a_u} \right)$$

The t test assuming unequal variances is the test derived in Tamhane and Logan (2004). The hypothesis $H_0\colon \mu_r = \mu_0$ is rewritten as $H_0\colon \mu_1 - \mu_0\mu_2 = 0$, and the Satterthwaite t test in the

section "Normal Difference (DIST=NORMAL TEST=DIFF)" on page 7907 for the two-independent-sample design is conducted on the original y_{1i} values ($i \in \{1, \ldots, n_1^\star\}$) and transformed values of y_{2i}

$$y_{2i}^\star = \mu_0 y_{2i} \ , \quad i \in \{1, \ldots, n_2^\star\}$$

with a null difference of 0. The degrees of freedom used in the unpooled t test differs from the df_u used in the unpooled confidence interval. The mean ratio μ_0 under the null hypothesis is used in place of the estimate $\hat{\mu}_r$:

$$\mathrm{df}_u^\star = \frac{\left(\dfrac{s_1^2}{\sum_{i=1}^{n_1^\star} f_{1i} w_{1i}} + \dfrac{\mu_0^2 s_2^2}{\sum_{i=1}^{n_2^\star} f_{2i} w_{2i}} \right)^2}{\dfrac{s_1^4}{(n_1-1)\left(\sum_{i=1}^{n_1^\star} f_{1i} w_{1i}\right)^2} + \dfrac{\mu_0^4 s_2^4}{(n_2-1)\left(\sum_{i=1}^{n_2^\star} f_{2i} w_{2i}\right)^2}}$$

The t value for the Satterthwaite-based unpooled test is computed as

$$t_u = \frac{\bar{y}_1 - \mu_0 \bar{y}_2}{\left(\dfrac{s_1^2}{\sum_{i=1}^{n_1^\star} f_{1i} w_{1i}} + \dfrac{\mu_0^2 s_2^2}{\sum_{i=1}^{n_2^\star} f_{2i} w_{2i}} \right)^{\frac{1}{2}}}$$

The p-value of the test is computed as

$$p\text{-value} = \begin{cases} P\left(t_u^2 > F_{1-\alpha, 1, \mathrm{df}_u^\star} \right) \ , & \text{2-sided} \\[2mm] P\left(t_u < t_{\alpha, \mathrm{df}_u^\star} \right) \ , & \text{lower 1-sided} \\[2mm] P\left(t_u > t_{1-\alpha, \mathrm{df}_u^\star} \right) \ , & \text{upper 1-sided} \end{cases}$$

AB/BA Crossover Design

Let "A" and "B" denote the two treatment values. Define the following notation:

$$n_1^\star = \text{number of observations with treatment sequence AB}$$
$$n_2^\star = \text{number of observations with treatment sequence BA}$$
$$y_{11i} = \text{response value of } i\text{th observation in sequence AB during period 1, } i \in \{1, \ldots, n_1^\star\}$$
$$y_{12i} = \text{response value of } i\text{th observation in sequence AB during period 2, } i \in \{1, \ldots, n_1^\star\}$$
$$y_{21i} = \text{response value of } i\text{th observation in sequence BA during period 1, } i \in \{1, \ldots, n_2^\star\}$$
$$y_{22i} = \text{response value of } i\text{th observation in sequence BA during period 2, } i \in \{1, \ldots, n_2^\star\}$$

So $\{y_{11i}, \ldots, y_{11n_1^\star}\}$ and $\{y_{22i}, \ldots, y_{22n_2^\star}\}$ are all observed at treatment level A, and $\{y_{12i}, \ldots, y_{12n_2^\star}\}$ and $\{y_{21i}, \ldots, y_{21n_1^\star}\}$ are all observed at treatment level B.

Define the *period difference* for an observation as the difference between period 1 and period 2 response values:

$$\mathrm{pd}_{kji} = y_{k1i} - y_{k2i}$$

for $k \in \{1, 2\}$ and $i \in \{1, \ldots, n_k^\star\}$. Similarly, the *period ratio* is the ratio between period 1 and period 2 response values:

$$\text{pr}_{kji} = y_{k1i}/y_{k2i}$$

The *crossover difference* for an observation is the difference between treatment A and treatment B response values:

$$\text{cd}_{kji} = \begin{cases} y_{k1i} - y_{k2i} \; , & k = 1 \\ y_{k2i} - y_{k1i} \; , & k = 2 \end{cases}$$

Similarly, the *crossover ratio* is the ratio between treatment A and treatment B response values:

$$\text{cr}_{kji} = \begin{cases} y_{k1i}/y_{k2i} \; , & k = 1 \\ y_{k2i}/y_{k1i} \; , & k = 2 \end{cases}$$

In the absence of the IGNOREPERIOD option in the PROC TTEST statement, the data are split into two groups according to treatment sequence and analyzed as a two-independent-sample design. If DIST=NORMAL, then the analysis of the treatment effect is based on the half period differences $\{\text{pd}_{kji}/2\}$, and the analysis for the period effect is based on the half crossover differences $\{\text{cd}_{kji}/2\}$. The computations for the normal difference analysis are the same as in the section "Normal Difference (DIST=NORMAL TEST=DIFF)" on page 7907 for the two-independent-sample design. The normal ratio analysis without the IGNOREPERIOD option is not supported for the AB/BA crossover design. If DIST=LOGNORMAL, then the analysis of the treatment effect is based on the square root of the period ratios $\{\sqrt{\text{pr}_{kji}}\}$, and the analysis for the period effect is based on the square root of the crossover ratios $\{\sqrt{\text{cr}_{kji}}\}$. The computations are the same as in the section "Lognormal Ratio (DIST=LOGNORMAL TEST=RATIO)" on page 7909 for the two-independent-sample design.

If the IGNOREPERIOD option is specified, then the treatment effect is analyzed as a paired analysis on the (treatment A, treatment B) response value pairs, regardless of treatment sequence. So the set of pairs is taken to be the concatenation of $\{(y_{111}, y_{121}), \ldots, (y_{11n_1^\star}, y_{12n_1^\star})\}$ and $\{(y_{221}, y_{211}), \ldots, (y_{22n_2^\star}, y_{22n_2^\star})\}$. The computations are the same as in the section "Paired Design" on page 7906.

See Senn (2002, Chapter 3) for a more detailed discussion of the AB/BA crossover design.

TOST Equivalence Test

The hypotheses for an equivalence test are

$$H_0 : \mu < \theta_L \quad \text{or} \quad \mu > \theta_U$$
$$H_1 : \theta_L \leq \mu \leq \theta_U$$

where θ_L and θ_U are the lower and upper bounds specified in the TOST option in the PROC TTEST statement, and μ is the analysis criterion (mean, mean ratio, or mean difference, depending on the analysis). Following the two one-sided tests (TOST) procedure of Schuirmann (1987), the equivalence test is conducted by performing two separate tests:

$$H_{a0} : \mu < \theta_L$$
$$H_{a1} : \mu \geq \theta_L$$

and

$$H_{b0}:\mu > \theta_U$$
$$H_{b1}:\mu \leq \theta_U$$

The overall p-value is the larger of the two p-values of those tests.

Rejection of H_0 in favor of H_1 at significance level α occurs if and only if the $100(1 - 2\alpha)\%$ confidence interval for μ is contained completely within (θ_L, θ_U). So, the $100(1 - 2\alpha)\%$ confidence interval for μ is displayed in addition to the usual $100(1 - \alpha)\%$ interval.

See Phillips (1990), Diletti, Hauschke, and Steinijans (1991), and Hauschke et al. (1999) for further discussion of equivalence testing for the designs supported in the TTEST procedure.

Displayed Output

For an AB/BA crossover design, the "CrossoverVarInfo" table shows the variables specified for the response and treatment values in each period of the design.

The summary statistics in the "Statistics" table and confidence limits in the "ConfLimits" table are displayed for certain variables and/or transformations or subgroups of the variables in the analysis, depending on the design. For a one-sample design, summary statistics are displayed for all variables in the analysis. For a paired design, statistics are displayed for the difference if you specify the TEST=DIFF option in the PROC TTEST statement, or for the ratio if you specify TEST=RATIO. For a two-independent-sample design, the statistics for each of the two groups and for the difference (if TEST=DIFF) or ratio (if TEST=RATIO) are displayed. For an AB/BA crossover design, statistics are displayed for each of the four cells in the design (all four combinations of the two periods and two treatments). If the IGNOREPERIOD option is absent, then if TEST=DIFF is specified, statistics are displayed for the treatment difference within each sequence and overall, and also for the period difference. If TEST=RATIO, statistics are displayed for the treatment ratio within each sequence and overall, and also for the period ratio. If the IGNOREPERIOD option is specified in the VAR statement, then statistics are displayed for the overall treatment difference if TEST=DIFF or for the overall treatment ratio if TEST=RATIO.

The "Statistics" table displays the following summary statistics:

- the name of the variable(s), displayed if the NOBYVAR option is used in the PROC TTEST statement

- the name of the classification variable (if two-independent-sample design) or treatment and period (if AB/BA crossover design)

- N, the number of nonmissing values

- the (arithmetic) Mean, displayed if the DIST=NORMAL option is specified in the PROC TTEST statement

- the Geometric Mean, displayed if the DIST=LOGNORMAL option is specified in the PROC TTEST statement

- Std Dev, the standard deviation, displayed if the DIST=NORMAL option is specified in the PROC TTEST statement

- the Coefficient of Variation, displayed if the DIST=LOGNORMAL option is specified in the PROC TTEST statement

- Std Err, the standard error of the mean, displayed if the DIST=NORMAL option is specified in the PROC TTEST statement

- the Minimum value

- the Maximum value

The "ConfLimits" table displays the following:

- the name of the variable(s), displayed if the NOBYVAR option is used in the PROC TTEST statement

- the name of the classification variable (if two-independent-sample design) or treatment and period (if AB/BA crossover design)

- the (arithmetic) Mean, displayed if the DIST=NORMAL option is specified in the PROC TTEST statement

- the Geometric Mean, displayed if the DIST=LOGNORMAL option is specified in the PROC TTEST statement

- $100(1 - \alpha)\%$ CL Mean, the lower and upper confidence limits for the mean. Separate pooled and Satterthwaite confidence limits are shown for the difference or ratio transformations in two-independent-sample designs and AB/BA crossover designs without the IGNOREPERIOD option.

- Std Dev, the standard deviation, displayed if the DIST=NORMAL option is specified in the PROC TTEST statement

- the Coefficient of Variation, displayed if the DIST=LOGNORMAL option is specified in the PROC TTEST statement

- $100(1 - \alpha)\%$ CL Std Dev, the equal-tailed confidence limits for the standard deviation, displayed if the DIST=NORMAL and CI=EQUAL options are specified in the PROC TTEST statement

- $100(1 - \alpha)\%$ UMPU CL Std Dev, the UMPU confidence limits for the standard deviation, displayed if the DIST=NORMAL and CI=UMPU options are specified in the PROC TTEST statement

- $100(1-\alpha)\%$ CL CV, the equal-tailed confidence limits for the coefficient of variation, displayed if the DIST=LOGNORMAL and CI=EQUAL options are specified in the PROC TTEST statement

- $100(1 - \alpha)\%$ UMPU CL CV, the UMPU confidence limits for the coefficient of variation, displayed if the DIST=LOGNORMAL and CI=UMPU options are specified in the PROC TTEST statement

The confidence limits in the "EquivLimits" table and test results in the "TTests" and "EquivTests" tables are displayed only for the test criteria—that is, the variables or transformations being tested. For a one-sample design, results are displayed for all variables in the analysis. For a paired design, results are displayed for the difference if you specify the TEST=DIFF option in the PROC TTEST statement, or for the ratio if you specify TEST=RATIO. For a two-independent-sample design, the results for the difference (if TEST=DIFF) or ratio (if TEST=RATIO) are displayed. For an AB/BA crossover design, results are displayed for the treatment difference (if TEST=DIFF) or ratio (if TEST=RATIO). If the IGNOREPERIOD option is absent, then results are also displayed for the period difference (if TEST=DIFF) or ratio (if TEST=RATIO).

The "EquivLimits" table, produced only if the TOST option is specified in the PROC TTEST statement, displays the following:

- the name of the variable(s), displayed if the NOBYVAR option is used in the PROC TTEST statement

- the (arithmetic) Mean, displayed if the DIST=NORMAL option is specified in the PROC TTEST statement

- the Geometric Mean, displayed if the DIST=LOGNORMAL option is specified in the PROC TTEST statement

- Lower Bound, the lower equivalence bound for the mean specified in the TOST option in the PROC TTEST statement

- $100(1 - 2\alpha)\%$ CL Mean, the lower and upper confidence limits for the mean relevant to the equivalence test. Separate pooled and Satterthwaite confidence limits are shown for two-independent-sample designs and AB/BA crossover designs without the IGNOREPERIOD option.

- Upper Bound, the upper equivalence bound for the mean specified in the TOST option in the PROC TTEST statement

- Assessment, the result of the equivalence test at the significance level specified by the ALPHA= option in the PROC TTEST statement, either "Equivalent" or "Not equivalent"

The "TTests" table is produced only if the TOST option is *not* specified in the PROC TTEST statement. Separate results for pooled and Satterthwaite tests (and also the Cochran and Cox test, if the COCHRAN option is specified in the PROC TTEST statement) are displayed for two-independent-sample designs and AB/BA crossover designs without the IGNOREPERIOD option. The table includes the following results:

- the name of the variable(s), displayed if the NOBYVAR option is used in the PROC TTEST statement

- t Value, the *t* statistic for comparing the mean to the null value as specified by the H0= option in the PROC TTEST statement

- DF, the degrees of freedom

- the p-value, the probability of obtaining a t statistic at least as extreme as the observed t value under the null hypothesis

The "EquivTests" table is produced only if the TOST option is specified in the PROC TTEST statement. Separate results for pooled and Satterthwaite tests are displayed for two-independent-sample designs and AB/BA crossover designs without the IGNOREPERIOD option. Each test consists of two separate one-sided tests. The overall p-value is the larger p-value from these two tests. The table includes the following results:

- the name of the variable(s), displayed if the NOBYVAR option is used in the PROC TTEST statement

- Null, the lower equivalence bound for the Upper test or the upper equivalence bound for the Lower test, as specified by the TOST option in the PROC TTEST statement

- t Value, the t statistic for comparing the mean to the Null value

- DF, the degrees of freedom

- the p-value, the probability of obtaining a t statistic at least as extreme as the observed t value under the null hypothesis

The "Equality" table gives the results of the test of equality of variances. It is displayed for two-independent-sample designs and AB/BA crossover designs without the IGNOREPERIOD option. The table includes the following results:

- the name of the variable(s), displayed if the NOBYVAR option is used in the PROC TTEST statement

- Num DF and Den DF, the numerator and denominator degrees of freedom

- F Value, the F' (folded) statistic

- Pr > F, the probability of a greater F' value. This is the two-tailed p-value.

ODS Table Names

PROC TTEST assigns a name to each table it creates. You can use these names to reference the table when using the Output Delivery System (ODS) to select tables and create output data sets. These names are listed in Table 93.5. For more information about ODS, see Chapter 20, "Using the Output Delivery System."

Table 93.5 ODS Tables Produced by PROC TTEST

ODS Table Name	Description	Syntax
ConfLimits	$100(1-\alpha)\%$ confidence limits for means, standard deviations, and/or coefficients of variation	By default
Equality	Tests for equality of variance	CLASS statement or VAR / CROSSOVER
EquivLimits	$100(1-2\alpha)\%$ confidence limits for means	PROC TTEST TOST
EquivTests	Equivalence t tests	PROC TTEST TOST
Statistics	Univariate summary statistics	By default
TTests	t tests	By default

ODS Graphics

ODS Graph Names

To request graphics with PROC TTEST, you must first enable ODS Graphics by specifying the `ODS GRAPHICS ON` statement. See Chapter 21, "Statistical Graphics Using ODS," for more information. You can reference every graph produced through ODS Graphics with a name. The names of the graphs that PROC TTEST generates are listed in Table 93.6, along with the required statements and options.

Table 93.6 ODS Graphics Produced by PROC TTEST

ODS Graph Name	Plot Description	Option
AgreementOfPeriods	Plot of period 2 against period 1 response values for an AB/BA crossover design	VAR / CROSSOVER PLOTS=AGREEMENT(TYPE=PERIOD)
AgreementOfTreatments	Plot of second treatment against first treatment response values for an AB/BA crossover design	VAR / CROSSOVER PLOTS=AGREEMENT
AgreementPlot	Plot of second response against first response for a paired design	PAIRED statement PLOTS=AGREEMENT
BoxPlot	Box plots, also with confidence band for one-sample or paired design	PLOTS=BOX PLOTS=SUMMARY(UNPACK)
Histogram	Histograms with overlaid kernel densities, and also normal densities if DIST=NORMAL	PLOTS=HISTOGRAM PLOTS=SUMMARY(UNPACK)
Interval	Confidence intervals for means	PLOTS=INTERVAL

Table 93.6   *continued*

ODS Graph Name	Plot Description	Option
ProfilesOverPrd	Plot of response profiles over periods 1 and 2 for an AB/BA crossover design	VAR / CROSSOVER PLOTS=PROFILES(TYPE=PERIOD)
ProfilesOverTrt	Plot of response profiles over first and second treatments for an AB/BA crossover design	VAR / CROSSOVER PLOTS=PROFILES
ProfilesPlot	Plot of response profiles over first and second response values for a paired design	PAIRED statement PLOTS=PROFILES
QQPlot	Normal quantile-quantile plots	PLOTS=QQ
SummaryPanel	Histograms with overlaid kernel densities (and also normal densities if DIST=NORMAL) and box plots (and also with confidence band for one-sample or paired design)	PLOTS=SUMMARY

Interpreting Graphs

Agreement Plots for Paired Designs

For paired designs, the second response of each pair is plotted against the first response, with the mean shown as a large bold symbol. If the WEIGHT statement is used, then the mean is the weighted mean. A diagonal line with slope=0 and y-intercept=1 is overlaid. The location of the points with respect to the diagonal line reveals the strength and direction of the difference or ratio. The tighter the clustering along the same direction as the line, the stronger the positive correlation of the two measurements for each subject. Clustering along a direction perpendicular to the line indicates negative correlation.

Period Agreement Plots for Crossover Designs

The response in the second period is plotted against the response in the first period, with plot symbols distinguishing the two treatment sequences and the two sequence means shown larger in bold. If the WEIGHT statement is used, then the means are weighted means. A diagonal line with slope=0 and y-intercept=1 is overlaid.

In the absence of a strong period effect, the points from each sequence will appear as mirror images about the diagonal line, farther apart with stronger treatment effects. Deviations from symmetry about the diagonal line indicate a period effect. The spread of points within each treatment sequence is an indicator of between-subject variability. The tighter the clustering along the same direction as the line (within each treatment sequence), the stronger the positive correlation of the two measurements for each subject. Clustering along a direction perpendicular to the line indicates negative correlation.

The period agreement plot is usually less informative than the treatment agreement plot. The exception is when the period effect is stronger than the treatment effect.

Treatment Agreement Plots for Crossover Designs

The response associated with the second treatment is plotted against the response associated with the first treatment, with plot symbols distinguishing the two treatment sequences and the two sequence means shown larger in bold. If the WEIGHT statement is used, then the means are weighted means. A diagonal line with slope=0 and y-intercept=1 is overlaid.

The location of the points with respect to the diagonal line reveals the strength and direction of the treatment effect. Substantial location differences between the two sequences indicates a strong period effect. The spread of points within each treatment sequence is an indicator of between-subject variability. The tighter the clustering along the same direction as the line (within each treatment sequence), the stronger the positive correlation of the two measurements for each subject. Clustering along a direction perpendicular to the line indicates negative correlation.

Box Plots

The box is drawn from the 25th percentile (lower quartile) to the 75th percentile (upper quartile). The vertical line inside the box shows the location of the median. If DIST=NORMAL, then a diamond symbol shows the location of the mean. The whiskers extend to the minimum and maximum observations, and circles beyond the whiskers identify outliers.

For one-sample and paired designs, a confidence interval for the mean is shown as a band in the background. If the analysis is an equivalence analysis (with the TOST option in the PROC TTEST statement), then the interval is a $100(1 - 2\alpha)\%$ confidence interval shown along with the equivalence bounds. The inclusion of this interval completely within the bounds is indicative of a significant p-value. If the analysis is not an equivalence analysis, then the confidence level is $100(1 - \alpha)\%$. If the SHOWH0 global plot option is used, then the null value for the test is shown. If the WEIGHT statement is used, then weights are incorporated in the confidence intervals.

Histograms

The WEIGHT statement is ignored in the computation of the normal and kernel densities.

Confidence Intervals

If the analysis is an equivalence analysis (with the TOST option in the PROC TTEST statement), then unless the TYPE=PERGROUP option is used, the interval is a $100(1 - 2\alpha)\%$ mean confidence interval shown along with the equivalence bounds. The inclusion of this interval completely within the bounds is indicative of a significant p-value.

If the analysis is not an equivalence analysis, or if the TYPE=PERGROUP option is used, then the confidence level is $100(1 - \alpha)\%$. If the SHOWH0 global plot option is used, then the null value for the test is shown.

If the SIDES=L or SIDES=U option is used in the PROC TTEST statement, then the unbounded side of the one-sided interval is represented with an arrowhead. Note that the actual location of the arrowhead is irrelevant.

If the WEIGHT statement is used, then weights are incorporated in the confidence intervals.

Profiles for Paired Designs

For paired designs, a line is drawn for each observation from left to right connecting the first response to the second response. The mean first response and mean second response are connected with a bold line. If the WEIGHT statement is used, then the means are weighted means. The more extreme the slope, the stronger the effect. A wide spread of profiles indicates high between-subject variability. Consistent positive slopes indicate strong positive correlation. Widely varying slopes indicate lack of correlation, while consistent negative slopes indicate strong negative correlation.

Profiles over Period for Crossover Designs

For each observation, the response for the first period is connected to the response for second period, regardless of the treatment applied in each period. The means for each treatment sequence are shown in bold. If the WEIGHT statement is used, then the means are weighted means.

In the absence of a strong period effect, the profiles for each sequence will appear as mirror images about an imaginary horizontal line in the center. Deviations from symmetry about this imaginary horizontal line indicate a period effect. A wide spread of profiles within sequence indicates high between-subject variability.

The TYPE=PERIOD plot is usually less informative than the TYPE=TREATMENT plot. The exception is when the period effect is stronger than the treatment effect.

Profiles over Treatment for Crossover Designs

For each observation, the response for the first treatment is connected to the response for the second treatment, regardless of the periods in which they occur. The means for each treatment sequence are shown in bold. If the WEIGHT statement is used, then the means are weighted means.

In general, the more extreme the slope, the stronger the treatment effect. Slope differences between the two treatment sequences measure the period effect. A wide spread of profiles within sequence indicates high between-subject variability.

Q-Q Plots

Q-Q plots are useful for diagnosing violations of the normality and homoscedasticity assumptions. If the data in a Q-Q plot come from a normal distribution, the points will cluster tightly around the reference line. You can use the UNIVARIATE procedure with the NORMAL option to numerically check the normality assumption.

Examples: TTEST Procedure

Example 93.1: Using Summary Statistics to Compare Group Means

This example, taken from Huntsberger and Billingsley (1989), compares two grazing methods using 32 steers. Half of the steers are allowed to graze continuously while the other half are subjected to controlled grazing time. The researchers want to know if these two grazing methods affect weight gain differently. The data are read by the following DATA step:

```
data graze;
   length GrazeType $ 10;
   input GrazeType $ WtGain @@;
   datalines;
controlled  45    controlled  62
controlled  96    controlled 128
controlled 120    controlled  99
controlled  28    controlled  50
controlled 109    controlled 115
controlled  39    controlled  96
controlled  87    controlled 100
controlled  76    controlled  80
continuous  94    continuous  12
continuous  26    continuous  89
continuous  88    continuous  96
continuous  85    continuous 130
continuous  75    continuous  54
continuous 112    continuous  69
continuous 104    continuous  95
continuous  53    continuous  21
;
run;
```

The variable GrazeType denotes the grazing method: "controlled" is controlled grazing and "continuous" is continuous grazing. The dollar sign ($) following GrazeType makes it a character variable, and the trailing at signs (@@) tell the procedure that there is more than one observation per line.

If you have summary data—that is, just means and standard deviations, as computed by PROC MEANS—then you can still use PROC TTEST to perform a simple *t* test analysis. This example demonstrates this mode of input for PROC TTEST. Note, however, that graphics are unavailable when summary statistics are used as input.

The MEANS procedure is invoked to create a data set of summary statistics with the following statements:

```
proc sort;
   by GrazeType;
proc means data=graze noprint;
   var WtGain;
```

```
    by GrazeType;
    output out=newgraze;
run;
```

The NOPRINT option eliminates all printed output from the MEANS procedure. The VAR statement tells PROC MEANS to compute summary statistics for the WtGain variable, and the BY statement requests a separate set of summary statistics for each level of GrazeType. The OUTPUT OUT= statement tells PROC MEANS to put the summary statistics into a data set called newgraze so that it can be used in subsequent procedures. This new data set is displayed in Output 93.1.1 by using PROC PRINT as follows:

```
proc print data=newgraze;
run;
```

The _STAT_ variable contains the names of the statistics, and the GrazeType variable indicates which group the statistic is from.

Output 93.1.1 Output Data Set of Summary Statistics

Obs	GrazeType	_TYPE_	_FREQ_	_STAT_	WtGain
1	continuous	0	16	N	16.000
2	continuous	0	16	MIN	12.000
3	continuous	0	16	MAX	130.000
4	continuous	0	16	MEAN	75.188
5	continuous	0	16	STD	33.812
6	controlled	0	16	N	16.000
7	controlled	0	16	MIN	28.000
8	controlled	0	16	MAX	128.000
9	controlled	0	16	MEAN	83.125
10	controlled	0	16	STD	30.535

The following statements invoke PROC TTEST with the newgraze data set, as denoted by the DATA= option:

```
proc ttest data=newgraze;
    class GrazeType;
    var WtGain;
run;
```

The CLASS statement contains the variable that distinguishes between the groups being compared, in this case GrazeType. The summary statistics and confidence intervals are displayed first, as shown in Output 93.1.2.

Output 93.1.2 Summary Statistics and Confidence Limits

```
                          The TTEST Procedure

                          Variable:  WtGain
```

GrazeType	N	Mean	Std Dev	Std Err	Minimum	Maximum
continuous	16	75.1875	33.8117	8.4529	12.0000	130.0
controlled	16	83.1250	30.5350	7.6337	28.0000	128.0
Diff (1–2)		−7.9375	32.2150	11.3897		

GrazeType	Method	Mean	95% CL Mean		Std Dev
continuous		75.1875	57.1705	93.2045	33.8117
controlled		83.1250	66.8541	99.3959	30.5350
Diff (1–2)	Pooled	−7.9375	−31.1984	15.3234	32.2150
Diff (1–2)	Satterthwaite	−7.9375	−31.2085	15.3335	

GrazeType	Method	95% CL Std Dev	
continuous		.	.
controlled		.	.
Diff (1–2)	Pooled	25.7434	43.0609
Diff (1–2)	Satterthwaite		

In Output 93.1.2, The GrazeType column specifies the group for which the statistics are computed. For each class, the sample size, mean, standard deviation and standard error, and maximum and minimum values are displayed. The confidence bounds for the mean are also displayed; however, since summary statistics are used as input, the confidence bounds for the standard deviation of the groups are not calculated.

Output 93.1.3 shows the results of tests for equal group means and equal variances.

Output 93.1.3 *t* Tests

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	30	−0.70	0.4912
Satterthwaite	Unequal	29.694	−0.70	0.4913

```
                     Equality of Variances
```

Method	Num DF	Den DF	F Value	Pr > F
Folded F	15	15	1.23	0.6981

A group test statistic for the equality of means is reported for both equal and unequal variances. Both tests indicate a lack of evidence for a significant difference between grazing methods ($t = -0.70$ and $p = 0.4912$ for the pooled test, $t = -0.70$ and $p = 0.4913$ for the Satterthwaite test). The equality of variances test does not indicate a significant difference in the two variances ($F' = 1.23$, $p = 0.6981$). Note that this test assumes that the observations in both data sets are normally distributed;

this assumption can be checked in PROC UNIVARIATE by using the NORMAL option with the raw data.

Although the ability to use summary statistics as input is useful if you lack access to the original data, some of the output that would otherwise be produced in an analysis on the original data is unavailable. There are also limitations on the designs and distributional assumptions that can be used with summary statistics as input. For more information, see the section "Input Data Set of Statistics" on page 7901.

Example 93.2: One-Sample Comparison with the FREQ Statement

This example examines children's reading skills. The data consist of Degree of Reading Power (DRP) test scores from 44 third-grade children and are taken from Moore (1995, p. 337). Their scores are given in the following DATA step:

```
data read;
   input score count @@;
   datalines;
40 2   47 2   52 2   26 1   19 2
25 2   35 4   39 1   26 1   48 1
14 2   22 1   42 1   34 2   33 2
18 1   15 1   29 1   41 2   44 1
51 1   43 1   27 2   46 2   28 1
49 1   31 1   28 1   54 1   45 1
;
run;
```

The following statements invoke the TTEST procedure to test if the mean test score is equal to 30.

```
ods graphics on;

proc ttest data=read h0=30;
   var score;
   freq count;
run;

ods graphics off;
```

The ODS GRAPHICS statement requests graphical output. The count variable contains the frequency of occurrence of each test score; this is specified in the FREQ statement. The output, shown in Output 93.2.1, contains the results.

Output 93.2.1 TTEST Results

```
                      The TTEST Procedure

                       Variable:  score

     N         Mean      Std Dev      Std Err      Minimum      Maximum

    44       34.8636     11.2303       1.6930      14.0000      54.0000


        Mean          95% CL Mean            Std Dev       95% CL Std Dev

      34.8636      31.4493  38.2780        11.2303       9.2788  14.2291


                       DF     t Value     Pr > |t|

                       43       2.87        0.0063
```

The SAS log states that 30 observations and two variables have been read. However, the sample size given in the TTEST output is N=44. This is due to specifying the count variable in the FREQ statement. The test is significant ($t = 2.87$, $p = 0.0063$) at the 5% level, so you can conclude that the mean test score is different from 30.

The summary panel in Output 93.2.2 shows a histogram with overlaid normal and kernel densities, a box plot, and the 95% confidence interval for the mean.

Output 93.2.2 Summary Panel
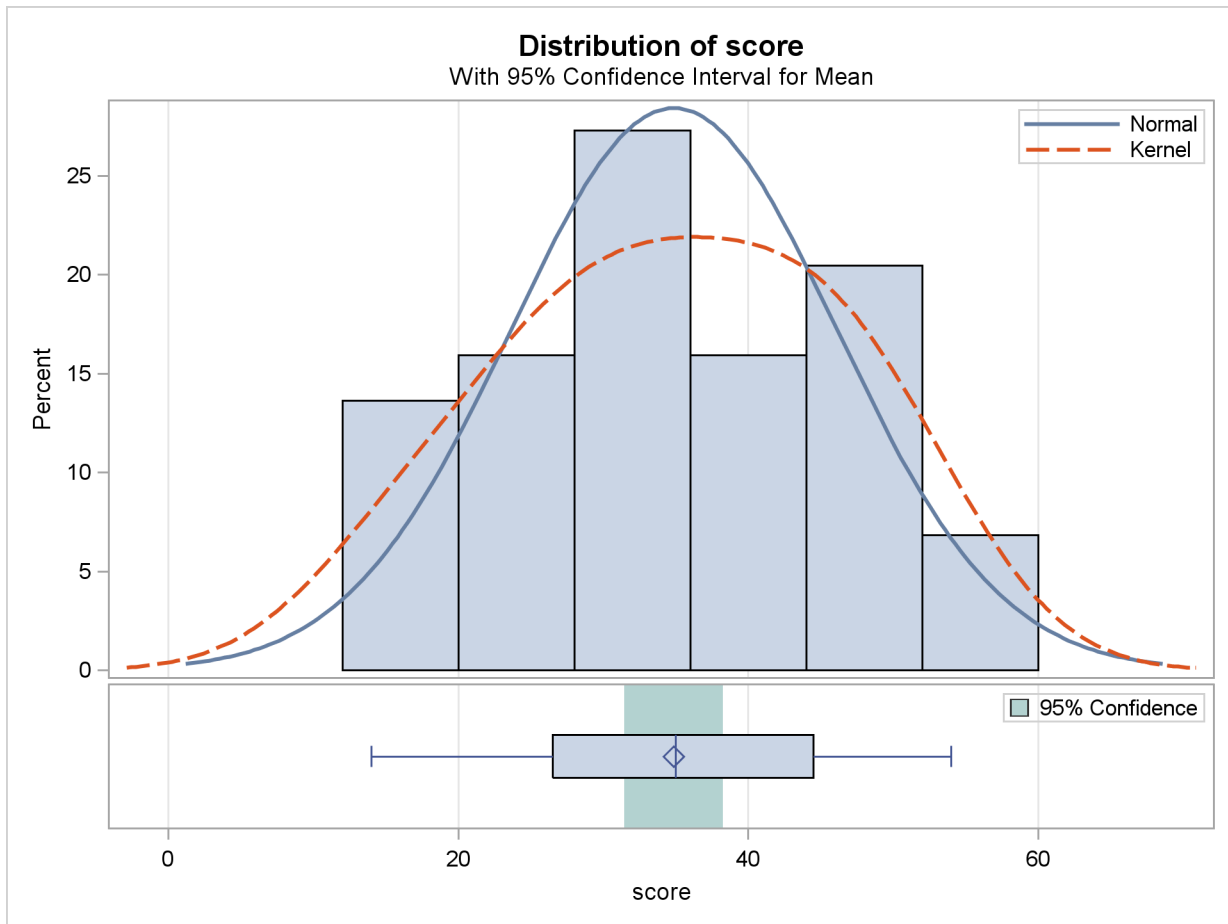
The Q-Q plot in Output 93.2.3 assesses the normality assumption.

Output 93.2.3 Q-Q Plot
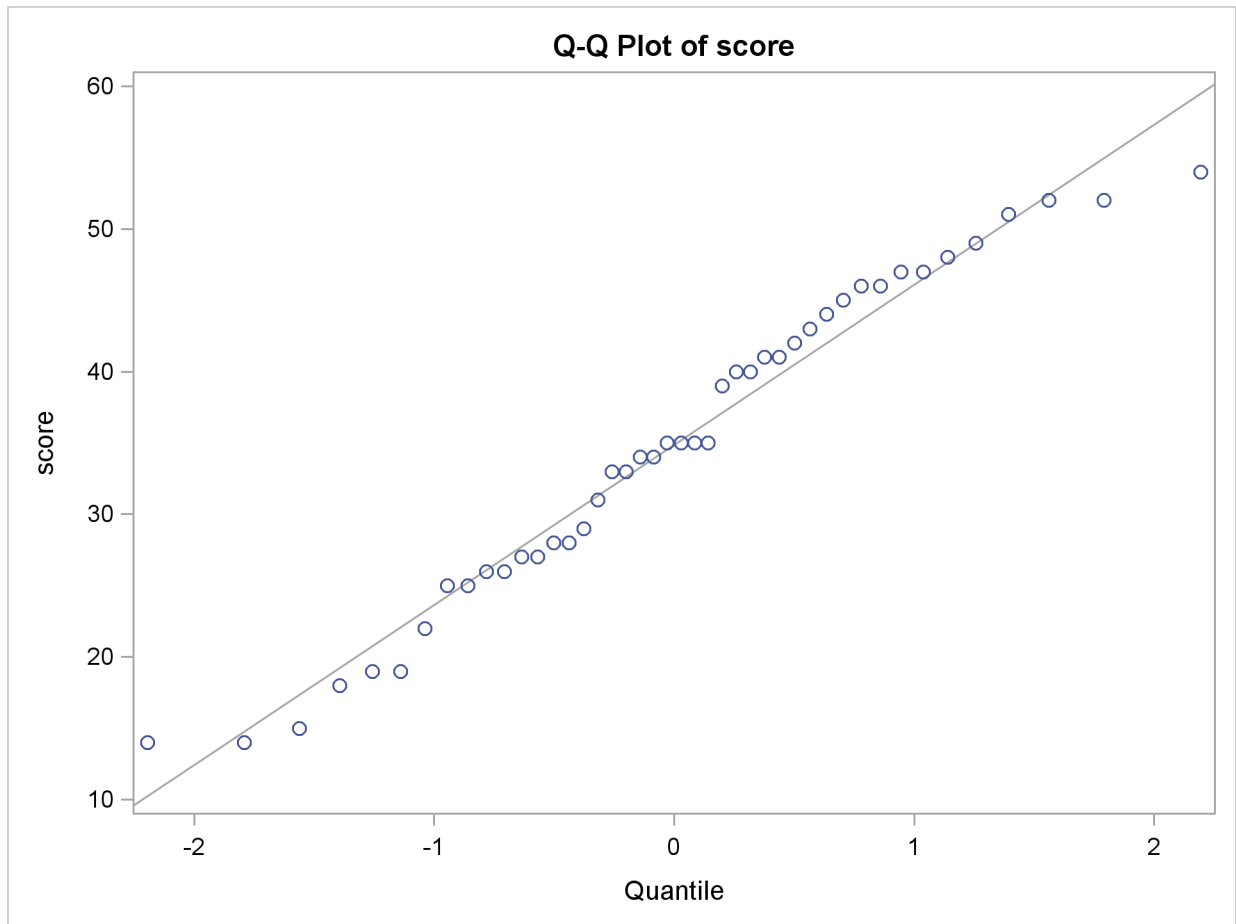
The tight clustering of the points around the diagonal line is consistent with the normality assumption. You could use the UNIVARIATE procedure with the NORMAL option to numerically check the normality assumption.

Example 93.3: Paired Comparisons

When it is not feasible to assume that two groups of data are independent, and a natural pairing of the data exists, it is advantageous to use an analysis that takes the correlation into account. Using this correlation results in higher power to detect existing differences between the means. The differences between paired observations are assumed to be normally distributed. Some examples of this natural pairing are as follows:

- pre- and post-test scores for a student receiving tutoring

- fuel efficiency readings of two fuel types observed on the same automobile

- sunburn scores for two sunblock lotions, one applied to the individual's right arm, one to the left arm

- political attitude scores of husbands and wives

In this example, taken from the *SUGI Supplemental Library User's Guide, Version 5 Edition*, a stimulus is being examined to determine its effect on systolic blood pressure. Twelve men participate in the study. Each man's systolic blood pressure is measured both before and after the stimulus is applied. The following statements input the data:

```
data pressure;
     input SBPbefore SBPafter @@;
     datalines;
120 128    124 131    130 131    118 127
140 132    128 125    140 141    135 137
126 118    130 132    126 129    127 135
;
run;
```

The variables SBPbefore and SBPafter denote the systolic blood pressure before and after the stimulus, respectively.

The statements to perform the test follow:

```
ods graphics on;

proc ttest;
   paired SBPbefore*SBPafter;
run;

ods graphics off;
```

The ODS GRAPHICS statement requests graphical output. The PAIRED statement is used to test whether the mean change in systolic blood pressure is significantly different from zero. The tabular output is displayed in Output 93.3.1.

Output 93.3.1 TTEST Results

```
                    The TTEST Procedure

                Difference:  SBPbefore - SBPafter

     N        Mean      Std Dev      Std Err      Minimum      Maximum

    12     -1.8333       5.8284       1.6825      -9.0000       8.0000


        Mean         95% CL Mean          Std Dev        95% CL Std Dev

     -1.8333     -5.5365   1.8698         5.8284         4.1288   9.8958


                      DF     t Value     Pr > |t|

                      11      -1.09        0.2992
```

The variables SBPbefore and SBPafter are the paired variables with a sample size of 12. The summary statistics of the difference are displayed (mean, standard deviation, and standard error) along with their confidence limits. The minimum and maximum differences are also displayed. The t test is not significant ($t = -1.09$, $p = 0.2992$), indicating that the stimuli did not significantly affect systolic blood pressure.

The summary panel in Output 93.3.2 shows a histogram, normal and kernel densities, box plot, and $100(1 - \alpha)\% = 95\%$ confidence interval of the SBPbefore $-$ SBPafter difference.

Output 93.3.2 Summary Panel

The agreement plot in Output 93.3.3 reveals that only three men have higher blood pressure before the stimulus than after.

Output 93.3.3 Agreement of Treatments

Agreement of SBPafter and SBPbefore
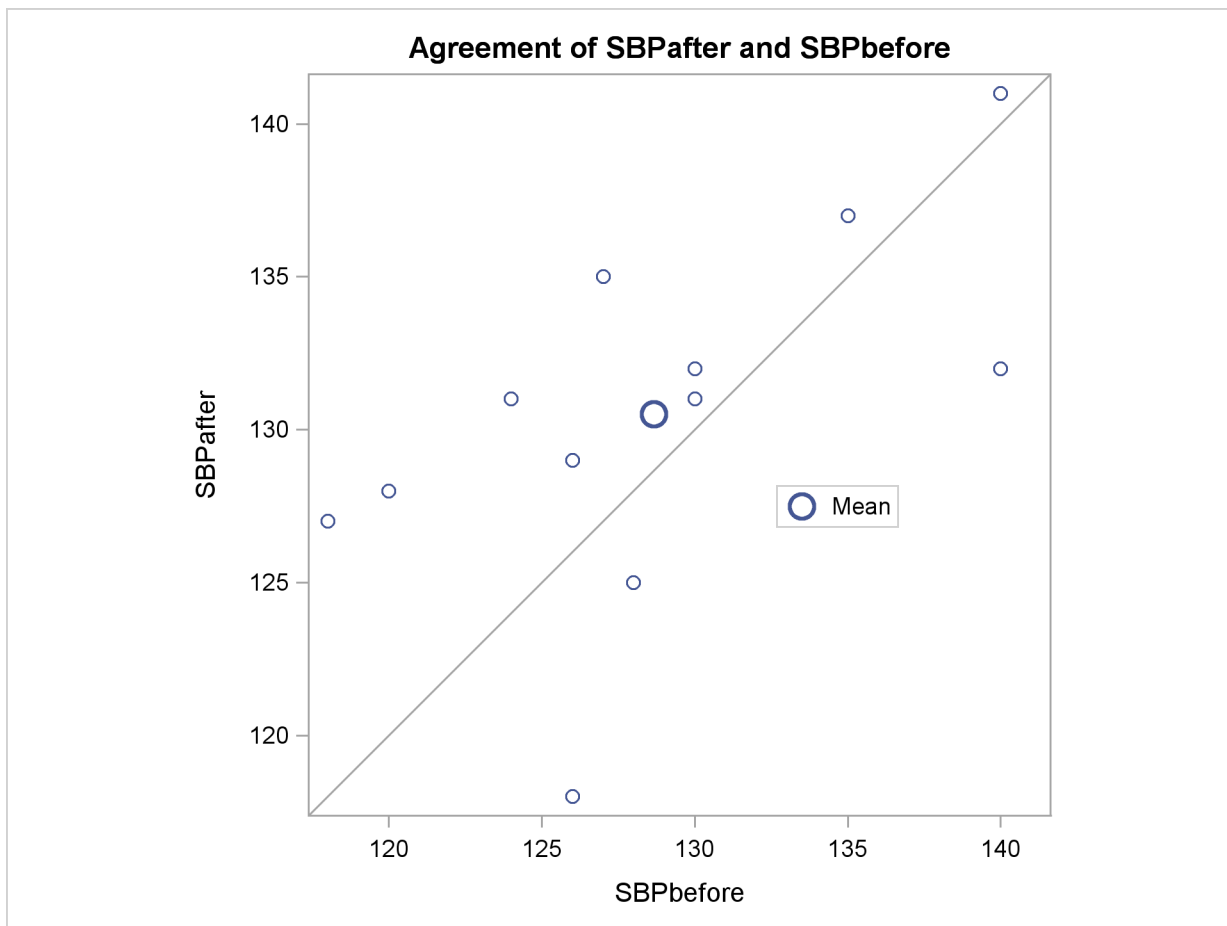
But the differences for these men are relatively large, keeping the mean difference only slightly negative.

The profiles plot in Output 93.3.4 is a different view of the same information contained in Output 93.3.3, plotting the blood pressure from before to after the stimulus.

Output 93.3.4 Profiles over Treatments
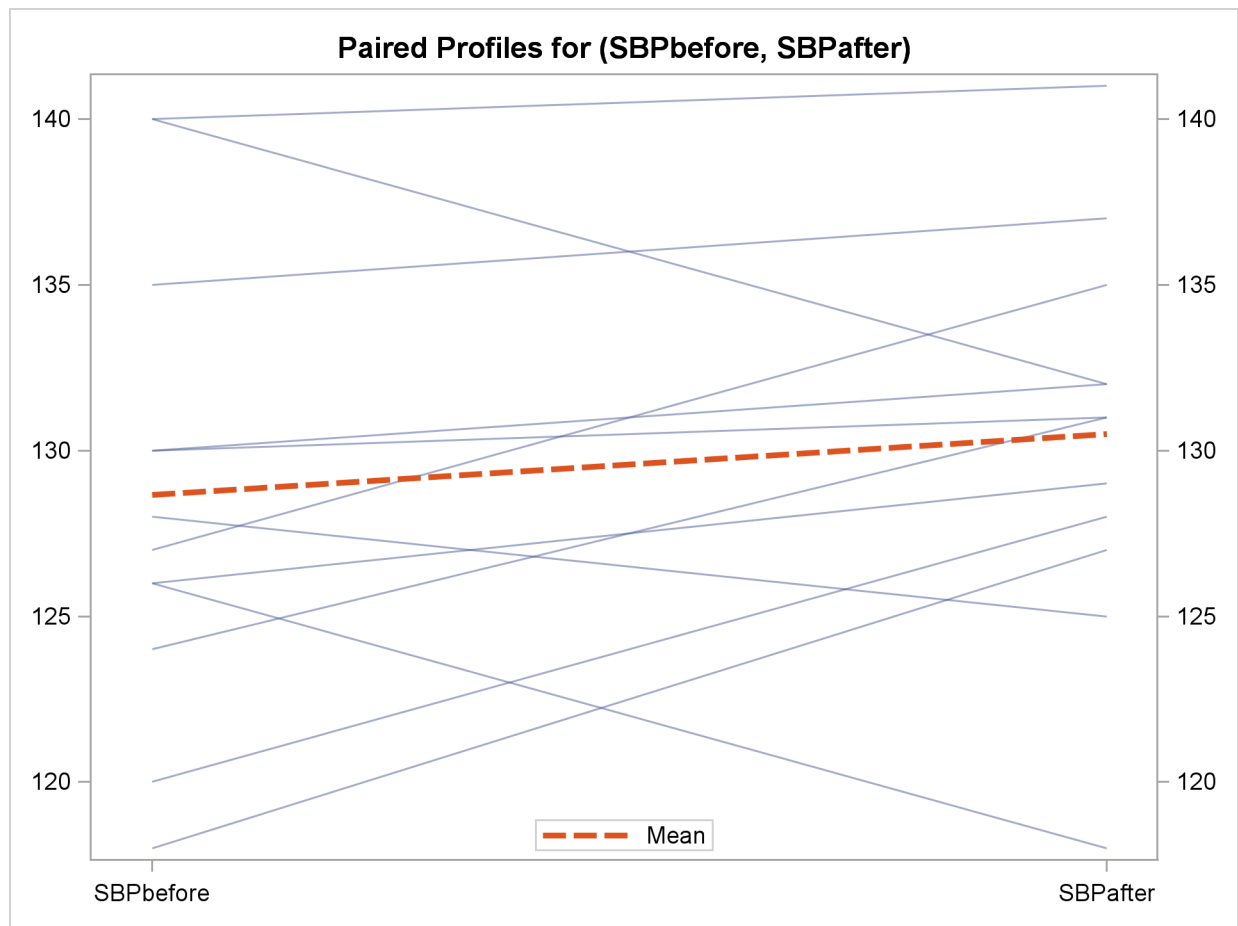
The Q-Q plot in Output 93.3.5 assesses the normality assumption.

Output 93.3.5 Q-Q Plot
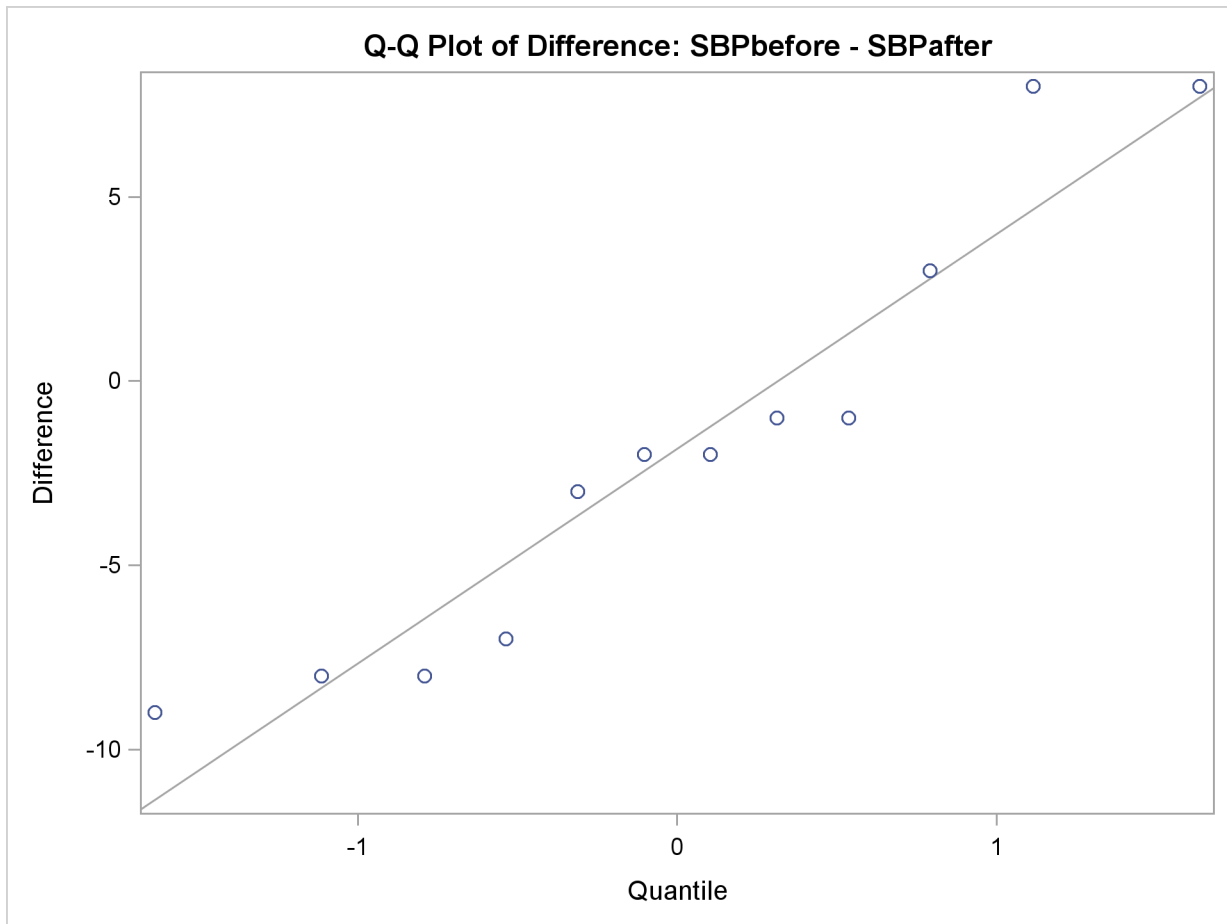
The Q-Q plot shows no obvious deviations from normality. You can check the assumption of normality more rigorously by using PROC UNIVARIATE with the NORMAL option.

Example 93.4: AB/BA Crossover Design

Senn (2002, Chapter 3) discusses a study comparing the effectiveness of two bronchodilators, formoterol ("for") and salbutamol ("sal"), in the treatment of childhood asthma. A total of 13 children are recruited for an AB/BA crossover design. A random sample of 7 of the children are assigned to the treatment sequence for/sal, receiving a dose of formoterol upon an intial visit ("period 1") and then a dose of salbutamol upon a later visit ("period 2"). The other 6 children are assigned to the sequence sal/for, receiving the treatments in the reverse order but otherwise in a similar manner. Periods 1 and 2 are sufficiently spaced so that no carryover effects are suspected. After a child inhales a dose of a bronchodilator, peak expiratory flow (PEF) is measured. Higher PEF indicates greater effectiveness. The data are assumed to be approximately normally distributed.

The data set is generated with the following statements:

```
data asthma;
   input Drug1 $ Drug2 $ PEF1 PEF2 @@;
   datalines;
for sal 310 270    for sal 310 260    for sal 370 300
for sal 410 390    for sal 250 210    for sal 380 350
for sal 330 365
sal for 370 385    sal for 310 400    sal for 380 410
sal for 290 320    sal for 260 340    sal for 90  220
;
run;
```

You can display the data by using the following statements, which produce Output 93.4.1:

```
proc print data=asthma;
run;
```

Output 93.4.1 Asthma Study Data

Obs	Drug1	Drug2	PEF1	PEF2
1	for	sal	310	270
2	for	sal	310	260
3	for	sal	370	300
4	for	sal	410	390
5	for	sal	250	210
6	for	sal	380	350
7	for	sal	330	365
8	sal	for	370	385
9	sal	for	310	400
10	sal	for	380	410
11	sal	for	290	320
12	sal	for	260	340
13	sal	for	90	220

The variables PEF1 and PEF2 represent the responses for the first and second periods, respectively. The variables Drug1 and Drug2 represent the treatment in each period.

You can analyze this crossover design by using the CROSSOVER option after a slash (/) in the VAR statement:

```
ods graphics on;

proc ttest data=asthma plots=interval;
   var PEF1 PEF2 / crossover= (Drug1 Drug2);
run;

ods graphics off;
```

The ODS GRAPHICS statement requests graphical output. With the default PROC TTEST options TEST=DIFF and DIST=NORMAL and the lack of the IGNOREPERIOD option in the VAR statement, both the treatment difference and the period difference are assessed. The PROC TTEST default options H0=0, SIDES=2, and ALPHA=0.05 specify a two-sided analysis with 95% confi-

dence limits comparing treatment and period differences to a default difference of zero. The default CI=EQUAL option in the PROC TTEST statement requests equal-tailed confidence intervals for standard deviations. The PLOTS=INTERVAL option produces TYPE=TREATMENT confidence intervals, in addition to the default plots AGREEMENT(TYPE=TREATMENT), BOX, HISTOGRAM, PROFILES(TYPE=TREATMENT), and QQ.

Output 93.4.2 summarizes the response and treatment variables for each period.

Output 93.4.2 Crossover Variable Information

```
                      The TTEST Procedure

                Response Variables:   PEF1, PEF2

                 Crossover Variable Information

             Period      Response      Treatment

                1          PEF1          Drug1
                2          PEF2          Drug2
```

Output 93.4.3 displays basic summary statistics (sample size, mean, standard deviation, standard error, minimum, and maximum) for each of the four cells in the design, the treatment difference within each treatment sequence, the overall treatment difference, and the overall period difference.

Output 93.4.3 Statistics

Sequence	Treatment	Period	N	Mean	Std Dev	Std Err
1	for	1	7	337.1	53.7631	20.3206
2	for	2	6	345.8	70.8814	28.9372
2	sal	1	6	283.3	105.4	43.0245
1	sal	2	7	306.4	64.7247	24.4636
1	Diff (1-2)		7	30.7143	32.9682	12.4608
2	Diff (1-2)		6	62.5000	44.6934	18.2460
Both	Diff (1-2)			46.6071	19.3702	10.7766
Both		Diff (1-2)		-15.8929	19.3702	10.7766

Sequence	Treatment	Period	Minimum	Maximum
1	for	1	250.0	410.0
2	for	2	220.0	410.0
2	sal	1	90.0000	380.0
1	sal	2	210.0	390.0
1	Diff (1-2)		-35.0000	70.0000
2	Diff (1-2)		15.0000	130.0
Both	Diff (1-2)			
Both		Diff (1-2)		

The treatment difference "Diff (1-2)" corresponds to the "for" treatment minus the "sal" treatment, because "for" appears before "sal" in the output, according to the ORDER=MIXED default PROC TTEST option. Its mean estimate is 46.6071, favoring formoterol over salbutamol. The standard

deviation (Std Dev) reported for a "difference" is actually the pooled standard deviation across both treatment sequence (for/sal and sal/for), assuming equal variances. The standard error (Std Err) is the standard deviation of the mean estimate.

The top half of the table in Output 93.4.4 shows 95% two-sided confidence limits for the means for the same criteria addressed in the table in Output 93.4.3.

Output 93.4.4 Confidence Limits

Sequence	Treatment	Period	Method	Mean	95% CL Mean	
1	for	1		337.1	287.4	386.9
2	for	2		345.8	271.4	420.2
2	sal	1		283.3	172.7	393.9
1	sal	2		306.4	246.6	366.3
1	Diff (1–2)			30.7143	0.2238	61.2048
2	Diff (1–2)			62.5000	15.5972	109.4
Both	Diff (1–2)		Pooled	46.6071	22.8881	70.3262
Both	Diff (1–2)		Satterthwaite	46.6071	21.6585	71.5558
Both		Diff (1–2)	Pooled	−15.8929	−39.6119	7.8262
Both		Diff (1–2)	Satterthwaite	−15.8929	−40.8415	9.0558

Sequence	Treatment	Period	Method	Std Dev	95% CL Std Dev	
1	for	1		53.7631	34.6446	118.4
2	for	2		70.8814	44.2447	173.8
2	sal	1		105.4	65.7841	258.5
1	sal	2		64.7247	41.7082	142.5
1	Diff (1–2)			32.9682	21.2445	72.5982
2	Diff (1–2)			44.6934	27.8980	109.6
Both	Diff (1–2)		Pooled	19.3702	13.7217	32.8882
Both	Diff (1–2)		Satterthwaite			
Both		Diff (1–2)	Pooled	19.3702	13.7217	32.8882
Both		Diff (1–2)	Satterthwaite			

For the mean differences, both pooled (assuming equal variances for both treatment sequences) and Satterthwaite (assuming unequal variances) intervals are shown. For example, the pooled confidence limits for the overall treatment mean difference (for − sal) assuming equal variances are 22.8881 and 70.3262.

The bottom half of Output 93.4.4 shows 95% equal-tailed confidence limits for the standard deviations within each cell and for the treatment difference within each sequence. It also shows confidence limits for the pooled common standard deviation assuming equal variances. Note that the pooled standard deviation of 19.3702 and associated confidence limits 13.7217 and 32.8882 apply to both difference tests (treatment and period), since each of those tests involves the same pooled standard deviation.

Output 93.4.5 shows the results of *t* tests of treatment and period differences.

Output 93.4.5 *t* Tests

Treatment	Period	Method	Variances	DF	t Value	Pr > |t|
Diff (1-2)		Pooled	Equal	11	4.32	0.0012
Diff (1-2)		Satterthwaite	Unequal	9.1017	4.22	0.0022
	Diff (1-2)	Pooled	Equal	11	-1.47	0.1683
	Diff (1-2)	Satterthwaite	Unequal	9.1017	-1.44	0.1838

Both pooled and Satterhthwaite versions of the test of treatment difference are highly significant ($p = 0.0012$ and $p = 0.0022$), and both versions of the test of period difference are insignificant ($p = 0.1683$ and $p = 0.1838$).

The folded F test of equal variances in each treatment sequence is shown in Output 93.4.6.

Output 93.4.6 Equality of Variances Test

	Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	5	6	1.84	0.4797

The insignificant result ($p = 0.48$) implies a lack of evidence for unequal variances. However, it does not demonstrate equal variances, and it is not very robust to deviations from normality.

Output 93.4.7 shows the distribution of the response variables PEF1 and PEF2 within each of the four cells (combinations of two treatments and two periods) of the AB/BA crossover design, in terms of histograms and normal and kernel density estimates.

Output 93.4.7 Comparative Histograms
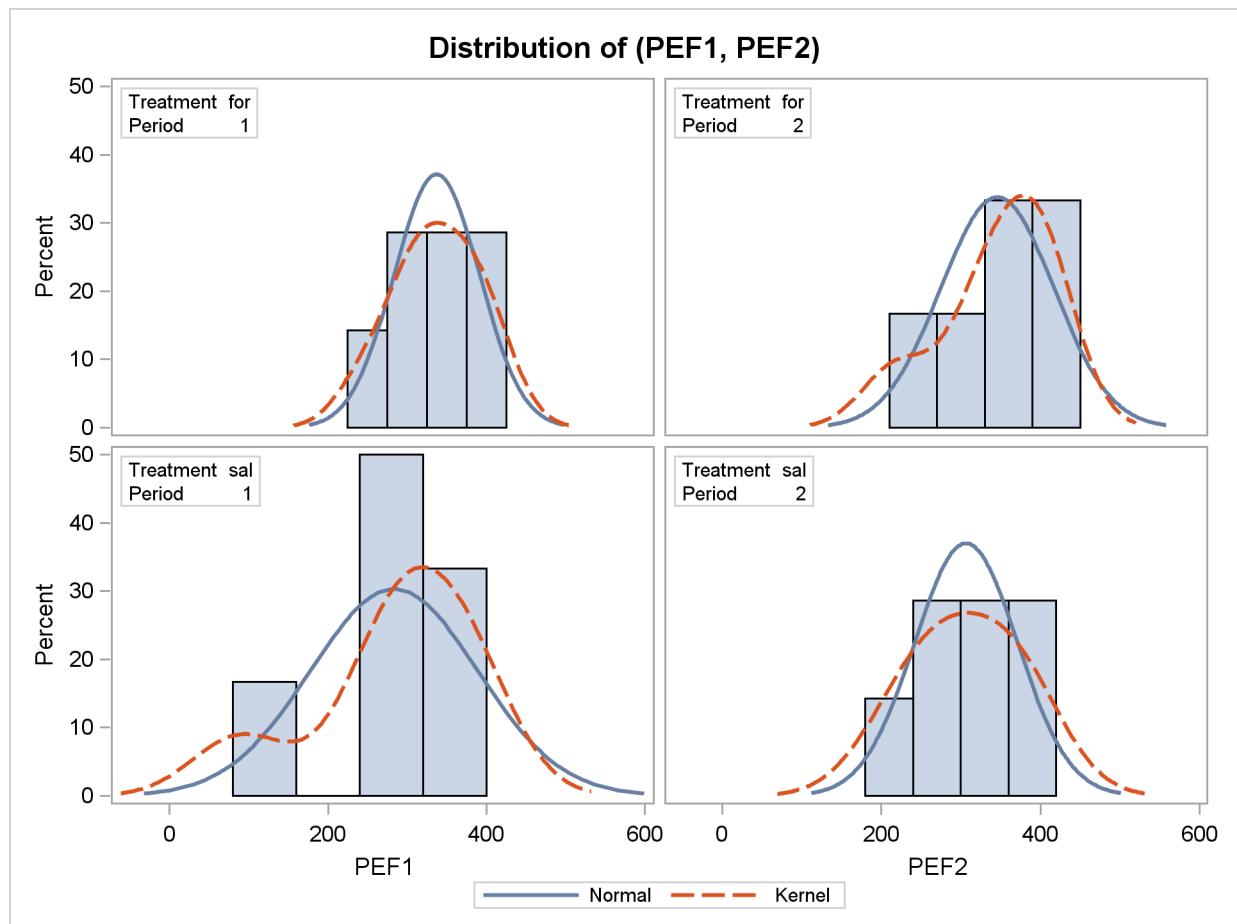
The distributions for the first treatment sequence (for/sal) appear to be somewhat symmetric, and the distributions for the sal/for sequence appear to be skewed to the left.

Output 93.4.8 shows a similar distributional summary but in terms of box plots.

Output 93.4.8 Comparative Box Plots

Box Plot of (PEF1, PEF2) by Treatment and Period
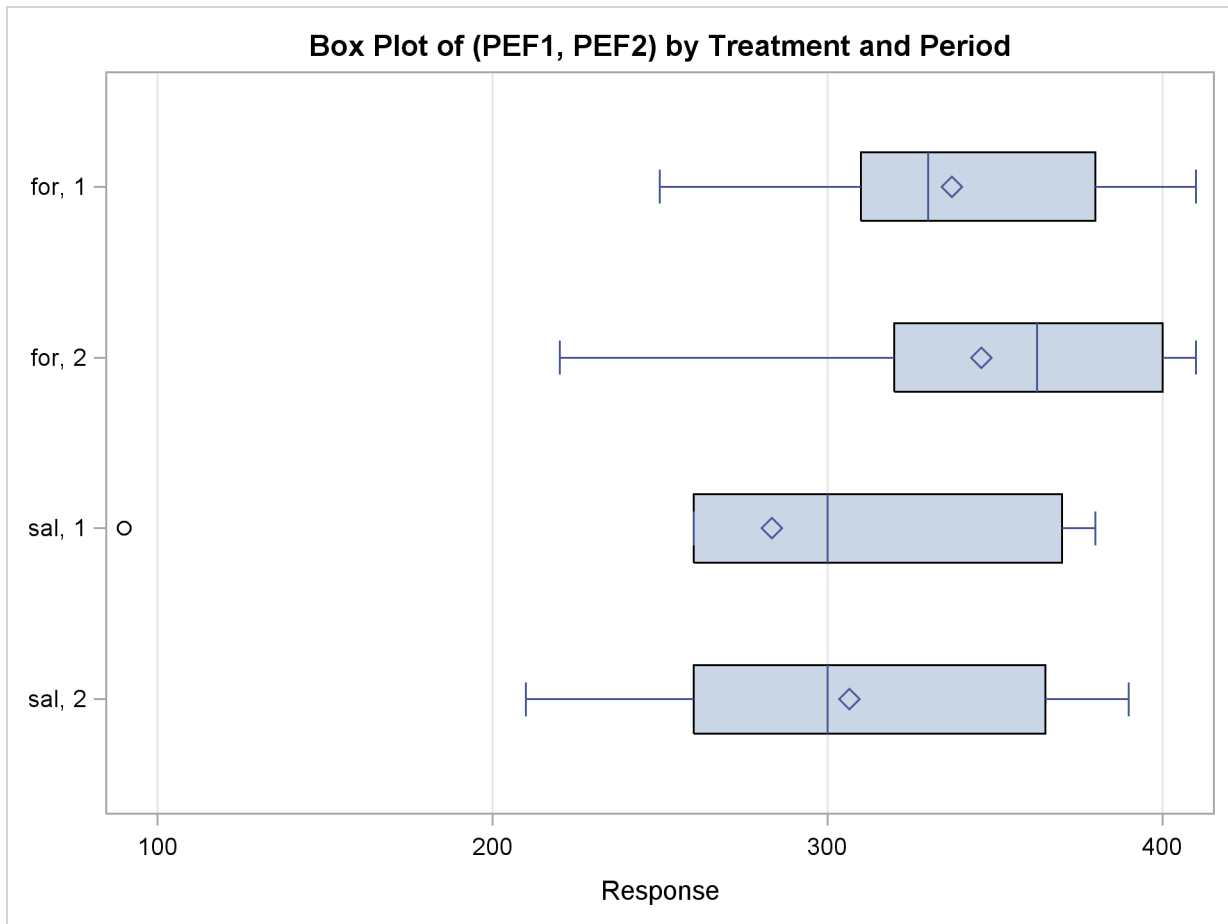
The relative locations of means and medians in each box plot corroborate the fact that the distributions for the sal/for sequence are skewed to the left. The distributions for the for/sal sequence appear to be skewed slightly to the right. The box plot for the salbutamol treatment in the first period shows an outlier (the circle on the far left side of the plot).

The treatment agreement plot in Output 93.4.9 reveals that only a single observation has a higher peak expiratory flow for salbutamol.

Output 93.4.9 Agreement of Treatments Plot
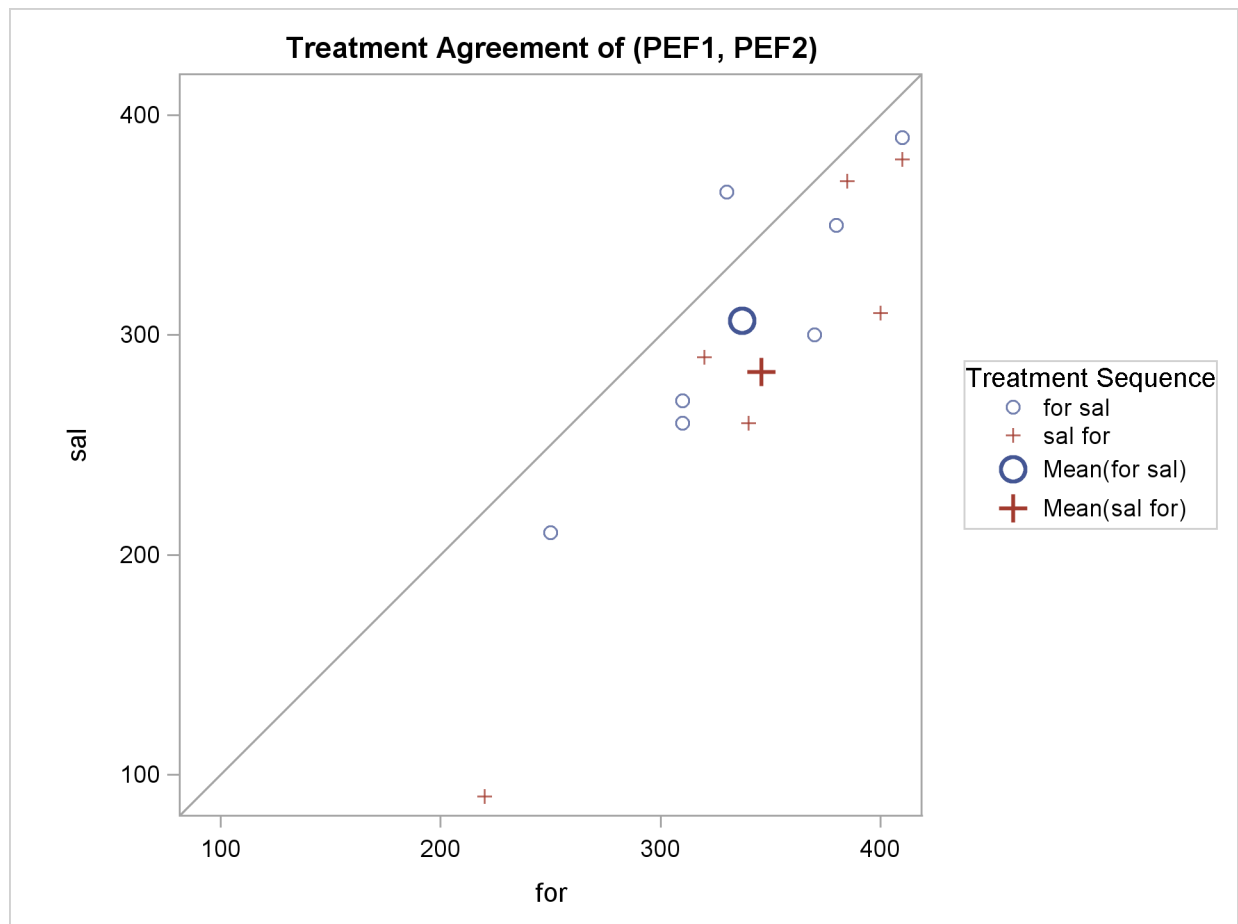
The mean for the sal/for treatment sequence is farther from the diagonal equivalence line, revealing that the treatment difference is more pronounced for the 6 observations in the sal/for sequence than for the 7 observations in the for/sal sequence. This fact is also seen numerically in Output 93.4.3 and Output 93.4.4, which show within-sequence treatment differences of 30.7 for for/sal and 62.5 for sal/for.

The profiles over treatment plot in Output 93.4.10 is a different view of the same information contained in Output 93.4.9, plotting the profiles from formoterol to salbutamol treatments. The lone observation for which the peak expiratory flow is higher for salbutamol appears as the only line with negative slope.

Output 93.4.10 Profiles over Treatment
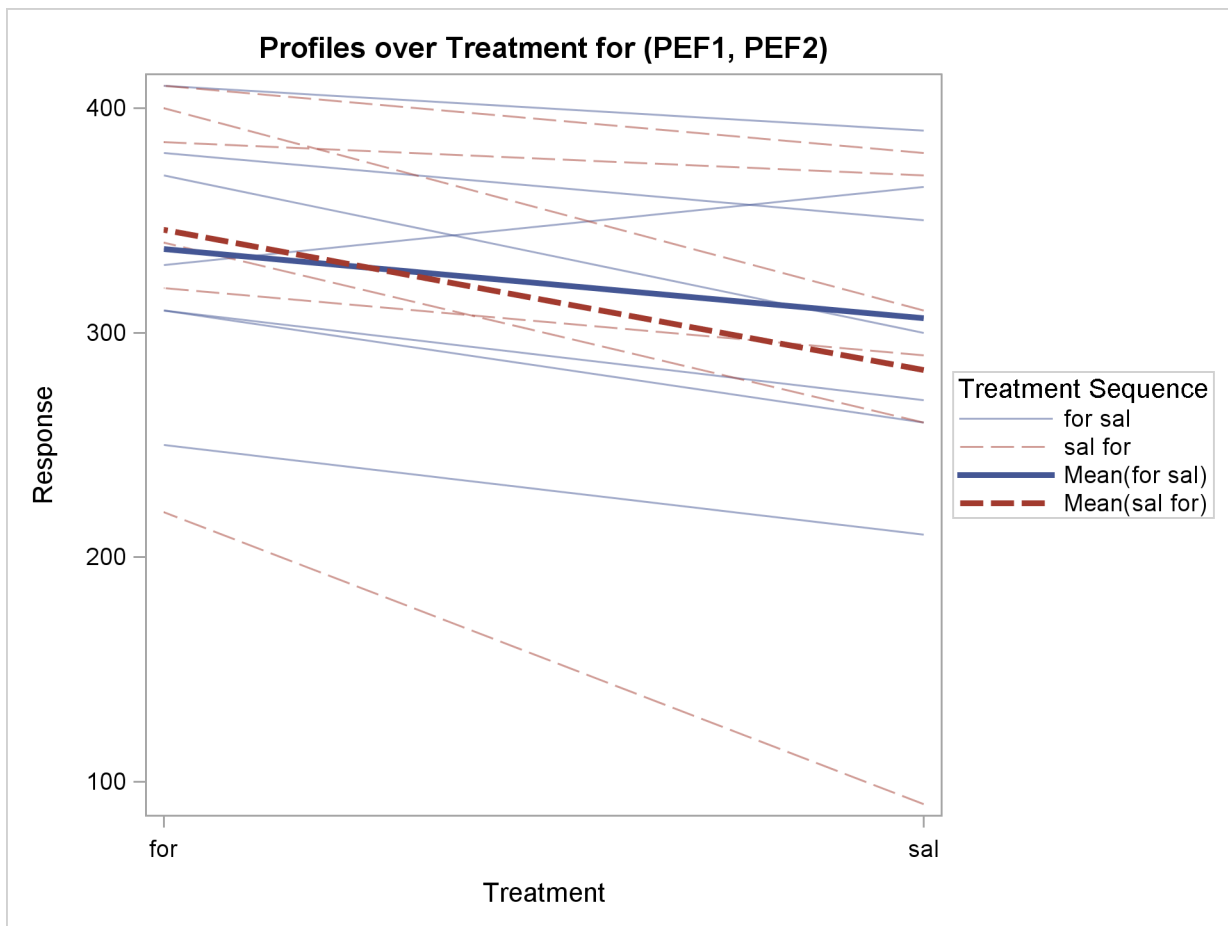
The Q-Q plots in Output 93.4.11 assess normality assumption within each of the four cells of the design.

Output 93.4.11 Q-Q Plots
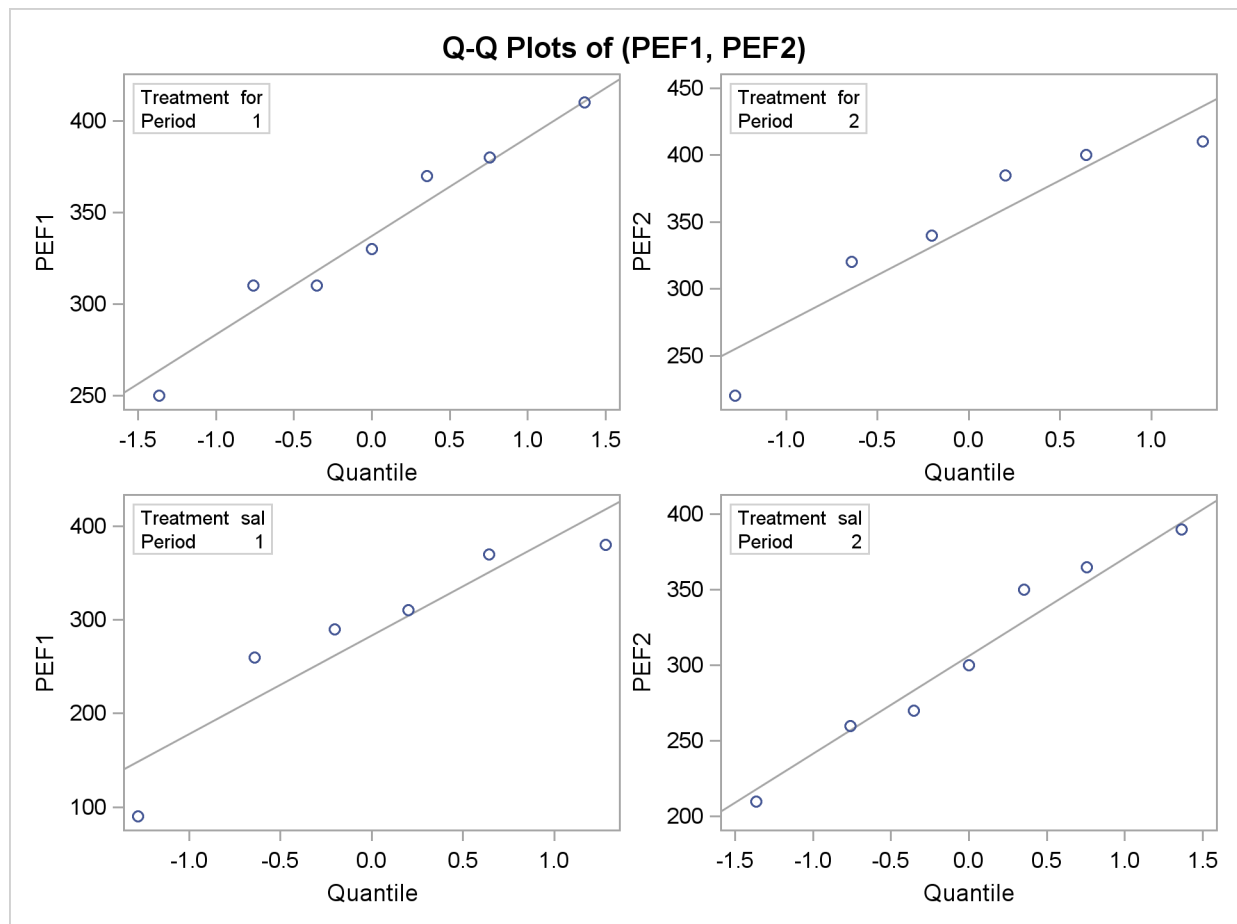
The two Q-Q plots for the sal/for sequence (lower left and upper right) suggest some possible normality violations in the tails, but the sample size is too small to make any strong conclusions. You could use the UNIVARIATE procedure with the NORMAL option to numerically check the normality assumptions.

Finally, Output 93.4.12 shows both pooled and Satterthwaite two-sided 95% confidence intervals for the treatment difference.

Output 93.4.12 Confidence Intervals for Treatment Difference
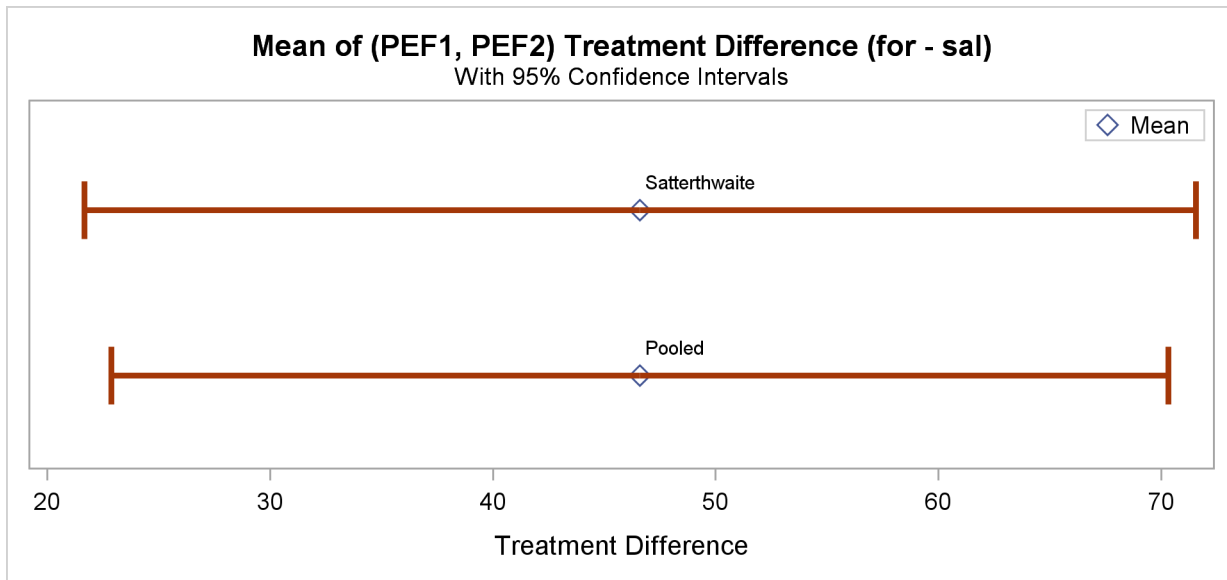
The pooled interval is slightly smaller than the Satterthwaite interval. (This is not always the case.)

Example 93.5: Equivalence Testing with Lognormal Data

Wellek (2003, p. 212) discusses an average bioequivalence study comparing the AUC (area under serum-concentration curve) measurements for two different drugs, denoted "Test" and "Reference," over a period of 20 hours. This example looks at a portion of Wellek's data, conducting an equivalence analysis with a paired design that uses AUC values on the original scale (assumed to be lognormally distributed). Each subject in the study received the Test drug upon one visit and then the Reference drug upon a later visit, sufficiently spaced so that no carryover effects would occur.

The goal is to test whether the geometric mean AUC ratio between Test and Reference is between 0.8 and 1.25, corresponding to the traditional FDA (80%, 125%) equivalence criterion. See the section "Arithmetic and Geometric Means" on page 7903 for a discussion of the use of geometric means for lognormal data.

The following SAS statements generate the data set:

```
data auc;
input TestAUC RefAUC @@;
datalines;
103.4 90.11   59.92 77.71   68.17 77.71   94.54 97.51
69.48 58.21   72.17 101.3   74.37 79.84   84.44 96.06
96.74 89.30   94.26 97.22   48.52 61.62   95.68 85.80
;
run;
```

You can display the data by using the following statements, which produce Output 93.5.1:

```
proc print data=auc;
run;
```

Output 93.5.1 AUC Data for Test and Reference Drugs

| | Test | |
Obs	AUC	RefAUC
1	103.40	90.11
2	59.92	77.71
3	68.17	77.71
4	94.54	97.51
5	69.48	58.21
6	72.17	101.30
7	74.37	79.84
8	84.44	96.06
9	96.74	89.30
10	94.26	97.22
11	48.52	61.62
12	95.68	85.80

The TestAUC and RefAUC variables represent the AUC measurements for each subject under the Test and Reference drugs, respectively. Use the following SAS statements to perform the equivalence analysis:

```
ods graphics on;

proc ttest data=auc dist=lognormal tost(0.8, 1.25);
   paired TestAUC*RefAUC;
run;

ods graphics off;
```

The ODS GRAPHICS statement requests graphical output. The DIST=LOGNORMAL option specifies the lognormal distributional assumption and requests an analysis in terms of geometric mean and coefficient of variation. The TOST option specifies the equivalence bounds 0.8 and 1.25.

Output 93.5.2 shows basic summary statistics for the ratio of TestAUC to RefAUC.

Output 93.5.2 Summary Statistics

The TTEST Procedure

Ratio: TestAUC / RefAUC

N	Geometric Mean	Coefficient of Variation	Minimum	Maximum
12	0.9412	0.1676	0.7124	1.1936

The geometric mean ratio of 0.9412 is the sample mean of the log-transformed data exponentiated to bring it back to the original scale. So the plasma concentration over the 20-hour period is slightly lower for the Test drug than for the Reference drug. The CV of 0.1676 is the ratio of the standard deviation to the (arithmetic) mean.

Output 93.5.3 shows the $100(1-\alpha)\% = 95\%$ confidence limits for the geometric mean ratio (0.8467 and 1.0462) and CV (0.1183 and 0.2884).

Output 93.5.3 Confidence Limits

Geometric Mean	95% CL Mean		Coefficient of Variation	95% CL CV	
0.9412	0.8467	1.0462	0.1676	0.1183	0.2884

Output 93.5.4 shows the $100(1-2\alpha)\% = 90\%$ confidence limits for the geometric mean ratio, 0.8634 and 1.0260.

Output 93.5.4 Equivalence Limits

Geometric Mean	Lower Bound	90% CL Mean		Upper Bound	Assessment
0.9412	0.8 <	0.8634	1.0260 <	1.25	Equivalent

The assessment of "Equivalent" reflects the fact that these limits are contained within the equivalence bounds 0.8 and 1.25. This result occurs if and only if the p-value of the test is less than the α value specified in the ALPHA= option in the PROC TTEST statement, and it is the reason that $100(1-2\alpha)\%$ confidence limits are shown in addition to the usual $100(1-\alpha)\%$ limits.

Output 93.5.5 shows the p-values for the two one-sided tests against the upper and lower equivalence bounds.

Output 93.5.5 TOST Equivalence Test

Test	Null	DF	t Value	P-Value
Upper	0.8	11	3.38	0.0031
Lower	1.25	11	−5.90	<.0001
Overall				0.0031

The overall p-value of 0.0031, the larger of the two one-sided p-values, indicates significant evidence of equivalence between the Test and Reference drugs.

The summary panel in Output 93.5.6 shows a histogram, kernel density, box plot, and $100(1-2\alpha)\% = 90\%$ confidence interval of the Test-to-Reference ratio of AUC, along with the equivalence bounds.

Output 93.5.6 Summary Panel
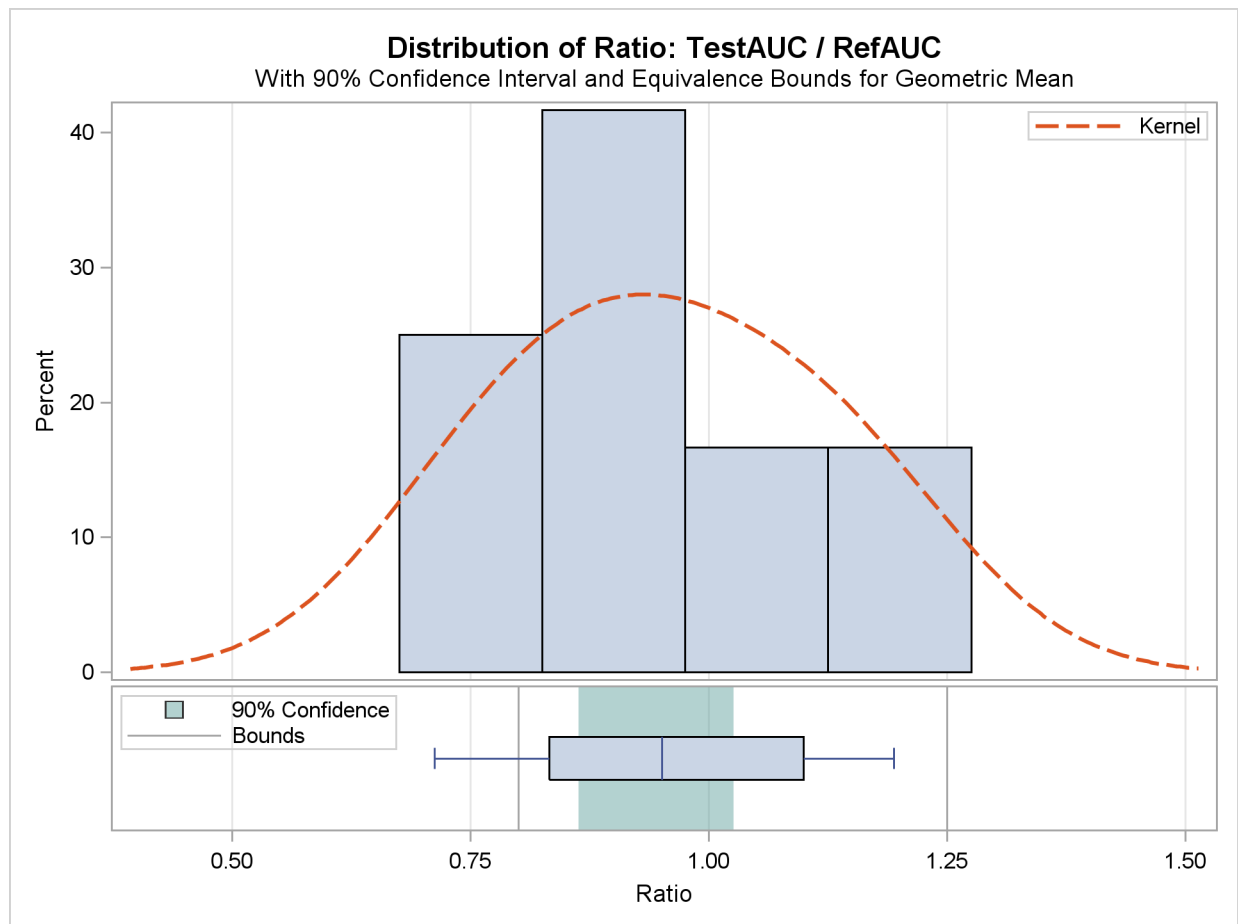
The confidence interval is closer to the lower equivalence bound than the upper bound and contained entirely within the bounds.

The agreement plot in Output 93.5.7 reveals that the only four subjects with higher AUC for the Test drug are at the far lower or far upper end of the AUC distribution. This might merit further investigation.

Output 93.5.7 Agreement Plot
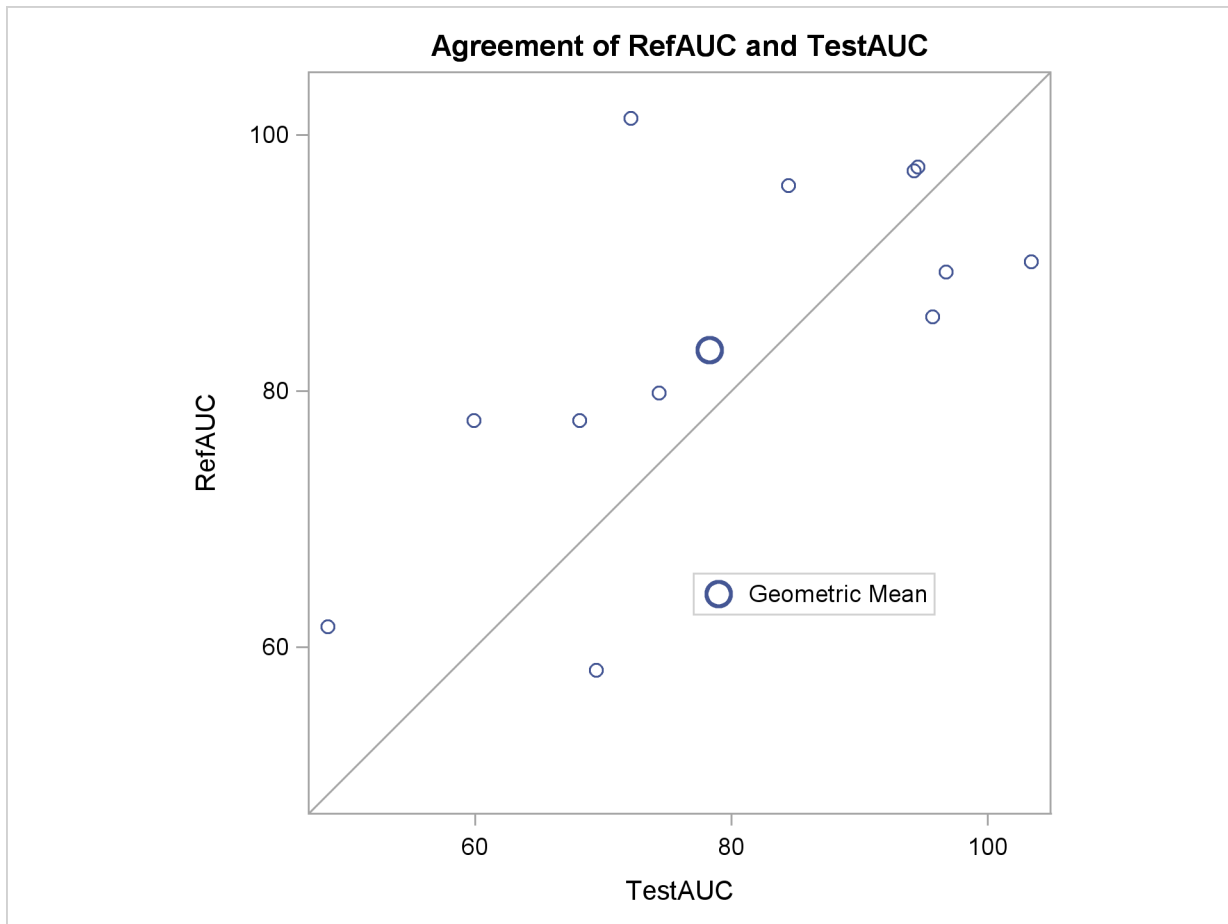
The profiles plot in Output 93.5.8 is a different view of the same information contained in Output 93.5.7, plotting the AUC from Test to Reference drug.

Output 93.5.8 Profiles Plot

Paired Profiles for (TestAUC, RefAUC)
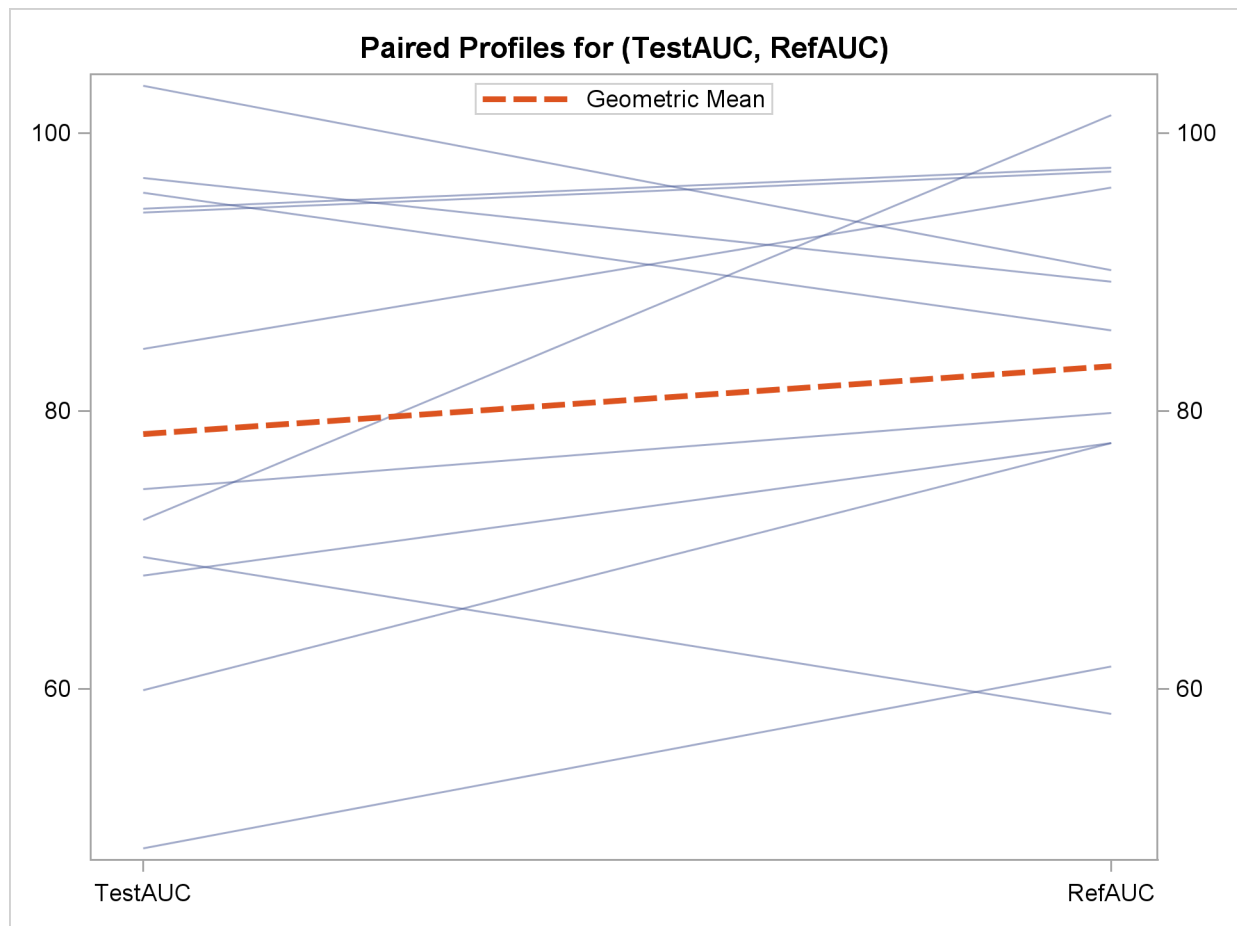
References

Best, D. I. and Rayner, C. W. (1987), "Welch's Approximate Solution for the Behren's-Fisher Problem," *Technometrics*, 29, 205–210.

Chow, S. and Liu, J. (2000), *Design and Analysis of Bioavailability and Bioequivalence Studies*, Second Edition, New York: Marcel Dekker.

Cochran, W. G. and Cox, G. M. (1950), *Experimental Designs*, New York: John Wiley & Sons.

Dilba, G., Schaarschmidt, F., and Hothorn, L. A. (2006), *mratios: A Package for Inference about Ratios of Normal Means,* UseR! Conference.

Diletti, D., Hauschke, D., and Steinijans, V. W. (1991), "Sample Size Determination for Bioequivalence Assessment by Means of Confidence Intervals," *International Journal of Clinical Pharmacology, Therapy and Toxicology*, 29, 1–8.

Fieller, E. C. (1954), "Some Problems in Interval Estimation," *Journal of the Royal Statistical Society*

Series B, 16, 175–185.

Hauschke, D., Kieser, M., Diletti, E., and Burke, M. (1999), "Sample Size Determination for Proving Equivalence Based on the Ratio of Two Means for Normally Distributed Data," *Statistics in Medicine*, 18, 93–105.

Huntsberger, David V. and Billingsley, Patrick P. (1989), *Elements of Statistical Inference*, Dubuque, IA: Wm. C. Brown.

Johnson, N. L. Kotz, S. and Balakrishnan, N. (1994), *Continuous Univariate Distributions, Volume 1*, Second Edition, New York: John Wiley & Sons.

Jones, B. and Kenward, M. G. (2003), *Design and Analysis of Cross-Over Trials*, Second Edition, Washington, DC: Chapman & Hall/CRC.

Lee, A. F. S. and Gurland, J. (1975), "Size and Power of Tests for Equality of Means of Two Normal Populations with Unequal Variances," *Journal of the American Statistical Association*, 70, 933–941.

Lehmann, E. L. (1986), *Testing Statistical Hypostheses*, New York: John Wiley & Sons.

Moore, David S. (1995), *The Basic Practice of Statistics*, New York: W. H. Freeman.

Phillips, K. F. (1990), "Power of the Two One-Sided Tests Procedure in Bioequivalence," *Journal of Pharmacokinetics and Biopharmaceutics*, 18, 137–144.

Posten, H. O., Yeh, Y. Y., and Owen, D. B. (1982), "Robustness of the Two-Sample t Test under Violations of the Homogeneity of Variance Assumption," *Communications in Statistics*, 11, 109–126.

Ramsey, P. H. (1980), "Exact Type I Error Rates for Robustness of Student's t Test with Unequal Variances," *Journal of Educational Statistics*, 5, 337–349.

Robinson, G. K. (1976), "Properties of Student's t and of the Behrens-Fisher Solution to the Two Mean Problem," *Annals of Statistics*, 4, 963–971.

SAS Institute Inc. (1986), *SUGI Supplemental Library User's Guide, Version 5 Edition*, Cary, NC: SAS Institute Inc.

Sasabuchi, S. (1988a), "A Multivariate Test with Composite Hypotheses Determined by Linear Inequalities When the Covariance Matrix Has an Unknown Scale Factor," *Memoirs of the Faculty of Science, Kyushu University, Series A*, 42, 9–19.

Sasabuchi, S. (1988b), "A Multivariate Test with Composite Hypotheses When the Covariance Matrix Is Completely Unknown," *Memoirs of the Faculty of Science, Kyushu University, Series A*, 42, 37–46.

Satterthwaite, F. W. (1946), "An Approximate Distribution of Estimates of Variance Components," *Biometrics Bulletin*, 2, 110–114.

Scheffe, H. (1970), "Practical Solutions of the Behrens-Fisher Problem," *Journal of the American Statistical Association*, 65, 1501–1508.

Schuirmann, D. J. (1987), "A Comparison of the Two One-Sided Tests Procedure and the Power

Approach for Assessing the Equivalence of Average Bioavailability," *Journal of Pharmacokinetics and Biopharmaceutics*, 15, 657–680.

Senn, S. (2002), *Cross-over Trials in Clinical Research*, Second Edition, New York: John Wiley & Sons.

Steel, R. G. D. and Torrie, J. H. (1980), *Principles and Procedures of Statistics*, Second Edition, New York: McGraw-Hill.

Tamhane, A. C. and Logan, B. R. (2004), "Finding the Maximum Safe Dose Level for Heteroscedastic Data," *Journal of Biopharmaceutical Statistics*, 14, 843–856.

Wang, Y. Y. (1971), "Probabilities of the Type I Error of the Welch Tests for the Behren's-Fisher Problem," *Journal of the American Statistical Association*, 66, 605–608.

Wellek, S. (2003), *Testing Statistical Hypotheses of Equivalence*, Boca Raton, FL: Chapman & Hall/CRC Press LLC.

Yuen, K. K. (1974), "The Two-Sample Trimmed t for Unequal Population Variances," *Biometrika*, 61, 165–170.

Subject Index

Syntax Index

Your Turn

We welcome your feedback.

- If you have comments about this book, please send them to **yourturn@sas.com**. Include the full title and page numbers (if applicable).

- If you have comments about the software, please send them to **suggest@sas.com**.

SAS® Publishing Delivers!

Whether you are new to the work force or an experienced professional, you need to distinguish yourself in this rapidly changing and competitive job market. SAS® Publishing provides you with a wide range of resources to help you set yourself apart. Visit us online at support.sas.com/bookstore.

SAS® Press

Need to learn the basics? Struggling with a programming problem? You'll find the expert answers that you need in example-rich books from SAS Press. Written by experienced SAS professionals from around the world, SAS Press books deliver real-world insights on a broad range of topics for all skill levels.

support.sas.com/saspress

SAS® Documentation

To successfully implement applications using SAS software, companies in every industry and on every continent all turn to the one source for accurate, timely, and reliable information: SAS documentation. We currently produce the following types of reference documentation to improve your work experience:

- Online help that is built into the software.
- Tutorials that are integrated into the product.
- Reference documentation delivered in HTML and PDF – **free** on the Web.
- Hard-copy books.

support.sas.com/publishing

SAS® Publishing News

Subscribe to SAS Publishing News to receive up-to-date information about all new SAS titles, author podcasts, and new Web site features via e-mail. Complete instructions on how to subscribe, as well as access to past issues, are available at our Web site.

support.sas.com/spn

§sas. | THE POWER TO KNOW®