Equivalence and Noninferiority Tests

(Comparing Variance to Target)
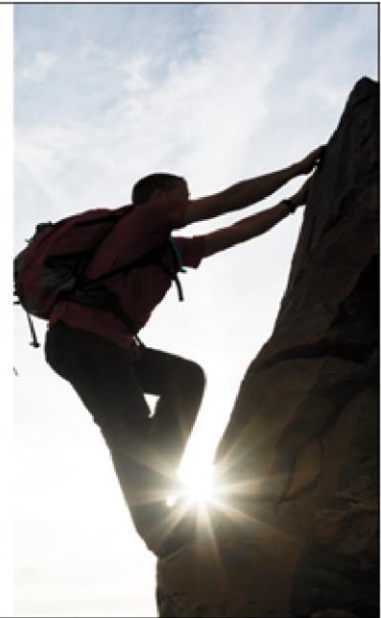
statgraphics® centurion
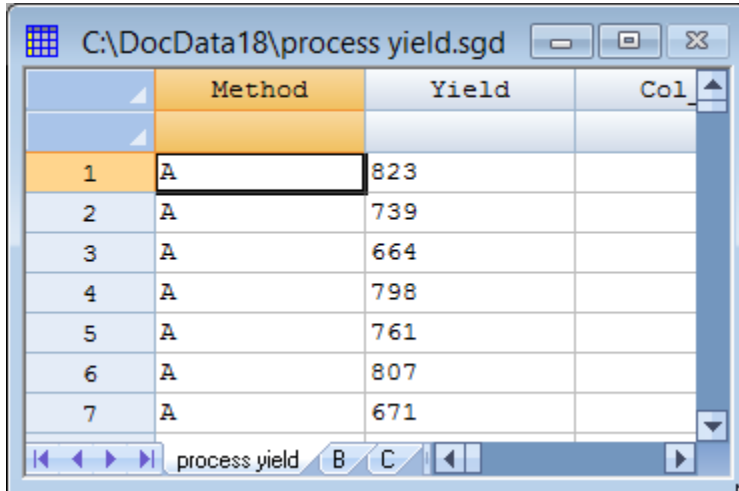
Revised: 12/11/2019

Summary

This procedure tests whether the standard deviation of a sample obtained from a single population may be considered to be equivalent to a target value. A standard deviation is considered to be "equivalent" if it falls within a specified interval surrounding that target value. Unlike standard hypothesis tests which are designed to prove that a standard deviation is significantly different than a specified value, equivalence tests are designed to prove that the standard deviation is essentially equivalent to the target.

The procedure may also be used to demonstrate noninferiority. A sample is considered to be "noninferior" if the difference between the standard deviation and the target value is no greater than (or no less than) a specified value. This situation corresponds to a one-sided test of equivalence.

Sample StatFolio: *equivalence to sigma target.sgp*

Sample Data:

The file *process yield.sgd* contains measurements of the yield of a product produced using 3 methods (A, B, and C). 50 batches of each product were produced. A portion of the data is shown below:

	Method	Yield	Col
1	A	823	
2	A	739	
3	A	664	
4	A	798	
5	A	761	
6	A	807	
7	A	671	

C:\DocData18\process yield.sgd

process yield / B / C

We wish to determine which methods have standard deviations that are equivalent to a target value of 50.

Data Input

To perform the desired equivalence tests, select **Compare – Equivalence and Noninferiority Tests - Comparison of Variance to Target** from the main menu. The first dialog box displayed asks the user to specify the manner in which the data have been entered:
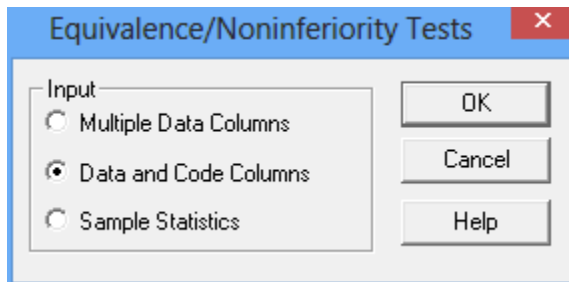
Data may be entered in any of 3 formats:

1. *Multiple Data Columns* – the data for each sample are placed in a separate column.

2. *Data and Code Columns* – all of the data are placed in a single column and a second column is created identifying the group that each observation corresponds to.

3. *Sample Statistics* – rather than entering the original data values, columns are created containing the sample sizes, means, standard deviations.

Multiple Data Columns

In this format, the data for each sample are placed in a separate column as in the datasheet shown below:

The data input dialog box for this format requests the names of the columns:

At least 2 columns containing data must be specified.

Data and Code Columns

In this format, all of the data is placed in a single column and a second column is created identifying which data belong to which samples:

The data input dialog box for this format requests the names of the data and code columns:

Note that the *Factor* column may be either numeric or character.

Sample Statistics

In this format, columns are created to contain the sample sizes, means, standard deviations and an optional label for each sample:

	Method	Count	Mean	Std deviation
1	A	50	744.26	46.5586
2	B	50	752.64	40.3782
3	C	50	775.68	49.4981
4				
5				
6				

The data input dialog box for this format requests the names of the columns with the sample statistics:
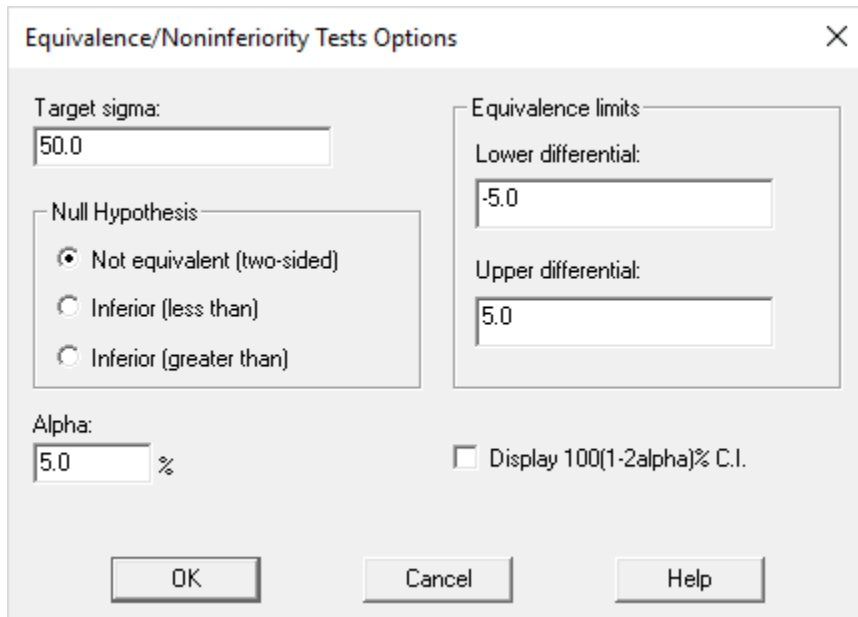
Analysis Options

Once the data is specified, a third dialog box is displayed on which to specify the hypothesis to be tested.

The most common type of test is a two-sided test of equivalence. In such a test, the null hypothesis is that the standard deviation of the population from which the sample has been obtained, σ, is not equivalent to a target value T. By not equivalent, we mean that the difference between the standard deviation and the target σ - T is either less than some lower differential Δ_L, or greater than some upper differential Δ_U:

Null hypothesis: σ - T $< \Delta_L$ or σ - T $> \Delta_U$

If this hypothesis is rejected, then we will have demonstrated that the difference between the standard deviation and the target satisfies $\Delta_L \leq \sigma$ - T $\leq \Delta_U$, which is our definition of equivalence.

To demonstrate equivalence, Statgraphics uses the TOST procedure of Schuirman (1987). This procedure consists of two one-sided tests: an upper-tailed test used to demonstrate that σ - T $\geq \Delta_L$ and a lower-tailed test used to demonstrate that σ - T $\leq \Delta_U$. Obtaining significant results on both tests allows an assertion of equivalence between the population and target standard deviations.

The fields on the *Analysis Options* dialog box specify:

- **Target standard deviation:** the target value T.

- **Null hypothesis:** whether to perform a two-tailed test of equivalence as described above or a one-tailed test of noninferiority. In the latter case, the null hypothesis is one of the following:

    "Less than" null hypothesis: σ - T $< \Delta_L$

    "Greater than" null hypothesis: σ - T $> \Delta_U$

- **Equivalence limits:** the value of the lower differential Δ_L and the upper differential Δ_U.

- **Alpha:** the significance level at which the tests will be performed.

- **Display 100(1-2alpha)% C.I.:** when displaying confidence intervals, use $(1-2\alpha)$ instead of $(1-\alpha)$.

Analysis Summary

The *Analysis Summary* for the sample data using the default options is shown below:

Equivalence & Noninferiority Tests - Comparison of Variance to Target

Dependent variable: Yield
Factor: Method

Sample Statistics

Sample	n	Minimum	Maximum	Mean	Std. deviation
A	50	664.0	844.0	744.26	46.5586
B	50	672.0	844.0	752.64	40.3782
C	50	667.0	892.0	775.68	49.4981

Equivalence Analysis
Target sigma: 50.0
Null hypothesis: Not equivalent (two-sided)
Lower equivalence differential: -5.0
Upper equivalence differential: 5.0

Comparison	Difference	95% CL	95% CL
A v target	-3.44136	40.0143	55.9506
B v target	-9.62181	34.7026	50.0
C v target	-0.501932	42.5405	59.4829

Comparison	Lower chi-squared value	Upper chi-squared value	Lower P-value	Upper P-value
A v target	52.4531	35.1133	0.3416	0.0676
B v target	39.4516	26.4098	0.8332	0.0034
C v target	59.2854	39.6869	0.1491	0.1737

| Comparison	Maximum P-value	Conclusion (alpha=5%)
A v target	0.3416	Equivalence has not been demonstrated.
B v target	0.8332	Equivalence has not been demonstrated.
C v target	0.1737	Equivalence has not been demonstrated.

The top of the output displays summary statistics for each sample. This is followed by an *Equivalence Analysis* which compares each sample standard deviation to the target value. In the example, the null hypothesis is that the standard deviation is not within the equivalence range of 45 to 55 calculated from 50±5.

The output then displays the estimated difference between each standard deviation and the target, together with a $100(1-2\alpha)\%$ conference interval for the mean. If the confidence interval is entirely within the equivalence range, then equivalence can be asserted. Otherwise, it cannot. In the example, no methods have both confidence limits entirely between 45 and 55.

An equivalent method for determining whether a standard deviation is equivalent to a target value is to run two one-sided tests, one against the target value plus the lower differential and another against the target value plus the upper differential. If both P-values are less than α, then equivalence can be asserted. The summary table shows the greater of the two P-values for each standard deviation and cannot asserts equivalence for any of the methods.
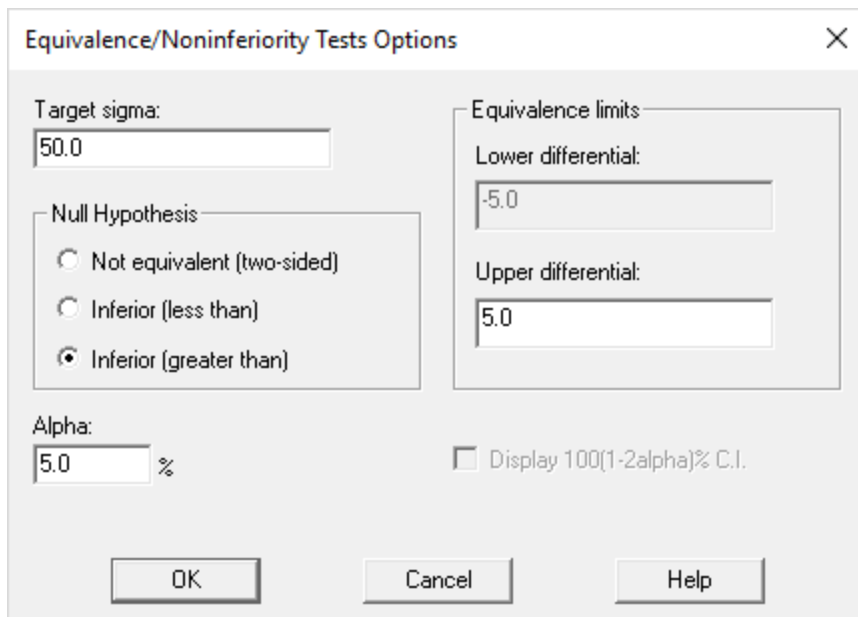
Equivalence Plot

This plot shows the confidence intervals for each standard deviation. If an interval is contained entirely in the region between the lower and upper equivalence limits, then the standard deviation may be asserted to be equivalent to the target value.

Equivalence Test - alpha = 5%
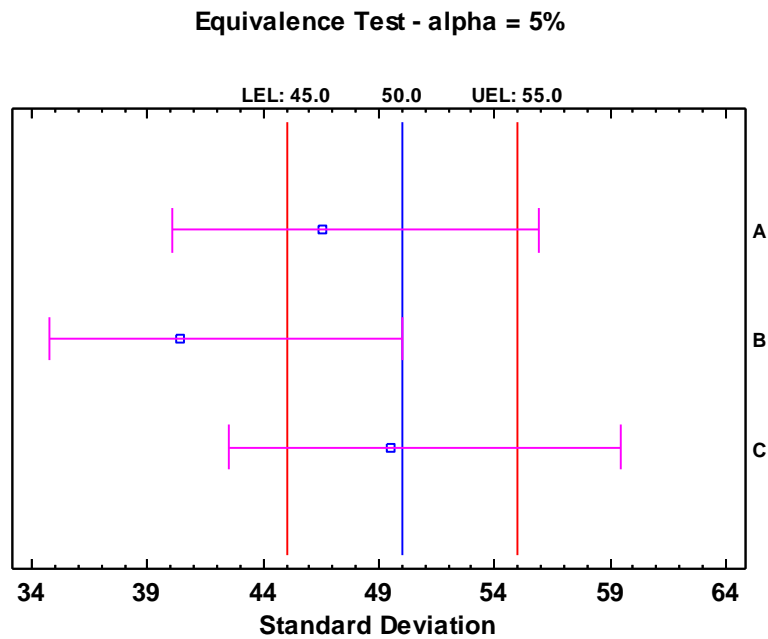
One-Sided Noninferiority Tests

In some circumstances, the desired goal is not one of showing that the difference between a standard deviation and a target value is within some specified range. Instead, the goal is either to show that the difference is no bigger than some value Δ_U or to show that the difference is no smaller than some value Δ_L. Rejection of a null hypothesis in such a one-sided situation leads to the assertion that the standard deviation is not inferior to the target (it might be either equivalent or superior).

For example, suppose it was desired to show that the standard deviation was no more than 5 units greater than the target. In such a case, the *Analysis Options* dialog box would be completed as shown below:

In this case, the null hypothesis is $\sigma - T > 25$. If this hypothesis can be rejected, then we can claim that the standard deviation is not inferior to the target.

For the sample data, the *Analysis Summary* is shown below:

Equivalence & Noninferiority Tests - Comparison of Variance to Target

Dependent variable: Yield

Factor: Method

Sample Statistics

Sample	n	Minimum	Maximum	Mean	Std. deviation
A	50	664.0	844.0	744.26	46.5586
B	50	672.0	844.0	752.64	40.3782
C	50	667.0	892.0	775.68	49.4981

Equivalence Analysis

Target sigma: 50.0

Null hypothesis: Inferior (greater than)

Upper equivalence differential: 5.0

Comparison	Difference	95% CL
A v target	-3.44136	55.9506
B v target	-9.62181	48.5234
C v target	-0.501932	59.4829

Comparison	Upper chi-squared value	Upper P-value
A v target	35.1133	0.0676
B v target	26.4098	0.0034
C v target	39.6869	0.1737

Comparison	Maximum P-value	Conclusion (alpha=5%)
A v target	0.0676	Noninferiority has not been demonstrated.
B v target	0.0034	Noninferiority has been demonstrated.
C v target	0.1737	Noninferiority has not been demonstrated.

For each sample, the output displays an upper confidence bound for the standard deviation. If the upper confidence bound is less than the target plus the upper equivalence differential, the P-value of a lower-tailed test comparing the difference to Δ_U will be less than alpha and noninferiority may be asserted.

The *Equivalence Plot* displays the one-sided confidence bounds for each standard deviation:

Noninferiority Test - alpha = 5%
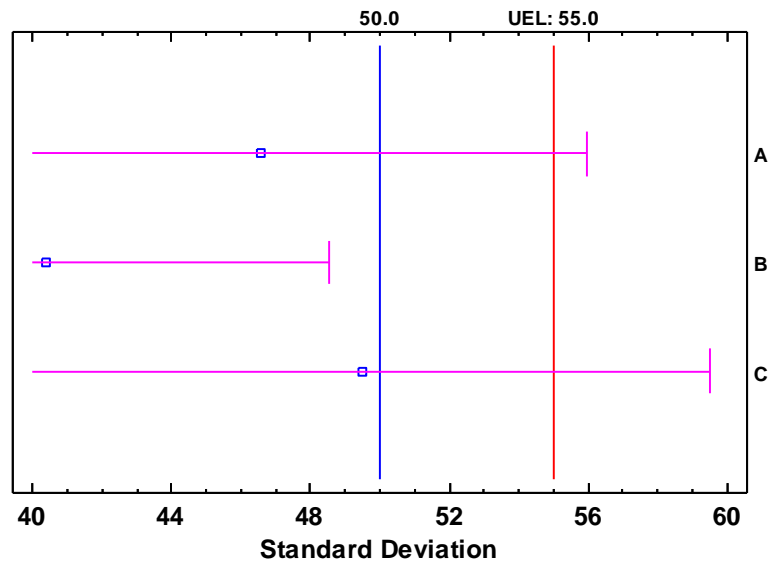
Noninferiority may be asserted for any standard deviation in which the confidence bound does not contain the UEL.

Calculations

By default, the confidence intervals are calculated by:

$$\left[min\left(T, \frac{(n-1)s^2}{X^2_{\alpha,n-1}} \right), max\left(T, \frac{(n-1)s^2}{X^2_{1-\alpha,n-1}} \right) \right]$$

If "Display 100(1-2alpha) C.I." is selected on the Analysis Options dialog box, the confidence intervals are calculated by:

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

References

Berger, R.L. and Hsu, J.C. (1995). "Bioequivalence trials, intersection-union tests, and equivalence confidence sets." Institute of Statistics Mimeo Series Number 2279.

Chow, S.-H. and Shao, J. (2002). <u>Statistics in Drug Research: Methodologies and Recent Developments</u>. New York: Marcel-Dekker.

Hsu, J.C., Hwang, J.T.G., Liu, H.-K., and Ruberg, S.J. (1994). "Confidence intervals associated with tests for bioequivalence." Biometrika 81: 103-114.

Schuirmann, D.J. (1987). "A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability." J. Pharmacokinet Biopharm. 15(6): 657-680.