Chapter 562

Power Comparison of Tests of Variances (Simulation)

Introduction

This procedure analyzes the power and significance level of tests which may be used to test statistical hypotheses about variances (homogeneity) in a one-way experimental design. For each scenario that is set up, two simulations are run. One simulation estimates the significance level and the other estimates the power.

The four tests that are compared in this procedure are

1. Bartlett's Test

2. Levene's Test

3. Brown-Forsythe Test

4. Conover Test using Ranks

Technical Details

Computer simulation allows us to estimate the power and significance level that is actually achieved by a test procedure in situations that are not mathematically tractable. Computer simulation was once limited to mainframe computers. But, in recent years, as computer speeds have increased, simulation studies can be completed on desktop and laptop computers in a reasonable period of time.

The steps to a simulation study are

1. Specify how the test is carried out. This includes indicating how the test statistic is calculated and how the significance level is specified.

2. Generate random samples from the distributions specified by the <u>alternative</u> hypothesis. Calculate the test statistics from the simulated data and determine if the null hypothesis is accepted or rejected. Tabulate the number of rejections and use this to calculate the test's power.

3. Generate random samples from the distributions specified by the <u>null</u> hypothesis. Calculate each test statistic from the simulated data and determine if the null hypothesis is accepted or rejected. Tabulate the number of rejections and use this to calculate the test's significance level.

4. Repeat steps 2 and 3 several thousand times, tabulating the number of times the simulated data leads to a rejection of the null hypothesis. The power is the proportion of simulated samples in step 2 that lead to rejection. The significance level is the proportion of simulated samples in step 3 that lead to rejection.

Generating Random Distributions

Two methods are available in PASS to simulate random samples. The first method generates the random variates directly, one value at a time. The second method generates a large pool (over 10,000) of random values and then draws the random numbers from this pool. This second method can cut the running time of the simulation by 70%.

The second method begins by generating a large pool of random numbers from the specified distributions. Each of these pools is evaluated to determine if its mean is within a small relative tolerance (0.0001) of the target mean. If the actual mean is not within the tolerance of the target mean, individual members of the population are replaced with new random numbers if the new random number moves the mean towards its target. Only a few hundred such swaps are required to bring the actual mean to within tolerance of the target mean. This population is then sampled with replacement using the uniform distribution. We have found that this method works well as long as the size of the pool is at least the maximum of twice the number of simulated samples desired and 10,000.

Tests

The details of the four tests that are compared here are found in the individual chapters that discuss each procedure and they are not duplicated here.

Procedure Options

This section describes the options that are specific to this procedure. These are located on the Design and Simulation tabs. For more information about the options of other tabs, go to the Procedure Window chapter.

Design Tab

The Design tab contains the parameters and options needed to described the experimental design, the data distributions, the error rates, and sample sizes.

Alpha and Simulations

Alpha

This option specifies one or more values for the probability of a type-I error, alpha. A type-I error occurs when a true null hypothesis is rejected. In this procedure, a type-I error occurs when you reject the null hypothesis of equal means when in fact the means are equal.

Alpha values must be between zero and one. Historically, 0.05 has been used for alpha. This means that about one test in twenty will falsely reject the null hypothesis. You should pick a value for alpha that represents the risk of a type-I error you are willing to take in your experimental situation.

You may enter a single value or a range of values such as *0.01 0.05 0.10* or *0.01 to 0.10 by 0.01*.

Simulations

This option specifies the number of simulations, *M*. The larger the number of simulations, the longer the running time and the more accurate the results.

The precision of the simulated power estimates are calculated from the binomial distribution. Thus, confidence intervals may be constructed for various power values using the binomial distribution. The following table gives an estimate of the precision that is achieved for various simulation sizes when the power is either 0.50 or 0.95. The table values are interpreted as follows: a 95% confidence interval of the true power is given by the power reported by the simulation plus and minus the 'Precision' amount given in the table.

Power Comparison of Tests of Variances (Simulation)

Simulation Size M	Precision when Power = 0.50	Precision when Power = 0.95
100	0.100	0.044
500	0.045	0.019
1000	0.032	0.014
2000	0.022	0.010
5000	0.014	0.006
10000	0.010	0.004
50000	0.004	0.002
100000	0.003	0.001

Notice that a simulation size of 1000 gives a precision of plus or minus 0.01 when the true power is 0.95. Also note that as the simulation size is increased beyond 5000, there is only a small amount of additional accuracy achieved.

Number of Groups and Sample Size Allocation

Number of Groups

Enter the number of groups in the design. The program can handle between 2 and 20 groups.

Group Allocation

Select the method used to allocate subjects to groups. The choices are

- **Equal (n = n1 = n2 = ...)**

  All group sample sizes are the same. The value(s) of n are entered in the *n (Group Size)* box that will appear immediately below this box. See also *n (Group Size)* below.

- **Enter base group size and group multipliers**

  The group sample sizes are found by multiplying the corresponding *Multiplier* value (found in the Power Simulation section) times the *Base Group Size* that will appear immediately below this box. See also *Base Group Size* and *Multiplier* below.

- **Enter total sample size and group percentages**

  The group sample sizes are found by taking the corresponding percentage (found in the Power Simulation section) of the *N (Total Size)* value that will appear immediately below this box. See also *N (Total Size)* and *Percent of N* below.

  Note that the *Percent of N* values are adjusted so they sum to 100%.

- **Enter individual group sample sizes**

  Enter each group's sample size directly in the corresponding *Sample Size* box in the Power Simulation section. See also *Sample Size* below.

n (Group Size)

This is the group sample size of all groups. One or more values, separated by blanks or commas, may be entered. A separate analysis is performed for each value listed here.

Base Group Size

This is the base sample size per group. One or more values, separated by blanks or commas, may be entered. A separate analysis is performed for each value listed here.

The individual group samples sizes are determined by multiplying this value times the corresponding *Multiplier* value entered in the Power Simulation section.

If the Multiplier numbers are represented by

m1, m2, m3, ...

and this value is represented by n, the group sample sizes are calculated as follows:

n1 = [n(m1)]

n2 = [n(m2)]

n3 = [n(m3)]

where the operator, [X] means the next integer after X, e.g. [3.1]=4.

For example, suppose there are three groups and the multipliers are set to 1, 2, and 3. If *n* is 5, the resulting group sample sizes will be 5, 10, and 15.

N (Total Size)

This is the total sample size. One or more values, separated by blanks or commas, may be entered. A separate analysis is performed for each value listed here.

The individual group samples sizes are determined by multiplying this value times the corresponding *Percent of N* value entered in the Power Simulation section.

If the Percent values are represented by

p1, p2, p3, ...

and this value is represented by N, the group sample sizes are calculated as follows:

n1 = [N(p1)]

n2 = [N(p2)]

n3 = [N(p3)]

where the operator, [X] means the next integer after X, e.g. [3.1]=4.

For example, suppose there are three groups and the percentages are set to 25, 25, and 50. If N is 36, the resulting group sample sizes will be 9, 9, and 18.

Power Simulation

These options specify the distributions to be used in the power simulation, one row per group. The first option specifies distribution. The second option, if visible, specifies the sample size of that group.

Distribution

Specify the distribution of each group under the alternative hypothesis, H1. This distribution is used in the simulation that determines the power.

A fundamental quantity in a power analysis is the amount of variation among the group means. In fact, in classical power analysis formulas, this variation is summarized as the standard deviation of the means. You must pay particular attention to the values you give to the means of these distributions because they are fundamental to the interpretation of the simulation.

For convenience in specifying a range of values, the parameters of the distribution can be specified using numbers or letters. If letters are used, their values are specified in the *Parameter Values* boxes below.

Following is a list of the distributions that are available and the syntax used to specify them. Each of the parameters should be replaced with a number or parameter name.

Distributions with Common Parameters

Beta(Shape1, Shape2, Min, Max)

Binomial(P,N)

Cauchy(Mean,Scale)

Constant(Value)

Exponential(Mean)

Gamma(Shape,Scale)

Gumbel(Location,Scale)

Laplace(Location,Scale)

Logistic(Location,Scale)

Lognormal(Mu,Sigma)

Multinomial(P1,P2,P3,...,Pk)

Normal(Mean,Sigma)

Poisson(Mean)

TukeyGH(Mu,S,G,H)

Uniform(Min, Max)

Weibull(Shape,Scale)

Distributions with Mean and SD Parameters

BetaMS(Mean,SD,Min,Max)

BinomialMS(Mean,N)

GammaMS(Mean,SD)

GumbelMS(Mean,SD)

LaplaceMS(Mean,SD)

LogisticMS(Mean,SD)

LognormalMS(Mean,SD)

UniformMS(Mean,SD)

WeibullMS(Mean,SD)

Details of writing mixture distributions, combined distributions, and compound distributions are found in the chapter on *Data Simulation* and will not be repeated here.

Finding the Value of the Mean of a Specified Distribution

The mean of a distribution created as a linear combination of other distributions is found by applying the linear combination to the individual means. However, the mean of a distribution created by multiplying or dividing other distributions is not necessarily equal to applying the same function to the individual means. For example, the mean of 4 Normal(4, 5) + 2 Normal(5, 6) is 4*4 + 2*5 = 26, but the mean of 4 Normal(4, 5) * 2 Norma (5, 6) is not exactly 4*4*2*5 = 160 (although it is close).

Multiplier

The group allocation *Multiplier* for this group is entered here. Typical values of this parameter are 0.5, 1, and 2.

- **Solve For = Power**

  The group sample size is found by multiplying this value times the *Base Group Size* value and rounding up to the next whole number.

- **Solve For = Sample Size**

  The group sample size is found by multiplying this value times the *base group size* that is being searched for and rounding up to the next whole number.

Percent of N

This is the percentage of the total sample size that is allocated to this group. The individual group samples sizes are determined by multiplying this value times the *N (Total Size)* value entered above.

If these values are represented by *p1, p2, p3, ...* the group sample sizes are calculated as follows:

n1 = [N(p1)]

n2 = [N(p2)]

n3 = [N(p3)]

where the operator, [X] means the next integer after X, e.g. [3.1]=4.

For example, suppose there are three groups and these percentages are set to 25, 25, and 50. If N is 36, the resulting group sample sizes will be 9, 9, and 18.

Sample Size

This option allows you to enter the sample size of this group directly.

Power Simulation – Parameter Values for Group Distributions

These options specify the names and values for the parameters used in the distributions.

Name

Up to six sets of named parameter values may be used in the simulation distributions. This option lets you select an appropriate name for each set of values. Possible names are

- **M1 to M5**

  You might use M1 to M5 for means.

- **SD**

  You might use SD for standard deviations.

- **A to T**

  You might use these letters for location, shape, and scale parameters.

Value(s)

These values are substituted for the *Parameter Name* (M1, M2, SD, A, B, C, etc.) in the simulation distributions. If more than one value is entered, a separate calculation is made for each value.

You can enter a single value such as "2" or a series of values such as "0 2 3" or "0 to 3 by 1".

Alpha Simulation

These options specify the distributions to be used in the alpha simulation, one row per group. The alpha simulation generates an estimate of the actual alpha value that will be achieved by the test procedure. The magnitude of the differences of the means of these distributions, which is often summarized as the standard deviation of the means, represents the magnitude of the mean differences specified under H0. Usually, the means are assumed to be equal under H0, so their standard deviation should be zero except for rounding.

Specify Alpha Distributions

This option lets you choose how you will specify the alpha distributions. Possible choices are

- **All equal to the group 1 distribution of the power simulation**

  Often, the first group will represent a neutral group such as the control group. This option indicates that the first group of the power distributions should be used for all of the alpha simulation groups. Typically, this is what you want since the null hypothesis assumes that all distributions are the same.

- **All equal to a custom distribution**

  Set all group distributions in the alpha simulation equal to the custom alpha distribution specified below.

- **Enter each distribution separately**

  Each group's simulation distribution is entered individually below.

Custom Alpha Distribution

Specify the alpha distribution to be used for all groups. The syntax is the same as that for the Power Simulation Distribution (see above) and will not be repeated here.

Individual Alpha Distributions

Specify a separate distribution for each group. The syntax is the same as that for the Power Simulation Distribution (see above) and will not be repeated here.

Keep in mind that the null hypothesis assumes that all distributions are the same.

Simulations Tab

The option on this tab controls the generation of the random numbers. For complicated distributions, a large pool of random numbers from the specified distributions is generated. Each of these pools is evaluated to determine if its mean is within a small relative tolerance (0.0001) of the target mean. If the actual mean is not within the tolerance of the target mean, individual members of the population are replaced with new random numbers if the new random number moves the mean towards its target. Only a few hundred such swaps are required to bring the actual mean to within tolerance of the target mean. This population is then sampled with replacement using the uniform distribution. We have found that this method works well as long as the size of the pool is at least the maximum of twice the number of simulated samples desired and 10,000.

Random Number Pool Size

This is the size of the pool of values from which the random samples will be drawn. Pools should be at least the minimum of 10,000 and twice the number of simulations. You can enter *Automatic* and an appropriate value will be calculated.

If you do not want to draw numbers from a pool, enter 0 here.

Example 1 – Power at Various Sample Sizes

This example repeats Example 1 of four individual procedures that are compared in this procedure.

A one-way design is being employed to compare the means of four groups using an *F* test. Past experimentation has shown that the data within each group can reasonably be assumed to be normally distributed. The researcher would like to compare the four homogeneity tests. Previous studies have shown that the standard deviation within a group is about 5. The researcher wants to investigate the power when the standard deviation of the second group is 6, 8, or 10. Treatment means of 10, 20, 10, and 10 are anticipated. The researcher wants to compute the power for a group sample size of 10. The group sample sizes are equal in a particular scenario. The value of alpha is 0.05.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Power Comparison of Tests of Variances (Simulation)** procedure window by expanding **Variances**, then **Many Variances**, and then clicking on **Power Comparison of Tests of Variances (Simulation)**. You may then make the appropriate entries as listed below or open **Example 1** by going to the **File** menu and choosing **Open Example Template**.

Option	**Value**
Design Tab	
Alpha..	**0.05**
Simulations	**5000**
Number of Groups	**4**
Group Allocation	**Equal**
n (Group Size)	**10**
Group 1 Distribution..............................	**Normal(10 S1)**
Group 2 Distribution..............................	**Normal(20 S2)**
Group 3 Distribution..............................	**Normal(10 S1)**
Group 4 Distribution..............................	**Normal(10 S1)**
S1...	**5**
S2...	**7 8 9**
Specify Alpha Distributions....................	**All equal to the group 1 distribution of the power simulation**

Annotated Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Results Report

Simulation Summary

Group	Power Distribution	Alpha Distribution
1	Normal(10 S1)	Normal(10 S1)
2	Normal(20 S2)	Normal(10 S1)
3	Normal(10 S1)	Normal(10 S1)
4	Normal(10 S1)	Normal(10 S1)

Number of Groups	4
Random Number Pool Size	10000
Number of Simulations Per Run	5000

Power Comparison of Tests of Variances (Simulation)

Power Comparison of Variance Tests

Row	Total Sample Size N	Bart-lett Test Power	Levene Test Power	Brown For-sythe Test Power	Con-over Test Power	Target Alpha	Std Dev of H1 σ's	Mean of H1 σ's	S1	S2
1	40	0.169	0.182	0.101	0.151	0.050	0.9	5.6	5.0	7.0
2	40	0.284	0.269	0.168	0.218	0.050	1.3	5.9	5.0	8.0
3	40	0.412	0.387	0.273	0.312	0.050	1.7	6.2	5.0	9.0

Run Time 3.03 seconds.

Definitions of the Power Comparison of Variance Tests Report
Row identifies this line of the report. It will be used cross-reference this line in other reports.
H0 is the null hypothesis that all σ's are equal.
H1 is the alternative hypothesis that at least one σ is different from the others.
Total Sample Size (N) is the total sample size of all groups combined.
Power is the probability of rejecting H0 when it is false calculated by the power simulation. It is usually
   different from the Target Power because of the discrete nature of the sample size, the violation of test
   assumptions, and the approximate nature of the test.
Bartlett Test is used to test equality of variance in normal populations. It requires normality.
Levene Test is used to test equallity of variances. It does not require the assumption of normality.
Brown-Forsythe Test is used to test equallity of variances. It does not require the assumption of normality.
Conover Test is nonparametric and used to test equallity of variances.
Target Alpha is the planned probability of rejecting a true H0. This is the planned Type-1 error rate.
Std Dev of H1 σ's is the standard deviation of the group standard deviations assumed by the power simulation.
   This is a measure of the effect size.
Mean of H1 σ's is the mean of the within-group standard deviations assumed during the power simulation. This
   measures the variation.

Alpha Comparison of Variance Tests

Row	Total Sample Size N	Bart-lett Test Alpha	Levene Test Alpha	Brown For-sythe Test Alpha	Con-over Test Alpha	Target Alpha	Std Dev of H1 σ's	Mean of H1 σ's	S1	S2
1	40	0.048	0.060	0.031	0.057	0.050	0.0	5.0	5.0	7.0
2	40	0.045	0.066	0.031	0.058	0.050	0.0	5.0	5.0	8.0
3	40	0.044	0.064	0.029	0.058	0.050	0.0	5.0	5.0	9.0

Definitions of the Alpha Comparison of Various Tests Report
Row identifies this line of the report. It will be used cross-reference this line in other reports.
H0 is the null hypothesis that all σ's are equal.
H1 is the alternative hypothesis that at least one σ is different from the others.
Total Sample Size (N) is the total sample size of all groups combined.
Alpha is the actual alpha achieved by the each test as calculated by the alpha simulation. The alpha
   simulation is completely independent of the power simulation.
Bartlett Test is used to test equallity of variance in normal populations. It requires normality.
Levene Test is used to test equallity of variances. It does not require the assumption of normality.
Brown-Forsythe Test is used to test equallity of variances. It does not require the assumption of normality.
Conover Test is nonparametric and used to test equality of variances.
Target Alpha is the planned probability of rejecting a true H0. This is the planned Type-1 error rate.
Std Dev of H0 σ's is the standard deviation of the group standard deviations assumed by the alpha simulation.
   This is a measure of the effect size.
Mean of H0 σ's is the mean of the within-group standard deviations assumed during the alpha simulation. This
   measures the variation.

These reports show the output for this run. The definitions of each column are shown below the report.

Detailed Results Reports

Details for Row 1 of the Numeric Results Report

Group	ni	Pct of N	H0 μ	H1 μ	H0 σ	H1 σ		
1	10	25.0	5.0	5.0	10.0	10.0		
2	10	25.0	5.0	7.0	10.0	20.0		
3	10	25.0	5.0	5.0	10.0	10.0		
4	10	25.0	5.0	5.0	10.0	10.0		
			Std Dev	Std Dev	Std Dev	Std Dev	Mean	Mean
Group	N		of H0 σ's	of H1 σ's	of H0 μ's	of H1 μ's	of H0 σ's	of H1 σ's
All	40	100.0	0.0	0.9	0.0	4.3	5.0	5.5

Details for Row 2 of the Numeric Results Report

Group	ni	Pct of N	H0 σ	H1 σ	H0 μ	H1 μ		
1	10	25.0	5.0	5.0	10.0	10.0		
2	10	25.0	5.0	8.0	10.0	20.0		
3	10	25.0	5.0	5.0	10.0	10.0		
4	10	25.0	5.0	5.0	10.0	10.0		
			Std Dev	Std Dev	Std Dev	Std Dev	Mean	Mean
Group	N		of H0 σ's	of H1 σ's	of H0 μ's	of H1 μ's	of H0 σ's	of H1 σ's
All	40	100.0	0.0	1.3	0.0	4.3	5.0	5.8

These reports show the details of each scenario.

Row (in Title)

This is the row number of the Numeric Results report about which this report gives the details.

Group

This is the number of the group shown on this line.

ni

This is the sample size of each group. This column is especially useful when the sample sizes are unequal.

Percent of N

This is the percentage of the total sample that is allocated to each group.

H0 μ and H1 μ

These are the means that were used in the alpha and power simulations, respectively.

H0 σ and H1 σ

These are the standard deviations that were obtained by the alpha and power simulations, respectively. Note that they often are not exactly equal to what was specified because of the error introduced by simulation.

### Std Dev of H0 (and H1) μ's

These are the standard deviations of the means that were obtained by the alpha and power simulations, respectively. Under H0, this value should be near zero. The H0 value lets you determine if your alpha simulation was correctly specified. The H1 value represents the magnitude of the effect size (when divided by an appropriate measure of the standard deviation).

Mean of H0 (and H1) σ's

These are the average of the individual group standard deviations that were obtained by the alpha and power simulations, respectively.

Plots Section
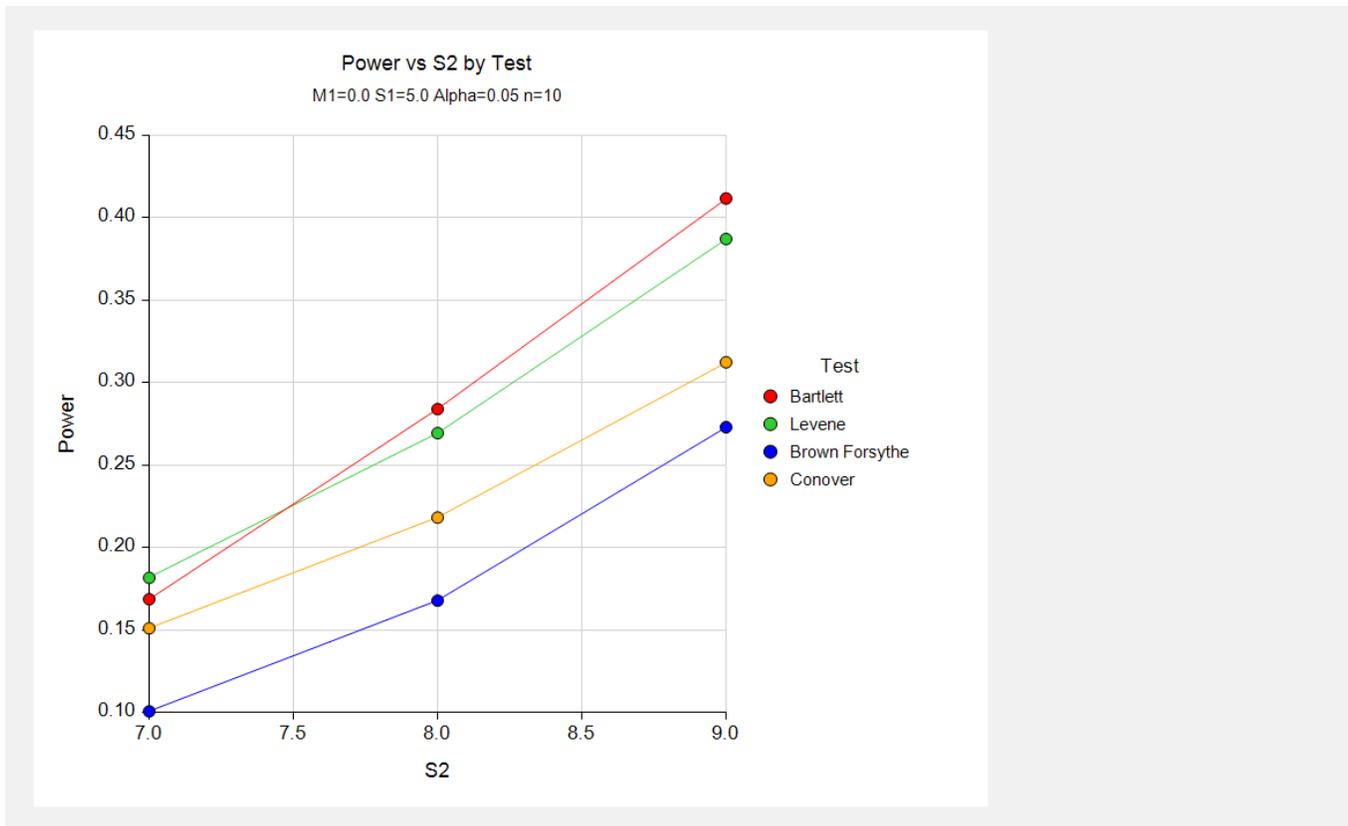
This plot gives a visual presentation to the results in the Numeric Report. We can quickly see the impact on the power of increasing the standard deviation. We can also see that Bartlett's test is the most powerful when the normality assumption is met, but that the Levene test is a close second.

Example 2 – Validation

To validate this procedure, we will compare the results to an example run in the *Tests for Two Variances* procedure. For this example, alpha was set to 0.05, the Scale was Standard Deviation, V1 was 1, and V2 was 2. The desired power was 0.8. The resulting sample size as N1 = N2 = 19.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Power Comparison of Tests of Variances (Simulation)** procedure window by expanding **Variances**, then **Many Variances**, and then clicking on **Power Comparison of Tests of Variances (Simulation)**. You may then make the appropriate entries as listed below or open **Example 2** by going to the **File** menu and choosing **Open Example Template**.

Option   **Value**

Design Tab

Alpha **0.05**
Simulations **5000**
Number of Groups **2**
Group Allocation .... **Equal**
n (Group Size) **19**
Group 1 Distribution .... **Normal(1 S1)**
Group 2 Distribution .... **Normal(1 S2)**
S1 .... **1**
S2 .... **2**
Specify Alpha Distributions .... **All equal to the group 1 distribution of the power simulation**

Reports Tab

Show Comparative Reports .... **Checked**
Show Detail Reports .... **Checked**
All other reports .... **Unchecked**

Comparative Plots Tab

All plots .... **Unchecked**

Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Results

Power Comparison of Variance Tests

Row	Total Sample Size N	Bart- lett Test Power	Levene Test Power	Brown For- sythe Test Power	Con- over Test Power	Target Alpha	Std Dev of H1 σ's	Mean of H1 σ's	S1	S2
1	38	0.822	0.738	0.674	0.695	0.050	0.5	1.6	1.0	2.0

Alpha Comparison of Variance Tests

Row	Total Sample Size N	Bart- lett Test Alpha	Levene Test Alpha	Brown For- sythe Test Alpha	Con- over Test Alpha	Target Alpha	Std Dev of H1 σ's	Mean of H1 σ's	S1	S2
1	38	0.047	0.057	0.036	0.056	0.050	0.0	1.0	1.0	2.0

Note that the power of Bartlett's test is calculated at just a little over 0.800. The powers of the other tests are about what we would suppose for tests designed to perform well with non-normal data.