

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 alternative 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 null 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.

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

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 1** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab

Alpha.....	0.05
Simulations	5000
Random Seed.....	5935245 (for Reproducibility)
Number of Groups	4
Group Allocation	Equal (n = n1 = n2 = ...)
n (Group Size)	10
Group 1 Power Distribution.....	Normal(10 S1)
Group 2 Power Distribution.....	Normal(20 S2)
Group 3 Power Distribution.....	Normal(10 S1)
Group 4 Power Distribution.....	Normal(10 S1)
Parameter 2 Name	S1
Parameter 2 Value(s).....	5
Parameter 3 Name	S2
Parameter 3 Value(s).....	7 8 9
Specify Alpha Distributions	All equal to the group 1 distribution of the power simulation

Power Comparison of Tests of Variances (Simulation)

Output

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

Simulation Summary and Comparison Results

Simulation Summary

Group	Simulation Distribution	
	Power (H1)	Alpha (H0)
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	5000
Random Seed	5935245 (User-Entered)
Run Time	1.81 seconds

Power Comparison of Variance Tests

Scenario	Total Sample Size N	Power					Target Alpha	Standard Deviation of σ 's H1	Average of σ 's H1	S1	S2
		Bartlett Test	Levene Test	Brown-Forsythe Test	Conover Test						
1	40	0.153	0.173	0.100	0.145		0.05	0.9	5.5	5	7
2	40	0.294	0.289	0.187	0.238		0.05	1.3	5.9	5	8
3	40	0.398	0.370	0.252	0.293		0.05	1.6	6.2	5	9

Row	Identifies this line of the report. It will be used cross-reference this line in other reports.
H0	The null hypothesis that all σ 's are equal.
H1	The alternative hypothesis that at least one σ is different from the others.
N	The total sample size of all groups combined.
Power	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	Used to test equality of variance in normal populations. It requires normality.
Levene Test	Used to test equality of variances. It does not require the assumption of normality.
Brown-Forsythe Test	Used to test equality of variances. It does not require the assumption of normality.
Conover Test	Nonparametric test used to test equality of variances.
Target Alpha	The planned probability of rejecting a true H0. This is the planned Type-1 error rate.
SD of σ 's H1	The standard deviation of the group standard deviations assumed by the power simulation. This is a measure of the effect size.
Average of σ 's H1	The average of the within-group standard deviations assumed during the power simulation. This measures the variation.
M1, S1, S2, etc.	The value(s) of the user-specified distribution parameters.

Power Comparison of Tests of Variances (Simulation)

Alpha Comparison of Variance Tests

Scenario	Total Sample Size N	Alpha				Target	Standard Deviation of σ 's H0	Average of σ 's H0	S1	S2
		Bartlett Test	Levene Test	Brown-Forsythe Test	Conover Test					
1	40	0.052	0.059	0.029	0.053	0.05	0	5	5	7
2	40	0.053	0.071	0.031	0.059	0.05	0	5	5	8
3	40	0.051	0.070	0.031	0.062	0.05	0	5	5	9
Row	Identifies this line of the report. It will be used cross-reference this line in other reports.									
H0	The null hypothesis that all σ 's are equal.									
H1	The alternative hypothesis that at least one σ is different from the others.									
N	The total sample size of all groups combined.									
Alpha	The actual alpha achieved by each test as calculated by the alpha simulation. The alpha simulation is completely independent of the power simulation.									
Bartlett Test	Used to test equality of variance in normal populations. It requires normality.									
Levene Test	Used to test equality of variances. It does not require the assumption of normality.									
Brown-Forsythe Test	Used to test equality of variances. It does not require the assumption of normality.									
Conover Test	Nonparametric test used to test equality of variances.									
Target Alpha	The planned probability of rejecting a true H0. This is the planned Type-1 error rate.									
SD of σ 's H0	The standard deviation of the group standard deviations assumed by the alpha simulation. This is a measure of the effect size.									
Average of σ 's H0	The average of the within-group standard deviations assumed during the alpha simulation. This measures the variation.									
M1, S1, S2, etc.	The value(s) of the user-specified distribution parameters.									
.										
.										
.										

References

Fleiss, Joseph L. 1986. The Design and Analysis of Clinical Experiments. John Wiley & Sons. New York.
 Gibbons, J.D. and Chakraborti, S. 2011. Nonparametric Statistical Inference, 5th Edition. CRC Press.
 Devroye, Luc. 1986. Non-Uniform Random Variate Generation. Springer-Verlag. New York.
 Matsumoto, M. and Nishimura, T. 1998. 'Mersenne twister: A 623-dimensionally equidistributed uniform pseudorandom number generator.' ACM Trans. On Modeling and Computer Simulations.

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

Detailed Results Reports

Details for Scenario 1 of the Numeric Results

Group	ni	Percent of N	σ H0	σ H1	μ H0	μ H1		
1	10	25	5	5.0	10	10.0		
2	10	25	5	7.0	10	20.0		
3	10	25	5	4.9	10	10.0		
4	10	25	5	4.9	10	10.0		
Group	N		SD of σ 's H0	SD of σ 's H1	SD of μ 's H0	SD of μ 's H1	Average of σ 's H0	Average of σ 's H1
All	40	100	0	0.9	0	4.3	5	5.5

(More Reports Follow)

These reports show the details of each scenario.

Scenario (in Title)

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

Power Comparison of Tests of Variances (Simulation)

Group

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

 n_i

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.

 σ | H0 and σ | H1

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.

 μ | H0 and μ | H1

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

SD of σ 's | H0, H1

These are the standard deviations of the group standard deviations 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.

SD of μ 's | H0, H1

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).

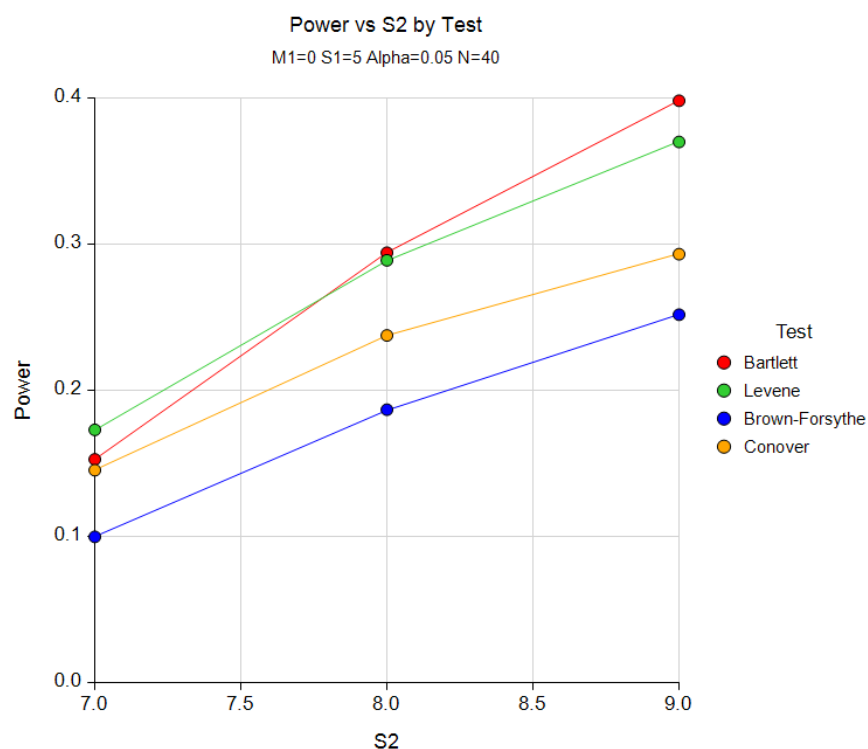
Average of σ 's | H0, H1

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

Power Comparison of Tests of Variances (Simulation)

Plots Section

Plots



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 $N1 = N2 = 19$.

For reproducibility, we'll use a random seed of 5744621.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 2** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab

Alpha.....	0.05
Simulations	5000
Random Seed	5744621 (for Reproducibility)
Number of Groups	2
Group Allocation	Equal (n = n1 = n2 = ...)
n (Group Size)	19
Group 1 Power Distribution.....	Normal(1 S1)
Group 2 Power Distribution.....	Normal(1 S2)
Parameter 2 Name	S1
Parameter 2 Value(s).....	1
Parameter 3 Name	S2
Parameter 3 Value(s).....	2
Specify Alpha Distributions	All equal to the group 1 distribution of the power simulation

Output

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

Simulation Summary

Group	Simulation Distribution	
	Power (H1)	Alpha (H0)
1	Normal(1 S1)	Normal(1 S1)
2	Normal(1 S2)	Normal(1 S1)

Number of Groups	2
Random Number Pool Size	10000
Number of Simulations	5000
Random Seed	5744621 (User-Entered)
Run Time	0.63 seconds

Power Comparison of Tests of Variances (Simulation)

Power Comparison of Variance Tests

Scenario	Total Sample Size N	Power				Target Alpha	Standard Deviation of σ 's H1	Average of σ 's H1	S1	S2
		Bartlett Test	Levene Test	Brown- Forsythe Test	Conover Test					
1	38	0.813	0.729	0.667	0.685	0.05	0.5	1.6	1	2

Alpha Comparison of Variance Tests

Scenario	Total Sample Size N	Alpha				Target	Standard Deviation of σ 's H0	Average of σ 's H0	S1	S2
		Bartlett Test	Levene Test	Brown- Forsythe Test	Conover Test					
1	38	0.054	0.058	0.037	0.055	0.05	0	1	1	2

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.