Chapter 657

Confidence Intervals for the Ratio of Two Variances using Relative Error

Introduction

This routine calculates the necessary sample size such that a variance ratio estimate from two independent samples will achieve a specified relative distance from the true variance ratio at a stated confidence level when the underlying data distributions are normal.

Caution: This procedure controls the relative width of the interval as a proportion of the true variance ratio. For controlling the absolute width of the interval see the procedure Confidence Intervals for the Ratio of Two Variances using Variances.

Technical Details

Following the results of Desu and Raghavarao (1990), let s_1^2 be the variance estimate based on a sample from a normal distribution with unknown μ_1 and unknown σ_1^2. Let s_2^2 be the variance estimate based on a sample from a normal distribution with unknown μ_2 and unknown σ_2^2. Let r be the proportion of σ_1^2/σ_2^2 such that s_1^2/s_2^2 is within $r\sigma_1^2/\sigma_2^2$ of σ_1^2/σ_2^2 with desired confidence $(1 - \alpha)$. That is, the desired condition to be satisfied is

$$\Pr\left(\left|\frac{s_1^2}{s_2^2} - \frac{\sigma_1^2}{\sigma_2^2}\right| \le r\frac{\sigma_1^2}{\sigma_2^2}\right) \ge 1 - \alpha$$

which can also be written as

$$\Pr\left(\left|\frac{\frac{s_1^2}{s_2^2} - \frac{\sigma_1^2}{\sigma_2^2}}{\frac{\sigma_1^2}{\sigma_2^2}}\right| \le r\right) \ge 1 - \alpha$$

which simplifies to

$$\Pr\left(1 - r \le F(df_1, df_2) \le 1 + r\right) \ge 1 - \alpha$$

since

$$\frac{s_1^2}{s_2^2} \bigg/ \frac{\sigma_1^2}{\sigma_2^2} \sim F(df_1, df_2)$$

If $G_{df_1, df}(.)$ is the distribution function of $F(df_1, df_2)$, then probability statement can be rewritten as

$$G_{df_1, df_2}(1 + r) - G_{df_1, df_2}(1 - r) \geq 1 - \alpha$$

This equation can be solved for any of the unknown quantities (df_1, df_2, r, α) in terms of the others.

Confidence Level

The confidence level, $1 - \alpha$, has the following interpretation. If thousands of random samples of size n_1 and n_2 are drawn from populations 1 and 2, respectively, and a variance ratio is calculated for each pair of samples, the proportion of those estimates that are within $r\sigma_1^2/\sigma_2^2$ of σ_1^2/σ_2^2 is $1 - \alpha$.

Procedure Options

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

Design Tab

The Design tab contains most of the parameters and options that you will be concerned with.

Solve For

Solve For

This option specifies the parameter to be solved for from the other parameters.

One-Sided or Two-Sided Interval

Interval Type

Specify whether the interval to be used will be a two-sided confidence interval, an interval that has only an upper limit, or an interval that has only a lower limit.

In each case the limits are based on the relative error of the true population variance.

Confidence

Confidence Level

The confidence level, $1 - \alpha$, has the following interpretation. If thousands of random samples of size n_1 and n_2 are drawn from populations 1 and 2, respectively, and a variance ratio is calculated for each pair of samples, the proportion of those estimates that are within $r\sigma_1^2/\sigma_2^2$ of σ_1^2/σ_2^2 is $1 - \alpha$.

Often, the values 0.95 or 0.99 are used. You can enter single values or a range of values such as *0.90, 0.95* or *0.90 to 0.99 by 0.01*.

Sample Size (When Solving for Sample Size)

Group Allocation

Select the option that describes the constraints on *N1* or *N2* or both.

The options are

- **Equal (N1 = N2)**

  This selection is used when you wish to have equal sample sizes in each group. Since you are solving for both sample sizes at once, no additional sample size parameters need to be entered.

- **Enter N1, solve for N2**

  Select this option when you wish to fix *N1* at some value (or values), and then solve only for *N2*. Please note that for some values of *N1*, there may not be a value of *N2* that is large enough to obtain the desired power.

- **Enter N2, solve for N1**

  Select this option when you wish to fix *N2* at some value (or values), and then solve only for *N1*. Please note that for some values of *N2*, there may not be a value of *N1* that is large enough to obtain the desired power.

- **Enter R = N2/N1, solve for N1 and N2**

  For this choice, you set a value for the ratio of *N2* to *N1*, and then PASS determines the needed *N1* and *N2*, with this ratio, to obtain the desired power. An equivalent representation of the ratio, *R*, is

  $$N2 = R * N1.$$

- **Enter percentage in Group 1, solve for N1 and N2**

  For this choice, you set a value for the percentage of the total sample size that is in Group 1, and then PASS determines the needed *N1* and *N2* with this percentage to obtain the desired power.

N1 (Sample Size, Group 1)

This option is displayed if Group Allocation = "Enter N1, solve for N2"

N1 is the number of items or individuals sampled from the Group 1 population.

N1 must be ≥ 2. You can enter a single value or a series of values.

N2 (Sample Size, Group 2)

This option is displayed if Group Allocation = "Enter N2, solve for N1"

N2 is the number of items or individuals sampled from the Group 2 population.

N2 must be ≥ 2. You can enter a single value or a series of values.

R (Group Sample Size Ratio)

This option is displayed only if Group Allocation = "Enter R = N2/N1, solve for N1 and N2."

R is the ratio of *N2* to *N1*. That is,

$$R = N2 / N1.$$

Use this value to fix the ratio of *N2* to *N1* while solving for *N1* and *N2*. Only sample size combinations with this ratio are considered.

N2 is related to *N1* by the formula:

$$N2 = [R \times N1],$$

where the value *[Y]* is the next integer $\geq Y$.

For example, setting $R = 2.0$ results in a Group 2 sample size that is double the sample size in Group 1 (e.g., $N1 = 10$ and $N2 = 20$, or $N1 = 50$ and $N2 = 100$).

R must be greater than 0. If $R < 1$, then $N2$ will be less than $N1$; if $R > 1$, then $N2$ will be greater than $N1$. You can enter a single or a series of values.

Percent in Group 1

This option is displayed only if Group Allocation = "Enter percentage in Group 1, solve for N1 and N2."

Use this value to fix the percentage of the total sample size allocated to Group 1 while solving for $N1$ and $N2$. Only sample size combinations with this Group 1 percentage are considered. Small variations from the specified percentage may occur due to the discrete nature of sample sizes.

The Percent in Group 1 must be greater than 0 and less than 100. You can enter a single or a series of values.

Sample Size (When <u>Not</u> Solving for Sample Size)

Group Allocation

Select the option that describes how individuals in the study will be allocated to Group 1 and to Group 2.

The options are

- **Equal (N1 = N2)**

  This selection is used when you wish to have equal sample sizes in each group. A single per group sample size will be entered.

- **Enter N1 and N2 individually**

  This choice permits you to enter different values for $N1$ and $N2$.

- **Enter N1 and R, where N2 = R * N1**

  Choose this option to specify a value (or values) for $N1$, and obtain $N2$ as a ratio (multiple) of $N1$.

- **Enter total sample size and percentage in Group 1**

  Choose this option to specify a value (or values) for the total sample size (N), obtain $N1$ as a percentage of N, and then $N2$ as $N - N1$.

Sample Size Per Group

This option is displayed only if Group Allocation = "Equal (N1 = N2)."

The Sample Size Per Group is the number of items or individuals sampled from each of the Group 1 and Group 2 populations. Since the sample sizes are the same in each group, this value is the value for $N1$, and also the value for $N2$.

The Sample Size Per Group must be ≥ 2. You can enter a single value or a series of values.

N1 (Sample Size, Group 1)

*This option is displayed if Group Allocation = "Enter N1 and N2 individually" or "Enter N1 and R, where N2 = R * N1."*

$N1$ is the number of items or individuals sampled from the Group 1 population.

$N1$ must be ≥ 2. You can enter a single value or a series of values.

N2 (Sample Size, Group 2)

This option is displayed only if Group Allocation = "Enter N1 and N2 individually."

N2 is the number of items or individuals sampled from the Group 2 population.

N2 must be $\geq$ 2. You can enter a single value or a series of values.

R (Group Sample Size Ratio)

*This option is displayed only if Group Allocation = "Enter N1 and R, where N2 = R * N1."*

R is the ratio of *N2* to *N1*. That is,

$$R = N2/N1$$

Use this value to obtain *N2* as a multiple (or proportion) of *N1*.

N2 is calculated from *N1* using the formula:

$$N2=[R \times N1],$$

where the value *[Y]* is the next integer $\geq$ *Y*.

For example, setting *R* = 2.0 results in a Group 2 sample size that is double the sample size in Group 1.

R must be greater than 0. If *R* < 1, then *N2* will be less than N1; if *R* > 1, then *N2* will be greater than *N1*. You can enter a single value or a series of values.

Total Sample Size (N)

This option is displayed only if Group Allocation = "Enter total sample size and percentage in Group 1."

This is the total sample size, or the sum of the two group sample sizes. This value, along with the percentage of the total sample size in Group 1, implicitly defines *N1* and *N2*.

The total sample size must be greater than one, but practically, must be greater than 3, since each group sample size needs to be at least 2.

You can enter a single value or a series of values.

Percent in Group 1

This option is displayed only if Group Allocation = "Enter total sample size and percentage in Group 1."

This value fixes the percentage of the total sample size allocated to Group 1. Small variations from the specified percentage may occur due to the discrete nature of sample sizes.

The Percent in Group 1 must be greater than 0 and less than 100. You can enter a single value or a series of values.

Precision

Relative Error

This is the distance from the true variance ratio as a proportion of the true variance ratio.

You can enter a single value or a list of values. The value(s) must be between 0 and 1.

Example 1 – Calculating Sample Size

Suppose a study is planned in which the researcher wishes to be 95% confident that estimated variance ratio is within 10% of the true variance ratio. In addition to 10% relative error, 5%, 15%, 20% and 25% will also be considered.

The goal is to determine the necessary sample size.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Confidence Intervals for the Ratio of Two Variances using Relative Error** procedure window by expanding **Variances**, then clicking on **Two Variances**, and then clicking on **Confidence Intervals for the Ratio of Two Variances using Relative Error**. 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
Solve For .... **Sample Size**
Interval Type .... **Two-Sided**
Confidence Level **0.95**
Group Allocation .... **Equal (N1 = N2)**
Relative Error .... **0.05 to 0.25 by 0.05**

Annotated Output

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

Numeric Results

Numeric Results for Two-Sided Relative Error Confidence Intervals for the Variance Ratio

Target Confidence Level	Actual Confidence Level	N1	N2	N	Relative Error
0.950	0.950	6154	6154	12308	0.050
0.950	0.950	1544	1544	3088	0.100
0.950	0.950	690	690	1380	0.150
0.950	0.950	392	392	784	0.200
0.950	0.950	253	253	506	0.250

References
Desu, M. M. and Raghavarao, D. 1990. Sample Size Methodology. Academic Press. New York.

Report Definitions
Target Confidence Level is the value of the confidence level that is entered into the procedure.
Actual Confidence Level is the value of the confidence level that is obtained from the procedure.
N1 and N2 are the number of items sampled from each population.
N is the total sample size, N1 + N2.
Relative Error is the distance from the true variance ratio as a proportion of the true variance ratio.
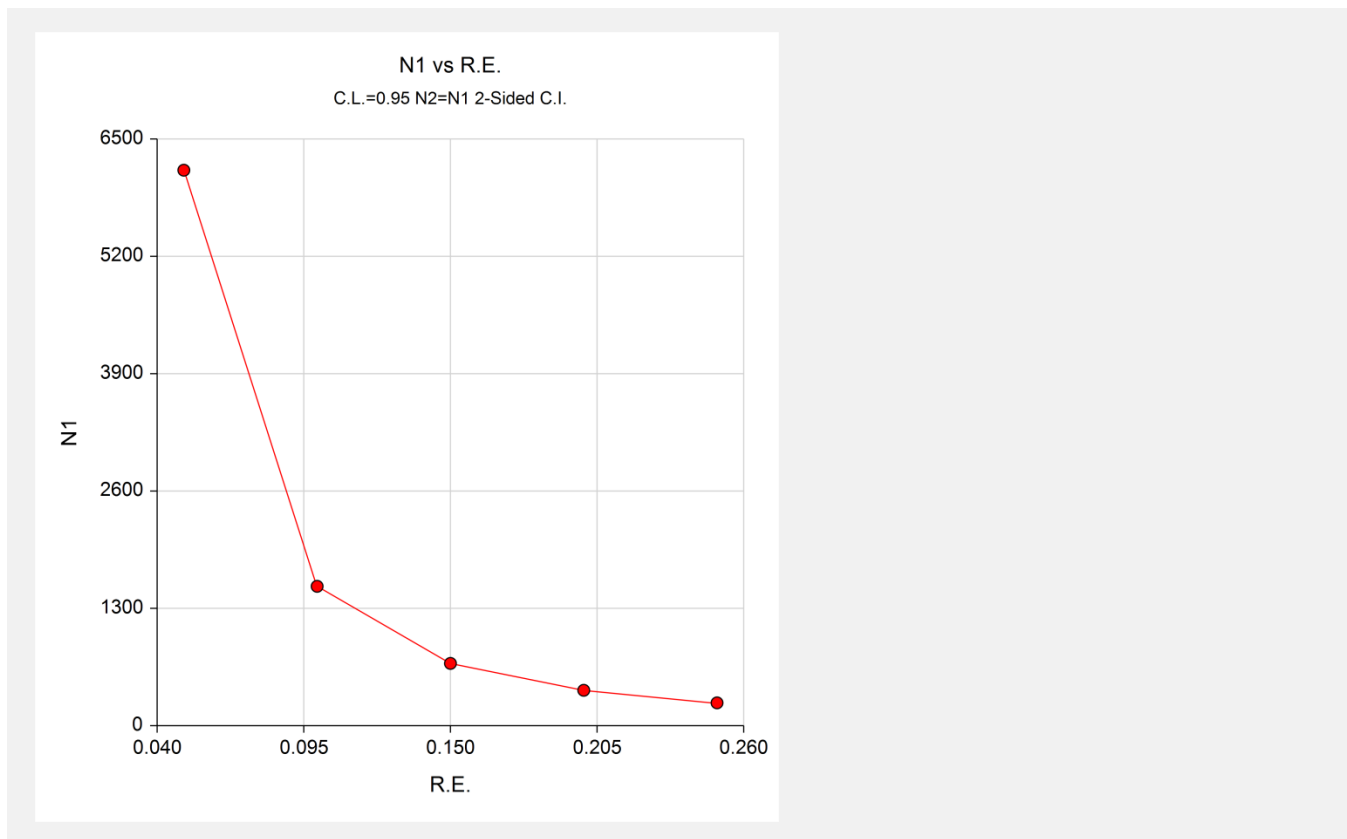
Summary Statements
With group sample sizes of 6154 and 6154, the probability is 0.950 (95% confidence) that the estimate of the variance ratio will be within 5% of the true variance ratio.

This report shows the calculated sample size for each of the scenarios.

Plots Section

This plot shows the sample size versus the relative error.

Example 2 – Validation using Direct Calculation

Suppose a study is planned in which the researcher wishes to be confident that estimated variance ratio is within 20% of the true variance ratio. For sample sizes of 200 per group, the resulting confidence level is 0.9003379878 - 0.0581710981 = 0.8421668897.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Confidence Intervals for the Ratio of Two Variances using Relative Error** procedure window by expanding **Variances**, then clicking on **Two Variances**, and then clicking on **Confidence Intervals for the Ratio of Two Variances using Relative Error**. 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	
Solve For	**Sample Size**
Interval Type	**Two-Sided**
Confidence Level	**0.8421668897**
Group Allocation	**Equal (N1 = N2)**
Relative Error	**0.2**

Output

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

Numeric Results

Target Confidence Level	Actual Confidence Level	N1	N2	N	Relative Error
0.842	0.842	200	200	400	0.200

PASS calculates the necessary sample size to be 200 per group, which is the same as the direct calculation sample size.