Chapter 805

Tests for Two Correlations

Introduction

The correlation coefficient (or correlation), ρ, is a popular parameter for describing the strength of the association between two variables. The correlation coefficient is the slope of the regression line between two variables when both variables have been standardized. It ranges between plus and minus one. This chapter covers the case in which you want to test the difference between two correlations, each coming from a separate sample.

Since the correlation is the standardized slope between two variables, you could also apply this procedure to the case in which you want to test whether the slopes in two groups are equal.

Test Procedure

In the following discussion, ρ is the population correlation coefficient and r is the value calculated from a sample. The testing procedure is as follows. H_0 is the null hypothesis that $\rho_1 = \rho_2$. H_A represents the alternative hypothesis that $\rho_1 \neq \rho_2$ (one-tailed hypotheses are also available). To construct the hypothesis test, transform the correlations using the Fisher-z transformation.

$$z_i = \frac{1}{2}\log\left(\frac{1+r_i}{1-r_i}\right)$$

$$Z_i = \frac{1}{2}\log\left(\frac{1+\rho_i}{1-\rho_i}\right)$$

This transformation is used because the combined distribution of r_1 and r_2 is too difficult to work with, but the distributions of z_1 and z_2 are approximately normal.

Note that the reverse transformation is

$$r_i = \frac{e^{z_i} - e^{-z_i}}{e^{z_i} + e^{-z_i}}$$

Once the correlations have been converted into z values, the normal distribution may be used to conduct the test of $Z_1 - Z_2$. The standard deviation of the difference is given by

$$\sigma_{z_1-z_2} = \sqrt{\frac{1}{N_1-3} + \frac{1}{N_2-3}}$$

The test statistic is given by

$$z = \frac{(z_1 - z_2) - (Z_1 - Z_2)}{\sigma_{z_1 - z_2}}$$

Note that the lower case $z's$ represent the values calculated from the two samples and the upper case $Z's$ represent the hypothesized population values.

Calculating the Power

1. Find z_α such that $1 - \Phi(z_\alpha) = \alpha$, where $\Phi(x)$ is the area under the standardized normal curve to the left of x.

2. Calculate: $Z_1 = \frac{1}{2}\log\left(\frac{1 + \rho_1}{1 - \rho_1}\right)$

3. Calculate: $Z_2 = \frac{1}{2}\log\left(\frac{1 + \rho_2}{1 - \rho_2}\right)$

4. Calculate: $\sigma_{z_1 - z_2} = \sqrt{\frac{1}{N_1 - 3} + \frac{1}{N_2 - 3}}$

5. Calculate: $x_a = Z_1 - Z_2 + z_\alpha \sigma_{z_1 - z_2}$

6. Calculate: $z_a = \frac{x_a}{\sigma_{z_1 - z_2}}$

7. Calculate: $\text{Power} = 1 - \Phi(z_a)$

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. Under most situations, you will select either *Power* or *Sample Size (N1)*.

Select *Sample Size (N1)* when you want to calculate the sample size needed to achieve a given power and alpha level.

Select *Power* when you want to calculate the power of an experiment.

Test

Alternative Hypothesis

This option specifies the alternative hypothesis. This implicitly specifies the direction of the hypothesis test. The null hypothesis is always H0: $\rho1 = \rho2$.

Possible selections are:

- **Ha: $\rho1 \neq \rho2$**

  This is the most common selection. It yields the *two-tailed* test. Use this option when you are testing whether the correlation values are different, but you do not want to specify beforehand which value is larger.

- **Ha: $\rho1 < \rho2$**

  This option yields a *one-tailed* test. When you use this option, you should be careful to enter values for *ρ1* and *ρ2* that follow this relationship.

- **Ha: $\rho1 > \rho2$**

  This option yields a *one-tailed* test. When you use this option, you should be careful to enter values for *ρ1* and *ρ2* that follow this relationship.

Power and Alpha

Power

This option specifies one or more values for power. Power is the probability of rejecting a false null hypothesis, and is equal to one minus Beta. Beta is the probability of a type-II error, which occurs when a false null hypothesis is not rejected. In this procedure, a type-II error occurs when you fail to reject the null hypothesis of equal correlations when in fact they are different.

Values must be between zero and one. Historically, the value of 0.80 (Beta = 0.20) was used for power. Now, 0.90 (Beta = 0.10) is also commonly used.

A single value may be entered here or a range of values such as *0.8 to 0.95 by 0.05* may be entered.

Alpha

This option specifies one or more values for the probability of a type-I error. A type-I error occurs when you reject the null hypothesis of equal correlations when in fact they are equal.

Values of alpha must be between zero and one. Historically, the value of 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 range of values such as *0.01 0.05 0.10* or *0.01 to 0.10 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 ≥ 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.

Effect Size

ρ1 (Correlation Group 1)

Specify the value of the population correlation coefficient of group one. Possible values range between plus and minus one.

You can enter a single value or a range of values separated by commas or blanks.

Note that the power depends on the specific values of ρ_1 and ρ_2, not just their difference. Hence, $\rho_1 = 0$ and $\rho_2 = 0.3$ will have a different power from $\rho_1 = 0.3$ and $\rho_2 = 0.6$.

ρ2 (Correlation Group 2)

Specify the value of the population correlation coefficient from group two under the alternative hypothesis. Possible values range between plus and minus one.

You can enter a single value or a range of values separated by commas or blanks.

Example 1 – Finding the Power

A researcher wants to compare the relationship between weight and heart rate in males and females. If the correlation between weight and heart rate is 0.3 in a sample of 100 males and 0.5 in a sample of 100 females, what is the power of a two sided test for the difference between correlations at the 0.01 and 0.05 significance levels? Also compute the power for samples of 20, 200, 300, 400, and 600.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Tests for Two Correlations** procedure window by expanding **Correlation**, then **Correlation**, then clicking on **Test (Inequality)**, and then clicking on **Tests for Two Correlations**. 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 .... **Power**
Alternative Hypothesis .... **Ha: ρ1 ≠ ρ2**
Alpha **0.01 0.05**
Group Allocation .... **Equal (N1 = N2)**
Sample Size Per Group **20 100 200 300 400 600**
ρ1 (Correlation Group 1) **0.3**
ρ2 (Correlation Group 2) **0.5**

Annotated Output

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

Numeric Results

Numeric Results when Ha: ρ1 ≠ ρ2

Power	N1	N2	N	ρ1	ρ2	ρ1 - ρ2	Alpha
0.03081	20	20	40	0.30000	0.50000	-0.20000	0.010
0.18250	100	100	200	0.30000	0.50000	-0.20000	0.010
0.42230	200	200	400	0.30000	0.50000	-0.20000	0.010
0.63541	300	300	600	0.30000	0.50000	-0.20000	0.010
0.78888	400	400	800	0.30000	0.50000	-0.20000	0.010
0.94144	600	600	1200	0.30000	0.50000	-0.20000	0.010
0.10760	20	20	40	0.30000	0.50000	-0.20000	0.050
0.38603	100	100	200	0.30000	0.50000	-0.20000	0.050
0.66271	200	200	400	0.30000	0.50000	-0.20000	0.050
0.83200	300	300	600	0.30000	0.50000	-0.20000	0.050
0.92196	400	400	800	0.30000	0.50000	-0.20000	0.050
0.98548	600	600	1200	0.30000	0.50000	-0.20000	0.050

Report Definitions
Power is the probability of rejecting a false null hypothesis.
N1 and N2 are the number of items sampled from each population.
N is the total sample size, N1 + N2.
ρ1 is the value of both correlations under the null hypothesis.
ρ2 is the correlation in group two under the alternative hypothesis.
ρ1 - ρ2 is the difference between population correlations at which power and sample size calculations are
   made.
Alpha is the probability of rejecting a true null hypothesis.
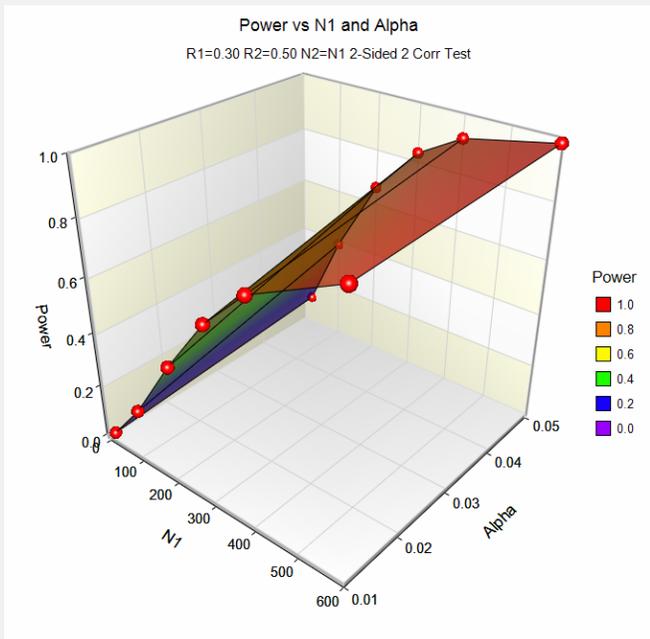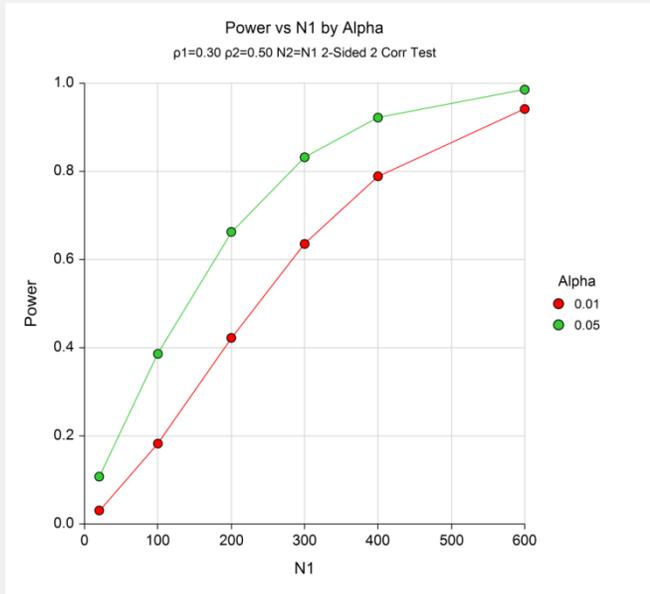
Tests for Two Correlations

This report shows the values of each of the parameters, one scenario per row. The definitions of each column are given in the Report Definitions section of the report, so they will not be repeated here.

The values from this table are plotted in the chart below.

Plots Section

These plots show the relationship between alpha, power, and sample size in this example.

Example 2 – Finding the Sample Size

Continuing with the previous example, suppose the researchers want to determine the exact sample size necessary to achieve 90% power at a 0.05 significance level.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Tests for Two Correlations** procedure window by clicking on **Correlation**, then **Tests for Two Correlations**. 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**
Alternative Hypothesis	**Ha: ρ1 ≠ ρ2**
Power	**0.90**
Alpha	**0.05**
Group Allocation	**Equal (N1 = N2)**
ρ1 (Correlation Group 1)	**0.3**
ρ2 (Correlation Group 2)	**0.5**

Output

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

Numeric Results

Numeric Results when Ha: ρ1 ≠ ρ2

Target Power	Actual Power	N1	N2	N	ρ1	ρ2	ρ1 - ρ2	Alpha
0.90	0.90040	369	369	738	0.30000	0.50000	-0.20000	0.050

PASS has calculated the sample size as 369 per group.

Example 3 – Validation using Zar

Zar (1984) page 314 presents an example of calculating the power for a test of two correlations. In his example, when $N1 = 95$, $N2 = 98$, $\rho1 = 0.84$, $\rho2 = 0.78$, and *alpha* = 0.05, the power is 22% for a two-sided test.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Tests for Two Correlations** procedure window by clicking on **Correlation**, then **Tests for Two Correlations**. 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	**Power**
Alternative Hypothesis	**Ha: ρ1 ≠ ρ2**
Alpha...	**0.05**
Group Allocation	**Enter N1 and N2 individually**
N1	**95**
N2	**98**
ρ1 (Correlation Group 1).........................	**0.84**
ρ2 (Correlation Group 2).........................	**0.78**

Output

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

Numeric Results

Numeric Results when Ha: ρ1 ≠ ρ2

Power	N1	N2	N	ρ1	ρ2	ρ1 - ρ2	Alpha
0.22498	95	98	193	0.84000	0.78000	0.06000	0.050

PASS also calculates the power to be 22%.