

**power onecorrelation** — Power analysis for a one-sample correlation test

<a href="#">Description</a>	<a href="#">Quick start</a>	<a href="#">Menu</a>	<a href="#">Syntax</a>
<a href="#">Options</a>	<a href="#">Remarks and examples</a>	<a href="#">Stored results</a>	<a href="#">Methods and formulas</a>
<a href="#">References</a>	<a href="#">Also see</a>		

## Description

`power onecorrelation` computes sample size, power, or target correlation for a one-sample correlation test. By default, it computes sample size for given power and the values of the correlation parameters under the null and alternative hypotheses. Alternatively, it can compute power for given sample size and values of the null and alternative correlations or the target correlation for given sample size, power, and the null correlation. Also see [\[PSS\] power](#) for a general introduction to the `power` command using hypothesis tests.

## Quick start

Sample size for a test of  $H_0: \rho = 0$  versus  $H_a: \rho \neq 0$  with null correlation  $r_0 = 0$ , alternative correlation  $r_a = 0.3$ , and default power of 0.8 and significance level  $\alpha = 0.05$

```
power onecorrelation 0 .3
```

As above, for alternative correlations of 0.2, 0.25, 0.3, 0.35, and 0.4

```
power onecorrelation 0 (.2(.05).4)
```

As above, but display a graph showing sample size versus alternative correlation

```
power onecorrelation 0 (.2(.05).4), graph
```

Sample size for  $r_0 = 0.1$ ,  $r_a = 0.3$  with power of 0.85 and  $\alpha = 0.01$

```
power onecorrelation .1 .3, power(.85) alpha(.01)
```

Power for a sample size of 30

```
power onecorrelation 0 .5, n(30)
```

Effect size and target correlation for a sample size of 20 and power of 0.8

```
power onecorrelation 0, n(20) power(.8)
```

As above, for sample sizes of 20, 30, 40, and 50

```
power onecorrelation 0, n(20(10)50) power(.8)
```

## Menu

Statistics > Power and sample size

### Syntax

*Compute sample size*

```
power onecorrelation  $r_0$   $r_a$  [ , power(numlist) options ]
```

*Compute power*

```
power onecorrelation  $r_0$   $r_a$  , n(numlist) [ options ]
```

*Compute effect size and target correlation*

```
power onecorrelation  $r_0$  , n(numlist) power(numlist) [ options ]
```

where  $r_0$  is the null (hypothesized) correlation or the value of the correlation under the null hypothesis, and  $r_a$  is the alternative (target) correlation or the value of the correlation under the alternative hypothesis.  $r_0$  and  $r_a$  may each be specified either as one number or as a list of values in parentheses (see [U] [11.1.8 numlist](#)).

<i>options</i>	Description
Main	
* <u>alpha</u> ( <i>numlist</i> )	significance level; default is <code>alpha(0.05)</code>
* <u>power</u> ( <i>numlist</i> )	power; default is <code>power(0.8)</code>
* <u>beta</u> ( <i>numlist</i> )	probability of type II error; default is <code>beta(0.2)</code>
* <u>n</u> ( <i>numlist</i> )	sample size; required to compute power or effect size
<u>nfractional</u>	allow fractional sample size
* <u>diff</u> ( <i>numlist</i> )	difference between the alternative correlation and the null correlation, $r_a - r_0$ ; specify instead of the alternative correlation $r_a$
<u>direction</u> ( <u>upper</u>   <u>lower</u> )	direction of the effect for effect-size determination; default is <code>direction(upper)</code> , which means that the postulated value of the parameter is larger than the hypothesized value
<u>onesided</u>	one-sided test; default is two sided
<u>parallel</u>	treat number lists in starred options or in command arguments as parallel when multiple values per option or argument are specified (do not enumerate all possible combinations of values)
Table	
[ <u>no</u> ] <u>table</u> [ ( <i>tablespec</i> ) ]	suppress table or display results as a table; see [PSS] <a href="#">power, table</a>
<u>saving</u> ( <i>filename</i> [ , <i>replace</i> ])	save the table data to <i>filename</i> ; use <code>replace</code> to overwrite existing <i>filename</i>
Graph	
<u>graph</u> [ ( <i>graphopts</i> ) ]	graph results; see [PSS] <a href="#">power, graph</a>
Iteration	
<u>init</u> (#)	initial value for sample size or correlation
<u>iterate</u> (#)	maximum number of iterations; default is <code>iterate(500)</code>
<u>tolerance</u> (#)	parameter tolerance; default is <code>tolerance(1e-12)</code>
<u>ftolerance</u> (#)	function tolerance; default is <code>ftolerance(1e-12)</code>
[ <u>no</u> ] <u>log</u>	suppress or display iteration log
[ <u>no</u> ] <u>dots</u>	suppress or display iterations as dots
<u>notitle</u>	suppress the title

\*Specifying a list of values in at least two starred options, or at least two command arguments, or at least one starred option and one argument results in computations for all possible combinations of the values; see [U] [11.1.8 numlist](#). Also see the `parallel` option.

`notitle` does not appear in the dialog box.

where *tablespec* is

```
column [ :label ] [ column [ :label ] [ ... ] ] [ , tableopts ]
```

*column* is one of the columns defined [below](#), and *label* is a column label (may contain quotes and compound quotes).

<i>column</i>	Description	Symbol
alpha	significance level	$\alpha$
power	power	$1 - \beta$
beta	type II error probability	$\beta$
N	number of subjects	$N$
delta	effect size	$\delta$
r0	null correlation	$\rho_0$
ra	alternative correlation	$\rho_a$
diff	difference between the alternative and null correlations	$\rho_a - \rho_0$
target	target parameter; synonym for ra	
_all	display all supported columns	

Column beta is shown in the default table in place of column power if specified.

## Options

### Main

`alpha()`, `power()`, `beta()`, `n()`, `nfractional`; see [PSS] **power**. The `nfractional` option is allowed only for sample-size determination.

`diff(numlist)` specifies the difference between the alternative correlation and the null correlation,  $r_a - r_0$ . You can specify either the alternative correlation  $r_a$  as a command argument or the difference between the two correlations in the `diff()` option. If you specify `diff(#)`, the alternative correlation is computed as  $r_a = r_0 + \#$ . This option is not allowed with the effect-size determination.

`direction()`, `onesided`, `parallel`; see [PSS] **power**.

### Table

`table`, `table()`, `notable`; see [PSS] **power, table**.

`saving()`; see [PSS] **power**.

### Graph

`graph`, `graph()`; see [PSS] **power, graph**. Also see the *column* table for a list of symbols used by the graphs.

### Iteration

`init(#)` specifies the initial value for the iteration procedure. Iteration is used to compute sample size or target correlation for a two-sided test. The default initial value for the estimated parameter is obtained from the corresponding closed-form one-sided computation using the significance level  $\alpha/2$ .

`iterate()`, `tolerance()`, `ftolerance()`, `log`, `nolog`, `dots`, `nodots`; see [PSS] **power**.

The following option is available with `power onecorrelation` but is not shown in the dialog box: `notitle`; see [PSS] **power**.

## Remarks and examples

Remarks are presented under the following headings:

*Introduction*  
*Using power onecorrelation*  
*Computing sample size*  
*Computing power*  
*Computing effect size and target correlation*  
*Performing hypothesis tests on correlation*

This entry describes the `power onecorrelation` command and the methodology for power and sample-size analysis for a one-sample correlation test. See [PSS] [intro](#) for a general introduction to power and sample-size analysis and [PSS] [power](#) for a general introduction to the `power` command using hypothesis tests.

## Introduction

Correlation analysis emanates from studies quantifying dependence between random variables, such as dependence between height and weight of individuals, between blood pressure and cholesterol levels, between the SAT scores of students and their first-year grade point average, between the number of minutes per week customers spend using a new fitness program and the number of pounds lost, and many more.

The correlation coefficient  $\rho$  is commonly used to measure the strength of such dependence. We consider Pearson's correlation obtained by standardizing the covariance between two random variables. As a result, the correlation coefficient has no units and ranges between  $-1$  and  $1$ .

This entry describes power and sample-size analysis for the inference about the population correlation coefficient performed using hypothesis testing. Specifically, we consider the null hypothesis  $H_0: \rho = \rho_0$  versus the two-sided alternative hypothesis  $H_a: \rho \neq \rho_0$ , the upper one-sided alternative  $H_a: \rho > \rho_0$ , or the lower one-sided alternative  $H_a: \rho < \rho_0$ . In most applications, the null value of the correlation,  $\rho_0$ , is set to zero.

Statistical inference on the correlation coefficient requires a distributional assumption between two random variables—bivariate normality with correlation  $\rho$ . The distribution of the sample correlation coefficient is rather complicated except under the null hypothesis of the true correlation being zero,  $H_0: \rho = 0$ , under which it is a Student's  $t$  distribution (for example, see [Graybill \[1961, 209\]](#)). The general inference of  $H_0: \rho = \rho_0$  is based on the asymptotic approximation.

One common approach in testing hypotheses concerning the correlation parameter is to use an inverse hyperbolic tangent transformation,  $\tanh^{-1}(x) = 0.5 \ln(1+x)/\ln(1-x)$ , also known as Fisher's  $z$  transformation when applied to the correlation coefficient ([Fisher 1915](#)). Specifically, if  $\hat{\rho}$  is the sample correlation coefficient and  $n$  is the sample size, [Fisher \(1915\)](#) showed that

$$\sqrt{n-3} \{ \tanh^{-1}(\hat{\rho}) - \tanh^{-1}(\rho) \} \sim N(0, 1)$$

for  $n$  as small as 10, although the approximation tends to perform better for  $n > 25$ . The null hypothesis  $H_0: \rho = \rho_0$  is equivalent to  $H_0: \tanh^{-1}(\rho) = \tanh^{-1}(\rho_0)$ . The latter test is referred to as Fisher's  $z$  test.

`power onecorrelation` performs computations based on the asymptotic Fisher's  $z$  test.

## Using power onecorrelation

`power onecorrelation` computes sample size, power, or target correlation for a one-sample correlation test. All computations are performed for a two-sided hypothesis test where, by default, the significance level is set to 0.05. You may change the significance level by specifying the `alpha()` option. You can specify the `onesided` option to request a one-sided test.

To compute sample size, you must specify the correlations under the null and alternative hypotheses,  $r_0$  and  $r_a$ , respectively, and, optionally, the power of the test in the `power()` option. The default power is set to 0.8.

To compute power, you must specify the sample size in the `n()` option and the correlations under the null and alternative hypotheses,  $r_0$  and  $r_a$ , respectively.

Instead of the alternative correlation,  $r_a$ , you may specify the difference  $r_a - r_0$  between the alternative correlation and the null correlation in the `diff()` option when computing sample size or power.

To compute effect size, the difference between the alternative and null correlations, and target correlation, you must specify the sample size in the `n()` option, the power in the `power()` option, the null correlation  $r_0$ , and, optionally, the direction of the effect. The direction is upper by default, `direction(upper)`, which means that the target correlation is assumed to be larger than the specified null value. This is also equivalent to the assumption of a positive effect size. You can change the direction to lower, which means that the target correlation is assumed to be smaller than the specified null value, by specifying the `direction(lower)` option. This is equivalent to assuming a negative effect size.

By default, the computed sample size is rounded up. You can specify the `nfractional` option to see the corresponding fractional sample size; see *Fractional sample sizes* in [PSS] **unbalanced designs** for an example. The `nfractional` option is allowed only for sample-size determination.

The sample-size and effect-size determinations for a two-sided test require iteration. The default initial value for the estimated parameter is obtained from the corresponding closed-form one-sided computation using the significance level  $\alpha/2$ . The default initial value may be changed by specifying the `init()` option. See [PSS] **power** for the descriptions of other options that control the iteration procedure.

In the following sections, we describe the use of `power onecorrelation` accompanied by examples for computing sample size, power, and target correlation.

## Computing sample size

To compute sample size, you must specify the correlations under the null and alternative hypotheses,  $r_0$  and  $r_a$ , respectively, and, optionally, the power of the test in the `power()` option. A default power of 0.8 is assumed if `power()` is not specified.

### ► Example 1: Sample size for a one-sample correlation test

Consider a study where the goal is to test the existence of a positive correlation between height and weight of individuals, that is,  $H_0: \rho = 0$  versus  $H_a: \rho > 0$ . Before conducting the study, we wish to determine the sample size required to detect a correlation of at least 0.5 with 80% power using a one-sided 5%-level test. We specify the values 0 and 0.5 of the null and alternative correlations after the command name. The only option we need to specify is `onesided`, which requests a one-sided test. We omit options `alpha(0.05)` and `power(0.8)` because the specified values are their defaults.

```

. power onecorrelation 0 0.5, onesided
Estimated sample size for a one-sample correlation test
Fisher's z test
Ho: r = r0 versus Ha: r > r0
Study parameters:
    alpha = 0.0500
    power = 0.8000
    delta = 0.5000
    r0 = 0.0000
    ra = 0.5000
Estimated sample size:
    N = 24

```

A sample of 24 individuals must be obtained to detect a correlation of at least 0.5 between height and weight with 80% power using an upper one-sided 5%-level test.

If we anticipate a stronger correlation of, say, 0.7, we will need a sample of only 12 subjects:

```

. power onecorrelation 0 0.7, onesided
Estimated sample size for a one-sample correlation test
Fisher's z test
Ho: r = r0 versus Ha: r > r0
Study parameters:
    alpha = 0.0500
    power = 0.8000
    delta = 0.7000
    r0 = 0.0000
    ra = 0.7000
Estimated sample size:
    N = 12

```

This is consistent with our expectation that in this example, as the alternative correlation increases, the required sample size decreases.

◀

## Computing power

To compute power, you must specify the sample size in the `n()` option and the correlations under the null and alternative hypotheses,  $r_0$  and  $r_a$ , respectively.

### ► Example 2: Power of a one-sample correlation test

Continuing with [example 1](#), we will suppose that we are designing a pilot study and would like to keep our initial sample small. We want to find out how much the power of the test would decrease if we sample only 15 individuals. To compute power, we specify the sample size in `n()`:

```

. power onecorrelation 0 0.5, n(15) onesided
Estimated power for a one-sample correlation test
Fisher's z test
Ho: r = r0 versus Ha: r > r0
Study parameters:
    alpha =    0.0500
      N =      15
    delta =    0.5000
      r0 =    0.0000
      ra =    0.5000
Estimated power:
    power =    0.6018

```

The power decreases to roughly 60%. The power to detect a correlation of 0.5 is low for a sample of 15 individuals. We should consider either a larger target correlation or a larger sample.

◀

### ▷ Example 3: Nonzero null hypothesis

To demonstrate power computation for a nonzero null hypothesis, we consider an example from Graybill (1961, 211) of a hypothetical study where the primary interest lies in the correlation between protein and fat for a particular breed of dairy cow. The sample contains 24 cows. Suppose we want to compute the power of a two-sided test with the null hypothesis  $H_0: \rho = 0.5$  versus the alternative hypothesis  $H_a: \rho \neq 0.5$ . Suppose that the correlation between protein and fat was estimated to be 0.3 in an earlier pilot study. We can compute power by typing

```

. power onecorrelation 0.5 0.3, n(24)
Estimated power for a one-sample correlation test
Fisher's z test
Ho: r = r0 versus Ha: r != r0
Study parameters:
    alpha =    0.0500
      N =     24
    delta =   -0.2000
      r0 =    0.5000
      ra =    0.3000
Estimated power:
    power =    0.1957

```

The power of the test is estimated to be 0.1957 for the above study, which is unacceptably low for practical purposes. We should consider either increasing the sample size or choosing a different target correlation.

In fact, the author considers two values of the alternative correlation, 0.3 and 0.2. We can compute the corresponding powers for the multiple values of the alternative correlation by enclosing them in parentheses, as we demonstrate below.



```
. power onecorrelation 0.5 (0.3 0.2), n(24)
Estimated power for a one-sample correlation test
Fisher's z test
Ho: r = r0 versus Ha: r != r0
```

alpha	power	N	delta	r0	ra
.05	.1957	24	-.2	.5	.3
.05	.3552	24	-.3	.5	.2

The power increases from 20% to 36% as the alternative correlation decreases from 0.3 to 0.2. It may seem counterintuitive that the power increased as the value of the alternative correlation decreased. It is important to remember that the power of the correlation test is an increasing function of the magnitude of the effect size,  $\delta = \rho_a - \rho_0$ , the difference between the alternative and null values of the correlation. So the power of 20% corresponds to  $|\delta| = |0.3 - 0.5| = 0.2$ , and the power of 36% corresponds to  $|\delta| = 0.3$ .

◀

For multiple values of parameters, the results are automatically displayed in a table, as we see above. For more examples of tables, see [PSS] [power, table](#). If you wish to produce a power plot, see [PSS] [power, graph](#).

## Computing effect size and target correlation

In a one-sample correlation analysis, the effect size is commonly defined as the difference between the alternative correlation and the null correlation,  $\delta = \rho_a - \rho_0$ . For a zero-null hypothesis, the effect size is the alternative correlation.

The distribution of the sample correlation coefficient is symmetric when the population correlation  $\rho$  is zero and becomes negatively skewed as  $\rho$  approaches 1 and positively skewed as  $\rho$  approaches  $-1$ . Also the sampling variance approaches zero as  $|\rho|$  approaches 1. Clearly, the power of a correlation test depends on the magnitude of the true correlation; therefore,  $\delta$  alone is insufficient for determining the power. In other words, for the same difference  $\delta = 0.1$ , the power to detect the difference between  $\rho_a = 0.2$  and  $\rho_0 = 0.1$  is not the same as that between  $\rho_a = 0.3$  and  $\rho_0 = 0.2$ ; the former difference is associated with the larger sampling variability and thus has lower power to be detected.

Sometimes, we may be interested in determining the smallest effect (for a given null correlation) and the corresponding target correlation that yield a statistically significant result for prespecified sample size and power. In this case, power, sample size, and null correlation must be specified. In addition, you must also decide on the direction of the effect: upper, meaning  $\rho_a > \rho_0$ , or lower, meaning  $\rho_a < \rho_0$ . The direction may be specified in the `direction()` option; `direction(upper)` is the default.

### ► Example 4: Minimum detectable value of the correlation

Continuing with [example 2](#), we will compute the minimum positive correlation that can be detected given a sample of 15 individuals and 80% power. To solve for target correlation, we specify the null correlation of 0 after the command, sample size `n(15)`, and power `power(0.8)`:

```

. power onecorrelation 0, n(15) power(0.8) onesided
Estimated target correlation for a one-sample correlation test
Fisher's z test
Ho: r = r0 versus Ha: r > r0; ra > r0
Study parameters:
      alpha =    0.0500
      power =    0.8000
         N =      15
        r0 =    0.0000
Estimated effect size and target correlation:
      delta =    0.6155
         ra =    0.6155

```

We find that the minimum value of the correlation as well as of the effect size that would yield a statistically significant result in this study is 0.6155.

In this example, we computed the correlation assuming an upper direction,  $\rho_a > \rho_0$ , or a positive effect,  $\delta > 0$ . To request a lower direction, or a negative effect, we can specify the `direction(lower)` option.

◀

## Performing hypothesis tests on correlation

After the data are collected, you can use the `pwcorr` command to test for correlation; see [R] [correlate](#).

### ▶ Example 5: Testing for correlation

For example, consider `auto.dta`, which contains various characteristics of 74 cars. Suppose that our study goal is to investigate the existence of a correlation between miles per gallon and weights of cars. The corresponding variables in the dataset are `mpg` and `weight`. We use `pwcorr` to perform the test.

```

. use http://www.stata-press.com/data/r15/auto
(1978 Automobile Data)
. pwcorr mpg weight, sig

```

	mpg	weight
mpg	1.0000	
weight	-0.8072	1.0000
	0.0000	

The test finds a statistically significant negative correlation of  $-0.8$  between `mpg` and `weight`; the  $p$ -value is less than 0.0001.

We use the parameters of this study to perform a sample-size analysis we would have conducted before the study.

```
. power onecorrelation 0 -0.8, power(0.9) alpha(0.01)
Performing iteration ...
Estimated sample size for a one-sample correlation test
Fisher's z test
Ho: r = r0 versus Ha: r != r0
Study parameters:
      alpha =    0.0100
      power =    0.9000
      delta =   -0.8000
      r0 =     0.0000
      ra =    -0.8000
Estimated sample size:
      N =          16
```

We find that the sample size required to detect a correlation of  $-0.8$  with 90% power using a 1%-level two-sided test is 16. The current sample contains many more cars (74), which would allow us to detect a potentially smaller (in absolute value) correlation.

We should point out that `pwcorr` performs an exact one-sample correlation test of  $H_0: \rho = 0$ , which uses a Student's  $t$  distribution. `power onecorrelation` performs computation for an approximate Fisher's  $z$  test, which uses the normal distribution.

◀

## Stored results

`power onecorrelation` stores the following in `r()`:

### Scalars

<code>r(alpha)</code>	significance level
<code>r(power)</code>	power
<code>r(beta)</code>	probability of a type II error
<code>r(delta)</code>	effect size
<code>r(N)</code>	sample size
<code>r(nfractional)</code>	1 if <code>nfractional</code> is specified, 0 otherwise
<code>r(onesided)</code>	1 for a one-sided test, 0 otherwise
<code>r(r0)</code>	correlation under the null hypothesis
<code>r(ra)</code>	correlation under the alternative hypothesis
<code>r(diff)</code>	difference between the alternative and null correlations
<code>r(separator)</code>	number of lines between separator lines in the table
<code>r(divider)</code>	1 if <code>divider</code> is requested in the table; 0 otherwise
<code>r(init)</code>	initial value for sample size or correlation
<code>r(maxiter)</code>	maximum number of iterations
<code>r(iter)</code>	number of iterations performed
<code>r(tolerance)</code>	requested parameter tolerance
<code>r(deltax)</code>	final parameter tolerance achieved
<code>r(ftolerance)</code>	requested distance of the objective function from zero
<code>r(function)</code>	final distance of the objective function from zero
<code>r(converged)</code>	1 if iteration algorithm converged, 0 otherwise

### Macros

<code>r(type)</code>	test
<code>r(method)</code>	<code>onecorrelation</code>
<code>r(direction)</code>	upper or lower
<code>r(columns)</code>	displayed table columns
<code>r(labels)</code>	table column labels
<code>r(widths)</code>	table column widths
<code>r(formats)</code>	table column formats

### Matrices

<code>r(pss_table)</code>	table of results
---------------------------	------------------

## Methods and formulas

Let  $\rho$  denote Pearson's correlation and  $\hat{\rho}$  its estimator, the sample correlation coefficient.

A one-sample correlation test involves testing the null hypothesis  $H_0: \rho = \rho_0$  versus the two-sided alternative hypothesis  $H_a: \rho \neq \rho_0$ , the upper one-sided alternative  $H_a: \rho > \rho_0$ , or the lower one-sided alternative  $H_a: \rho < \rho_0$ .

The exact distribution of the sample correlation coefficient  $\hat{\rho}$  is complicated for testing a general hypothesis  $H_0: \rho = \rho_0 \neq 0$ . An alternative is to apply an inverse hyperbolic tangent,  $\tanh^{-1}(\cdot)$ , or Fisher's  $z$  transformation to  $\hat{\rho}$  to approximate its sampling distribution by a normal distribution (Fisher 1915). The hypothesis  $H_0: \rho = \rho_0$  is then equivalent to  $H_0: \tanh^{-1}(\rho) = \tanh^{-1}(\rho_0)$ .

Let  $Z = \tanh^{-1}(\hat{\rho}) = 0.5 \ln\{(1 + \hat{\rho})/(1 - \hat{\rho})\}$  denote Fisher's  $z$  transformation of  $\hat{\rho}$ . Then  $Z$  follows a normal distribution with mean  $\mu_z = \tanh^{-1}(\rho)$  and standard deviation  $\sigma_z = 1/\sqrt{n-3}$ , where  $n$  is the sample size (for example, see Graybill [1961, 211]; Anderson [2003, 134]).

Let  $\alpha$  be the significance level,  $\beta$  be the probability of a type II error, and  $z_{1-\alpha/k}$  and  $z_\beta$  be the  $(1 - \alpha/k)$ th and the  $\beta$ th quantiles of the standard normal distribution. Denote  $\delta_z = \mu_a - \mu_0 = \tanh^{-1}(\rho_a) - \tanh^{-1}(\rho_0)$ .

The power  $\pi = 1 - \beta$  is computed using

$$\pi = \begin{cases} \Phi\left(\frac{\delta_z}{\sigma_z} - z_{1-\alpha}\right) & \text{for an upper one-sided test} \\ \Phi\left(-\frac{\delta_z}{\sigma_z} - z_{1-\alpha}\right) & \text{for a lower one-sided test} \\ \Phi\left(\frac{\delta_z}{\sigma_z} - z_{1-\alpha/2}\right) + \Phi\left(-\frac{\delta_z}{\sigma_z} - z_{1-\alpha/2}\right) & \text{for a two-sided test} \end{cases} \quad (1)$$

where  $\Phi(\cdot)$  is the cdf of a standard normal distribution.

The sample size  $n$  for a one-sided test is computed using

$$n = 3 + \left(\frac{z_{1-\alpha} - z_\beta}{\delta_z}\right)^2$$

See, for example, Lachin (1981) for the case of  $\rho_0 = 0$ .

The minimum detectable value of the correlation is obtained by first finding the corresponding minimum value of  $\mu_a$  and then applying the inverse Fisher's  $z$  transformation to the  $\mu_a$ :

$$\rho_a = \frac{e^{2\mu_a} - 1}{e^{2\mu_a} + 1}$$

For a one-sided test,  $\mu_a = \mu_0 + \sigma_z(z_{1-\alpha} - z_\beta)$  when  $\mu_a > \mu_0$ , and  $\mu_a = \mu_0 - \sigma_z(z_{1-\alpha} - z_\beta)$  when  $\mu_a < \mu_0$ .

For a two-sided test, the sample size  $n$  and  $\mu_a$  are obtained by iteratively solving the two-sided power equation in (1) for  $n$  and  $\mu_a$ , respectively. The initial values are obtained from the respective formulas for the one-sided computation with the significance level  $\alpha/2$ .

If the `nfractional` option is not specified, the computed sample size is rounded up.

## References

- Anderson, T. W. 2003. *An Introduction to Multivariate Statistical Analysis*. 3rd ed. New York: Wiley.
- Fisher, R. A. 1915. Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika* 10: 507–521.
- Graybill, F. A. 1961. *An Introduction to Linear Statistical Models, Vol. 1*. New York: McGraw–Hill.
- Lachin, J. M. 1981. Introduction to sample size determination and power analysis for clinical trials. *Controlled Clinical Trials* 2: 93–113.

## Also see

- [PSS] **power** — Power and sample-size analysis for hypothesis tests
- [PSS] **power, graph** — Graph results from the power command
- [PSS] **power, table** — Produce table of results from the power command
- [PSS] **Glossary**
- [R] **correlate** — Correlations (covariances) of variables or coefficients