

**power oneslope** — Power analysis for a slope test in a simple linear regression

<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="#">Reference</a>	<a href="#">Also see</a>		

## Description

`power oneslope` computes sample size, power, or the target slope coefficient for a test of a slope in a simple linear regression. By default, it computes sample size given power and the slope coefficient. Alternatively, it computes power given sample size and the slope coefficient, or it computes the target slope coefficient given sample size, power, and the coefficient under the null. See [\[PSS\] power](#) for a general introduction to the `power` command using hypothesis tests.

## Quick start

Sample size for a two-sided test of  $H_0: b = 1.0$  versus  $H_a: b \neq 1.0$  with null slope  $b_0 = 1.0$ , alternative slope  $b_a = 1.2$ , covariate standard deviation of 3, and error standard deviation of 1.5 using default power of 0.8 and significance level  $\alpha = 0.05$

```
power oneslope 1.0 1.2, sdx(3) sderror(1.5)
```

As above, but for a one-sided test with power of 0.9

```
power oneslope 1.0 1.2, sdx(3) sderror(1.5) power(.9) onesided
```

Specify  $b_0$  and the difference between  $b_a$  and  $b_0$  instead of specifying  $b_a = 1.2$

```
power oneslope 1.0, sdx(3) sderror(1.5) diff(0.2)
```

As above, but specify correlation between the dependent variable and the covariate instead of the error standard deviation

```
power oneslope 1.0, sdx(3) corr(.923) diff(0.2)
```

Power for sample size of 75

```
power oneslope 1.0 1.2, sdx(3) sderror(1.5) n(75)
```

Power for sample sizes of 50, 60, 70, and 80

```
power oneslope 1.0 1.2, sdx(3) sderror(1.5) n(50(10)80)
```

As above, but display results in a graph of power versus sample size

```
power oneslope 1.0 1.2, sdx(3) sderror(1.5) n(50(10)80) graph
```

Effect size and target slope for sample size of 40, power of 0.9, and  $\alpha = 0.01$

```
power oneslope 1.0, sdx(3) sderror(1.5) n(40) power(.9) alpha(0.01)
```

## Menu

Statistics > Power and sample size

## Syntax

*Compute sample size*

```
power oneslope  $b_0$   $b_a$  [ , power(numlist) options ]
```

*Compute power*

```
power oneslope  $b_0$   $b_a$  , n(numlist) [ options ]
```

*Compute effect size and target slope*

```
power oneslope  $b_0$  , n(numlist) power(numlist) [ options ]
```

where  $b_0$  is the null (hypothesized) slope or the value of the slope coefficient under the null hypothesis and  $b_a$  is the alternative (target) slope or the value of the slope coefficient under the alternative hypothesis.  $b_0$  and  $b_a$  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 slope and the null slope coefficients, $b_a - b_0$ ; specify instead of the alternative slope $b_a$
* <u>sdx</u> ( <i>numlist</i> )	standard deviation of the covariate of interest; default is <code>sdx(1)</code>
* <u>sderror</u> ( <i>numlist</i> )	standard deviation of the error term of the regression model; may not be combined with <code>sdv()</code> or <code>corr()</code> ; default is <code>sderror(1)</code>
* <u>sdv</u> ( <i>numlist</i> )	standard deviation of the dependent variable; may not be combined with <code>sderror()</code> or <code>corr()</code>
* <u>corr</u> ( <i>numlist</i> )	correlation between the dependent variable and the covariate of interest; may not be combined with <code>sderror()</code> or <code>sdv()</code>
<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> [ , <u>replace</u> ])	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 slope
<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	$\delta$
b0	null slope coefficient	$b_0$
ba	alternative slope coefficient	$b_a$
diff	difference between alternative and null slope coefficients	$b_a - b_0$
sdx	standard deviation of covariate	$\sigma_x$
sderror	standard deviation of error term	$\sigma$
sdv	standard deviation of dependent variable	$\sigma_y$
corr	correlation between dependent variable and covariate	$\rho$
target	target parameter; synonym for ba	
_all	display all supported columns	

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

Columns diff, sdv, and corr are shown in the default table 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 slope and the null slope coefficients,  $b_a - b_0$ . You can specify either the alternative slope  $b_a$  as a command argument or the difference between the two slopes in `diff()`. If you specify `diff(#)`, the alternative slope is computed as  $b_a = b_0 + \#$ . This option is not allowed with effect-size determination.

`sdx(numlist)` specifies the standard deviation for the covariate of interest. The default is `sdx(1)`.

`sderror(numlist)` specifies the standard deviation of the error term of the regression model. The default is `sderror(1)`. This option may not be combined with `sdv()` or `corr()`.

`sdv(numlist)` specifies the standard deviation of the dependent variable in the regression model. This option may not be combined with `sderror()` or `corr()`.

`corr(numlist)` specifies the correlation between the covariate of interest and the dependent variable. This option may not be combined with `sderror()` or `sdv()`.

`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 of the sample size for the sample-size determination or the initial value of the slope for the effect-size determination. The default is to use a closed-form normal approximation to compute an initial value for the sample size and a bisection search method to compute an initial value for the effect size.

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

The following option is available with `power oneslope` but is not shown in the dialog box:

`notitle`; see [PSS] [power](#).

## Remarks and examples

[stata.com](http://www.stata.com)

Remarks are presented under the following headings:

*Introduction*

*Using power oneslope*

*Computing sample size*

*Computing power*

*Computing effect size and target slope*

*Performing hypothesis tests on the slope coefficient*

`power oneslope` computes sample size, power, and the target slope coefficient for a linear regression slope test. See [PSS] [intro](#) for a general introduction to power and sample-size analysis, and see [PSS] [power](#) for a general introduction to the `power` command using hypothesis tests.

## Introduction

In a simple linear regression, researchers estimate the relationship between a covariate and the outcome without controlling for other covariates. Consider a study where a researcher would like to compare the slope from a simple linear regression with a hypothesized slope. For example, health researchers may want to know whether the effect of time spent exercising on body mass index (BMI) has changed compared with 10 years ago. Or pharmacological researchers may be interested in knowing whether the slope of the dose–response curve reaches a given magnitude.

We can test whether the covariate of interest,  $x$ , has an effect on the outcome,  $y$ , by using a  $t$  test. In the case of the dose–response study, for example, we may have a model,

$$y_i = a + b_{\text{dose}}x_i + \varepsilon_i \quad i = 1, 2, \dots, n$$

where the errors  $\varepsilon_i$ 's are independently and normally distributed with mean 0 and standard deviation  $\sigma$ . The null hypothesis would be  $H_0: b_{\text{dose}} = b_0$ , where  $b_{\text{dose}}$  is the slope coefficient for the dose of the drug that was administered and  $b_0$  is the hypothesized value of the slope. In general, for a slope coefficient  $b$ , we consider the null hypothesis  $H_0: b = b_0$  versus the two-sided alternative hypothesis  $H_a: b \neq b_0$ , the upper one-sided alternative  $H_a: b > b_0$ , or the lower one-sided alternative  $H_a: b < b_0$ .

The `power oneslope` command provides power and sample-size analysis for the comparison of a slope coefficient with a reference value using a  $t$  test in a simple linear regression. For power and sample-size analysis in a multiple linear regression, see [PSS] [power rsquared](#) and [PSS] [power pcorr](#).

## Using `power oneslope`

`power oneslope` computes sample size, power, or the target slope,  $b_a$ , for a test of slope in a simple linear regression. By default, all computations are performed for a two-sided hypothesis test where the significance level is set to 0.05. You can change the significance level by specifying the `alpha()` option. You can request a one-sided test by specifying the `onesided` option.

By default, all computations use one as the standard deviation of the covariate of interest and as the standard deviation of the error. You can change these values by specifying the `sdx()` and `sderror()` options. Instead of the `sderror()` option, you can combine the `sdx()` option with the `sdyc()` option to specify the standard deviation of the dependent variable or combine the `sdx()` option with the `corr()` option to specify the correlation between the covariate and the dependent variable.

To compute sample size, you must specify the slope under the null hypothesis ( $b_0$ ) and the slope under the alternative hypothesis ( $b_a$ ), and you may specify 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, the null slope  $b_0$ , and the alternative slope  $b_a$ .

When computing sample size or power, you may specify the difference between the alternative slope and null slope,  $b_a - b_0$ , in the `diff()` option instead of specifying the alternative slope.

To compute the effect size  $\delta$ , you must specify the sample size in the `n()` option, the power in the `power()` option, and the null slope  $b_0$ , and you may specify the direction of the effect.  $\delta$  is defined as the difference between the alternative and null values of the slope multiplied by the ratio of standard deviation of the covariate to that of the error term,  $\delta = (b_a - b_0)\sigma_x/\sigma$ . The direction is upper by default, `direction(upper)`, which means that the target slope is assumed to be larger than the specified null value. This is also equivalent to the assumption of a positive effect size. You may change the direction to lower by specifying the `direction(lower)` option, which means that the target slope is assumed to be smaller than the specified null value. 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.

`power oneslope`'s computations of sample size and effect size require iteration. A noncentral Student's  $t$  distribution is used for the computations. The degrees of freedom depends on the sample size, and the noncentrality parameter depends on the sample size and effect size. The default initial value of the sample size is obtained using a closed-form normal approximation. The default initial value of the effect size is obtained using a bisection search method. The default initial values may be changed by specifying the `init()` option. See [PSS] [power](#) for the descriptions of other options that control the iteration procedure.

## Computing sample size

To compute sample size, you must specify the slope coefficients under the null,  $b_0$ , and alternative,  $b_a$ , hypotheses, and you may specify 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 linear regression slope test

Consider an example from [Dupont and Plummer \(1998\)](#) that discusses the effectiveness of a dieting program in encouraging patients to exercise regularly. Suppose we are conducting a similar study and we want to know the effect of average time spent per day exercising on BMI, measured in  $\text{kg}/\text{m}^2$ , after six months in the program. The parameter of interest is the slope coefficient  $b$ , which measures the effect of exercising on BMI, and our null hypothesis is  $H_0: b = 0$  versus a two-sided  $H_a: b \neq 0$ .

We wish to compute the sample size required to detect a drop in BMI of  $0.0667 \text{ kg}/\text{m}^2$  per minute of exercise, the effect reported in [Dupont and Plummer](#), with 80% power using a 5%-level two-sided test. Using values obtained from previous studies, we specify a standard deviation of 7.5 minutes for time spent exercising in `sdx()` and  $4.0 \text{ kg}/\text{m}^2$  for BMI in `sdv()`.

```
. power oneslope 0 -0.0667, sdx(7.5) sdv(4)
Performing iteration ...
Estimated sample size for a linear regression slope test
t test
Ho: b = b0 versus Ha: b != b0
Study parameters:
    alpha =    0.0500
    power =    0.8000
    delta =   -0.1261
    b0 =      0.0000
    ba =   -0.0667
    sdx =     7.5000
    sderror =  3.9686
    sdv =     4.0000
Estimated sample size:
    N =      496
```

We find that a sample of 496 subjects is required to detect a drop of  $0.0667 \text{ kg}/\text{m}^2$  BMI per minute with 80% power using a 5%-level two-sided test. The effect size (`delta`) and standard deviation of the error term (`sderror`) are calculated using the given information about the null and alternative slopes and the standard deviations of the dependent and independent variables; see [Methods and formulas](#).

As we mentioned in [Using power oneslope](#), sample-size computation requires iteration. The iteration log is suppressed by default, but you can display it by specifying the `log` option.

◀

### ▷ Example 2: Specifying difference between slopes

Instead of specifying the alternative slope as in [example 1](#), we can obtain the same results by specifying the difference between the slope coefficient under the alternative and null hypotheses in the `diff()` option. We specify `diff(-0.0667)` and the same standard deviation of the covariate of interest in the `sdv()` option and standard deviation of the outcome in the `sdv()` option that we used in [example 1](#).

In this case, the difference is equal to  $b_a$ , specified in [example 1](#). If we wanted to test for a null value other than zero, then  $b_a$  would not equal the difference.

```

. power oneslope 0, diff(-0.0667) sdx(7.5) sdy(4)
Performing iteration ...
Estimated sample size for a linear regression slope test
t test
Ho: b = b0 versus Ha: b != b0
Study parameters:
    alpha =    0.0500
    power  =    0.8000
    delta  =   -0.1261
    b0     =    0.0000
    ba     =   -0.0667
    diff   =   -0.0667
    sdx    =    7.5000
    sderror=    3.9686
    sdy    =    4.0000
Estimated sample size:
    N =      496

```

Although the sample size estimate is the same, the difference between the alternative and null values is also reported in the output when we specify the `diff()` option. Notice that the command computed the value of  $b_a$ .

◀

### ▶ Example 3: Specifying standard deviation of the error

Instead of  $\sigma_y$ , the standard deviation of the dependent variable, we can specify the standard deviation of the error term in the regression model using the `sderror()` option. The value of the standard deviation is obtained from the relation  $\sigma = \sqrt{\sigma_y^2 - b_a^2 \sigma_x^2} = \sqrt{4^2 - \{(-0.0667)^2 \times 7.5^2\}} \approx 3.9686$ .

```

. power oneslope 0 -0.0667, sdx(7.5) sderror(3.9686)
Performing iteration ...
Estimated sample size for a linear regression slope test
t test
Ho: b = b0 versus Ha: b != b0
Study parameters:
    alpha =    0.0500
    power  =    0.8000
    delta  =   -0.1261
    b0     =    0.0000
    ba     =   -0.0667
    sdx    =    7.5000
    sderror=    3.9686
Estimated sample size:
    N =      496

```

◀

### ▶ Example 4: Specifying correlation

We can also specify the correlation between time per day spent exercising and BMI using the `corr()` option. The value of the correlation  $\rho$  is obtained from the relation  $\rho = b_a \times (\sigma_x / \sigma_y) = -0.0667 \times (7.5/4) \approx -0.1251$ .

```
. power oneslope 0 -0.0667, sdx(7.5) corr(-0.1251)
Performing iteration ...
Estimated sample size for a linear regression slope test
t test
Ho: b = b0 versus Ha: b != b0
Study parameters:
    alpha =    0.0500
    power =    0.8000
    delta =   -0.1261
    b0 =     0.0000
    ba =   -0.0667
    sdx =    7.5000
    sderror =  3.9674
    corr =   -0.1251
Estimated sample size:
    N =      496
```

The sample sizes in the two examples now are the same, but the standard deviation of the errors is slightly different because of limited precision of the specified correlation.

◀

## Computing power

To compute power, you must specify the sample size in the `n()` option and the slopes under the null,  $b_0$ , and alternative,  $b_a$ , hypotheses.

### ▶ Example 5: Power of a linear regression slope test

Continuing with [example 1](#), suppose that we are designing a new study and anticipate that we will observe a similar effect but that we will be able to recruit only a sample of 400 subjects. Given the study parameters from example 1, we compute the power by specifying the sample size of 400 in the `n()` option:

```
. power oneslope 0 -0.0667, n(400) sdx(7.5) sdy(4)
Estimated power for a linear regression slope test
t test
Ho: b = b0 versus Ha: b != b0
Study parameters:
    alpha =    0.0500
    N =       400
    delta =   -0.1261
    b0 =     0.0000
    ba =   -0.0667
    sdx =    7.5000
    sderror =  3.9686
    sdy =    4.0000
Estimated power:
    power =    0.7106
```

With a smaller sample size, the power of the test decreases to about 71%.

◀

### ▷ Example 6: Multiple values of study parameters

To investigate the effect of sample size on power, we can specify a list of sample sizes in the `n()` option:

```
. power oneslope 0 -0.0667, n(50 100 200 400 800) sdx(7.5) sdy(4)
Estimated power for a linear regression slope test
t test
Ho: b = b0 versus Ha: b != b0
```

alpha	power	N	delta	b0	ba	sdx	sderror	sdv
.05	.141	50	-.1261	0	-.0667	7.5	3.969	4
.05	.239	100	-.1261	0	-.0667	7.5	3.969	4
.05	.4263	200	-.1261	0	-.0667	7.5	3.969	4
.05	.7106	400	-.1261	0	-.0667	7.5	3.969	4
.05	.9453	800	-.1261	0	-.0667	7.5	3.969	4

As expected, when the sample size increases, the power tends to get closer to 1.

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 slope

The effect size  $\delta$  for a linear regression slope test is defined as the difference between the alternative and null values of the slope multiplied by the ratio of standard deviations of the covariate to the error term,  $\delta = (b_a - b_0)\sigma_x/\sigma$ .

Sometimes, we may want to determine the minimum detectable effect and the corresponding alternative or target slope coefficient that yield a statistically significant result for a prespecified sample size and power. In this case, we must specify power, sample size, and the null slope coefficient. In addition, we must also decide on the direction of the effect: upper, which means  $b_a > b_0$ , or lower, which means  $b_a < b_0$ . The default direction is upper but may be changed by specifying `direction(lower)`.

### ▷ Example 7: Minimum detectable value of the slope coefficient

Continuing with [example 5](#), we may also be interested in finding the minimum value of the slope coefficient that can be detected with a sample of 400 subjects and 80% power. To compute this, we specify the null value of 0 as the command argument and the required options `n(400)` and `power(0.8)` and continue to use `sdx(7.5)` and `sdv(4)`.

```

. power oneslope 0, n(400) power(0.8) sdx(7.5) sdy(4)
Performing iteration ...
Estimated target slope for a linear regression slope test
t test
Ho: b = b0 versus Ha: b != b0; ba > b0
Study parameters:
      alpha =    0.0500
      power =    0.8000
         N =      400
         b0 =    0.0000
         sdx =    7.5000
      sderror =    3.9611
         sdy =    4.0000

Estimated effect size and target slope:
      delta =    0.1404
         ba =    0.0742

```

The minimum detectable value of the slope coefficient is 0.0742, which corresponds to an effect size of 0.1404. Compared with [example 1](#), this example uses a smaller sample size of 400 subjects, so the minimum detectable difference for the slope coefficient is larger than the absolute value of the alternative slope specified in [example 1](#).

In this example, we assumed the effect to be in the upper direction. By symmetry, the effect size in the lower direction will be  $-0.1404$ , which can be obtained by specifying `direction(lower)`.

◀

## Performing hypothesis tests on the slope coefficient

Suppose we wish to test the hypothesis that the slope coefficient is different from a reference value on the collected data. We can use the `regress` command to estimate the slope and perform a hypothesis test.

### ▶ Example 8: Testing for the slope coefficient

Suppose our study goal is to investigate the effect that the weight of cars (`weight`) has on their mileage (`mpg`). We use `regress` with the `beta` option to estimate the standardized slope using data from `auto.dta`:

```

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

```

Source	SS	df	MS	Number of obs	=	74
Model	1591.9902	1	1591.9902	F(1, 72)	=	134.62
Residual	851.469256	72	11.8259619	Prob > F	=	0.0000
				R-squared	=	0.6515
				Adj R-squared	=	0.6467
Total	2443.45946	73	33.4720474	Root MSE	=	3.4389

  

mpg	Coef.	Std. Err.	t	P> t	Beta
weight	-.0060087	.0005179	-11.60	0.000	-.8071749
_cons	39.44028	1.614003	24.44	0.000	.

The standardized slope is about  $-0.81$ . The  $p$ -value  $< 0.001$  for the test of the slope gives us statistical evidence to reject the null hypothesis of  $H_0: b_{\text{weight}} = 0$  versus a two-sided alternative  $H_a: b_{\text{weight}} \neq 0$  at the 5% significance level.

Suppose we wish to design a new similar study. We use the estimated standardized slope from this study to perform a sample-size analysis for our new study. Because we use the standardized slope, we can use the default value of 1 for `sd $x$ ()`, and we specify `sd $y$ (1)`.

```
. power oneslope 0 -0.81, sd $y$ (1)
Performing iteration ...
Estimated sample size for a linear regression slope test
t test
Ho: b = b0 versus Ha: b != b0
Study parameters:
    alpha =    0.0500
    power =    0.8000
    delta =   -1.3812
    b0 =     0.0000
    ba =    -0.8100
    sd $x$  =    1.0000
    sderror =  0.5864
    sd $y$  =    1.0000
Estimated sample size:
    N =      7
```

We find that a sample size of only 7 is required to detect a slope coefficient of  $-0.81$  with 80% power using a 5%-level two-sided test.

◀

## Stored results

`power oneslope` 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(b0)</code>	slope coefficient under the null hypothesis
<code>r(ba)</code>	slope coefficient under the alternative hypothesis
<code>r(diff)</code>	difference between the alternative and null slopes
<code>r(sd<math>x</math>)</code>	standard deviation of the covariate of interest
<code>r(sderror)</code>	standard deviation of the error term of the regression model
<code>r(sd<math>y</math>)</code>	standard deviation of the dependent variable
<code>r(corr)</code>	correlation between the covariate of interest and the dependent variable
<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 for slope
<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	
r(type)	test
r(method)	oneslope
r(direction)	upper or lower
r(columns)	displayed table columns
r(labels)	table column labels
r(widths)	table column widths
r(formats)	table column formats
Matrices	
r(pss_table)	table of results

## Methods and formulas

Consider a simple linear regression model of a dependent variable  $y$  on a single covariate  $x$ ,

$$y_i = a + bx_i + \varepsilon_i$$

for  $i = 1, \dots, n$ , where  $i$  denotes an individual observation in a sample of  $n$  subjects,  $a$  is the intercept,  $b$  is the slope coefficient, and  $\varepsilon_i$  is the error term.  $\varepsilon_i$ 's are independently and normally distributed with mean 0 and standard deviation  $\sigma$ . Let  $b_0$  denote the null value of the slope coefficient and  $b_a$  denote the alternative value of the slope coefficient.

Testing a linear regression slope involves testing the null hypothesis  $H_0: b = b_0$  versus the two-sided alternative hypothesis  $H_a: b \neq b_0$ , the upper one-sided alternative  $H_a: b > b_0$ , or the lower one-sided alternative  $H_a: b < b_0$ .

Let  $\sigma_y$  denote the standard deviation of  $y$ ,  $\sigma_x$  denote the standard deviation of  $x$ , and  $\rho$  denote the correlation between  $y$  and  $x$ . These terms have the following relations with the standard deviation of the error term  $\sigma$  (Dupont and Plummer 1998):

$$\sigma = b\sigma_x \sqrt{\frac{1}{\rho^2} - 1} = \sqrt{\sigma_y^2 - b^2\sigma_x^2}$$

Let  $\hat{b}$  denote the estimator of the population slope coefficient  $b$ ,  $\hat{\sigma}$  denote the estimated standard deviation of the error term, and  $s = \hat{\sigma}/\sigma_x$  denote the estimated standard error of  $\sqrt{n}\hat{b}$ . The sampling distribution of the test statistic  $t = \sqrt{n}(\hat{b} - b_0)/s$  under the null hypothesis follows a Student's  $t$  distribution with  $n - 2$  degrees of freedom.

Let  $\alpha$  be the significance level,  $\beta$  be the probability of a type II error, and  $t_{n-2,\alpha}$  denote the  $\alpha$ th quantile of a Student's  $t$  distribution with  $n - 2$  degrees of freedom. Let  $\delta = (b_a - b_0)\sigma_x/\hat{\sigma}$  denote the effect size. Under the alternative hypothesis, the test statistic follows a noncentral Student's  $t$  distribution.

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

$$\pi = \begin{cases} 1 - T_{n-2,\lambda}(t_{n-2,1-\alpha}) & \text{for an upper one-sided test} \\ T_{n-2,\lambda}(-t_{n-2,1-\alpha}) & \text{for a lower one-sided test} \\ 1 - T_{n-2,\lambda}(t_{n-2,1-\alpha/2}) + T_{n-2,\lambda}(-t_{n-2,1-\alpha/2}) & \text{for a two-sided test} \end{cases} \quad (1)$$

where  $T_{n-2,\lambda}(\cdot)$  is the cumulative noncentral Student's  $t$  distribution with a noncentrality parameter  $\lambda = \sqrt{n}\delta$ .

Sample size and minimum detectable value of the slope are obtained by iteratively solving nonlinear equations in (1) for  $n$  and  $\delta$ . The default initial value for the sample size is calculated using the following closed-form expression,

$$n = \left( \frac{z_{1-\alpha/k} - z_{\beta}}{\delta} \right)^2$$

where  $k = 1$  for a one-sided and  $k = 2$  for a two-sided test, respectively, and  $z_{1-\alpha/k}$  and  $z_{\beta}$  are the  $(1 - \alpha/k)$ th and the  $\beta$ th quantiles of the standard normal distribution. If the `nfractional` option is not specified, the computed sample size is rounded up.

## Reference

Dupont, W. D., and W. D. Plummer, Jr. 1998. Power and sample size calculations for studies involving linear regression. *Controlled Clinical Trials* 19: 589–601.

## Also see

[PSS] **power** — Power and sample-size analysis for hypothesis tests

[PSS] **power rsquared** — Power analysis for an  $R^2$  test in a multiple linear regression

[PSS] **power pcorr** — Power analysis for a partial-correlation test in a multiple linear regression

[PSS] **power, graph** — Graph results from the power command

[PSS] **power, table** — Produce table of results from the power command

[PSS] **Glossary**

[R] **regress** — Linear regression

[R] **test** — Test linear hypotheses after estimation