Title

stata.com

power oneway — Power analysis for one-way analysis of variance

Description

`power oneway` computes sample size, power, or effect size for one-way analysis of variance (ANOVA). By default, it computes sample size for given power and effect size. Alternatively, it can compute power for given sample size and effect size or compute effect size for given sample size, power, and number of groups. Also see [PSS-2] **power** for a general introduction to the `power` command using hypothesis tests.

Quick start

Sample size for a test of H_0: $\mu_1 = \mu_2 = \mu_3$ for cell means of 21, 19, and 18 and within-group variance of 16 with default power of 0.8 and significance level $\alpha = 0.05$

```
power oneway 21 19 18, varerror(16)
```

Same as above, but specify a between-group variance of 1.55 for three groups

```
power oneway, varerror(16) varmeans(1.55) ngroups(3)
```

Same as above, but specify that delta (Cohen's f), given by $\sqrt{\sigma^2_m/\sigma^2_e}$, equals 0.31

```
power oneway, ngroups(3) delta(.31)
```

Same as above, but specify delta equal to 0.2, 0.25, 0.3, 0.35, and 0.4

```
power oneway, ngroups(3) delta(.2(.05).4)
```

Same as above, but show results in a graph of sample size versus effect size (delta)

```
power oneway, ngroups(3) delta(.2(.05).4) graph
```

Sample size for contrast of H_0: $\mu_1 = (\mu_2 + \mu_3)/2$ or, equivalently, H_0: $-\mu_1 + 0.5\mu_2 + 0.5\mu_3 = 0$

```
power oneway 21 19 18, varerror(16) contrast(-1 0.5 0.5)
```

For an unbalanced design where the sample size in group 1 is 1.5 times the sample sizes for groups 2 and 3

```
power oneway 21 19 18, varerror(16) grweights(3 2 2)
```

Specify group means for 5 groups using matrix `means` and within-group variance of 34

```
matrix means = (21, 19, 19, 22, 27)
power oneway means, varerror(34)
```

Power for sample sizes of 99, 108, 117, and 126 with balanced group sizes

```
power oneway 21 19 18, varerror(16) n(99(9)126)
```

Same as above, but with sample sizes of 25, 25, and 50 for groups 1, 2 and 3, respectively

```
power oneway 21 19 18, varerror(16) n1(25) n2(25) n3(50)
```

Same as above, sample size allocations treated as parallel sets, each with total sample size 100

```
power oneway 21 19 18, varerror(16) n1(50 25 25) n2(25 50 25) ///
     n3(25 25 50) parallel
```

Effect size and target between-group variance for a sample size of 150 for three groups, power of 0.9, and $\alpha = 0.01$

```
power oneway, n(150) ngroups(3) power(.9) alpha(0.1)
```

Menu

Statistics > Power, precision, and sample size

Syntax

Compute sample size

power oneway *meanspec* [, power(*numlist*) *options*]

Compute power

power oneway *meanspec*, n(*numlist*) [*options*]

Compute effect size and target between-group variance

power oneway, n(*numlist*) power(*numlist*) ngroups(#) [varerror(*numlist*) *options*]

where *meanspec* is either a matrix *matname* containing group means or individual group means specified in a matrix form:

m_1 m_2 [m_3 ... m_J]

m_j, where $j = 1, 2, \ldots, J$, is the alternative group mean or the group mean under the alternative hypothesis for the jth group. Each m_j may be specified either as one number or as a list of values in parentheses; see **[U] 11.1.8 numlist**.

matname is the name of a Stata matrix with J columns containing values of alternative group means. Multiple rows are allowed, in which case each row corresponds to a different set of J group means or, equivalently, column j corresponds to *numlist* for the jth group mean.

options	Description
Main	
* alpha(*numlist*)	significance level; default is alpha(0.05)
* power(*numlist*)	power; default is power(0.8)
* beta(*numlist*)	probability of type II error; default is beta(0.2)
* n(*numlist*)	total sample size; required to compute power or effect size
nfractional	allow fractional sample sizes
* npergroup(*numlist*)	number of subjects per group; implies balanced design
* n#(*numlist*)	number of subjects in group #
grweights(*wgtspec*)	group weights; default is one for each group, meaning equal group sizes
ngroups(#)	number of groups
* varmeans(*numlist*)	variance of the group means or between-group variance
* varerror(*numlist*)	error (within-group) variance; default is varerror(1)
contrast(*contrastspec*)	contrast specification for group means
parallel	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	
[no]table[(*tablespec*)]	suppress table or display results as a table; see [PSS-2] **power, table**
saving(*filename* [, replace])	save the table data to *filename*; use replace to overwrite existing *filename*
Graph	
graph[(*graphopts*)]	graph results; see [PSS-2] **power, graph**
Iteration	
init(#)	initial value for sample size or effect size; default is to use a bisection algorithm to bound the solution
iterate(#)	maximum number of iterations; default is iterate(500)
tolerance(#)	parameter tolerance; default is tolerance(1e-12)
ftolerance(#)	function tolerance; default is ftolerance(1e-12)
[no]log	suppress or display iteration log
[no]dots	suppress or display iterations as dots
notitle	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.

collect is allowed; see [U] **11.1.10 Prefix commands**.

notitle does not appear in the dialog box.

wgtspec	Description
$\#_1$ $\#_2$ ... $\#_J$	J group weights. Weights must be positive and must be integers unless option `nfractional` is specified. Multiple values for each group weight $\#_j$ can be specified as a *numlist* enclosed in parentheses.
matname	matrix with J columns containing J group weights. Multiple rows are allowed, in which case each row corresponds to a different set of J weights or, equivalently, column j corresponds to a *numlist* for the jth weight.

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).

column	Description	Symbol
`alpha`	significance level	α
`power`	power	$1-\beta$
`beta`	type II error probability	β
`N`	total number of subjects	N
`N_per_group`	number of subjects per group	N/N_g
`N_avg`	average number of subjects per group	N_{avg}
`N#`	number of subjects in group #	$N_{\#}$
`delta`	effect size	δ
`N_g`	number of groups	N_g
`m#`	group mean #	$\mu_{\#}$
`Cm`	mean contrast	$C{\cdot}\mu$
`c0`	null mean contrast	c_0
`Var_m`	group means (between-group) variance	σ^2_m
`Var_Cm`	contrast variance	$\sigma^2_{C\mu}$
`Var_e`	error (within-group) variance	σ^2_e
`grwgt#`	group weight #	$w_{\#}$
`target`	target parameter; synonym for `Var_m` or `Var_Cm`	
`_all`	display all supported columns	

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

Column `N_per_group` is available and is shown in the default table only for balanced designs.

Columns `N_avg` and `N#` are shown in the default table only for unbalanced designs.

Columns `m#` are shown in the default table only if group means are specified.

Column `Var_m` is not shown in the default table if the `contrast()` option is specified.

Columns `Cm`, `c0`, and `Var_Cm` are shown in the default table only if the `contrast()` option is specified.

Columns `grwgt#` are not shown in the default table.

Options

Main

alpha(), power(), beta(), n(), nfractional; see [PSS-2] **power**.

npergroup(*numlist*) specifies the group size. Only positive integers are allowed. This option implies a balanced design. npergroup() cannot be specified with n(), n#(), or grweights().

n#(*numlist*) specifies the size of the #th group. Only positive integers are allowed. All group sizes must be specified. For example, all three options n1(), n2(), and n3() must be specified for a design with three groups. n#() cannot be specified with n(), npergroup(), or grweights().

grweights(*wgtspec*) specifies J group weights for an unbalanced design. The weights may be specified either as a list of values or as a matrix, and multiple sets of weights are allowed; see *wgtspec* for details. The weights must be positive and must also be integers unless the nfractional option is specified. grweights() cannot be specified with npergroup() or n#().

ngroups(#) specifies the number of groups. At least two groups must be specified. This option is required if *meanspec* is not specified. This option is also required for effect-size determination unless grweights() is specified.

varmeans(*numlist*) specifies the variance of the group means or the between-group variance. varmeans() cannot be specified with *meanspec* or contrast(), nor is it allowed with effect-size determination.

varerror(*numlist*) specifies the error (within-group) variance. The default is varerror(1).

contrast(*contrastspec*) specifies a contrast for group means containing J contrast coefficients that must sum to zero. *contrastspec* is

$\#_1\ \#_2\ \left[\#_3\ \ldots \#_J\right]$ [, null(*numlist*) <u>onesid</u>ed]

null(*numlist*) specifies the null or hypothesized value of the mean contrast. The default is null(0).

onesided requests a one-sided t test. The default is F test.

parallel; see [PSS-2] **power**.

Table

table, table(), notable; see [PSS-2] **power, table**.

saving(); see [PSS-2] **power**.

Graph

graph, graph(); see [PSS-2] **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 effect size δ for the effect-size determination. The default uses a bisection algorithm to bracket the solution.

iterate(), tolerance(), ftolerance(), log, nolog, dots, nodots; see [PSS-2] **power**.

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

notitle; see [PSS-2] **power**.

Remarks and examples

Remarks are presented under the following headings:

This entry describes the `power oneway` command and the methodology for power and sample-size analysis for one-way ANOVA. See [PSS-2] **Intro (power)** for a general introduction to power and sample-size analysis and [PSS-2] **power** for a general introduction to the `power` command using hypothesis tests.

Introduction

The comparison of multiple group means using one-way ANOVA models is a commonly used approach in a wide variety of statistical studies. The term "one way" refers to a single factor containing an arbitrary number of groups or levels. In what follows, we will assume that the factor levels are fixed. For two groups, the ANOVA model is equivalent to an unpaired two-sample t test; see [PSS-2] **power twomeans** for the respective power and sample-size analysis. One-way ANOVA uses an F test based on the ratio of the between-group variance to the within-group variance to compare means of multiple groups.

For example, consider a type of drug with three levels of dosage in treating a medical condition. An investigator may wish to test whether the mean response of the drug is the same across all levels of dosage. This is equivalent to testing the null hypothesis H_0: $\mu_1 = \mu_2 = \mu_3$ versus the alternative hypothesis H_a: $\mu_1 \neq \mu_2$ or $\mu_1 \neq \mu_3$ or $\mu_2 \neq \mu_3$; that is, at least one of the three group means is different from all the others. Rejection of the null hypothesis, however, does not provide any specific information about the individual group means. Therefore, in some cases, investigators may want to form a hypothesis for a mean contrast, $c = \sum_{j=1}^{k} c_j\mu_j$, a linear combination of group means where weights c_j sum to zero, and compare individual means by testing a hypothesis H_0: $c = c_0$ versus H_a: $c \neq c_0$.

This entry describes power and sample-size analysis for the inference about multiple population means using hypothesis testing based on one-way ANOVA. Specifically, we consider the null hypothesis H_0: $\mu_1 = \cdots = \mu_J$, which tests the equality of J group means against the alternative hypothesis H_a: $\mu_i \neq \mu_j$for some i, j. The test statistic for this hypothesis is based on the ratio of the between-group variance to the within-group variance and has an F distribution under the null hypothesis. The corresponding test is known as an overall F test, which tests the equality of multiple group means. This test is nondirectional.

For testing a single mean contrast, H_0: $c = c_0$ versus H_a: $c \neq c_0$, a test statistic is a function of the ratio of the contrast variance to the error (or within-group) variance, and either an F test or a t test can be used for a two-sided alternative. For a one-sided alternative, H_a: $c > c_0$ or H_a: $c < c_0$, only a t test can be used.

Power and sample-size computations use the distribution of the test statistic under the alternative hypothesis, which is a noncentral F distribution for the considered tests. Power is a function of the noncentrality parameter, and the noncentrality parameter is a function of the ratio of the standard deviation of the tested effect to the standard deviation of the errors. As such, the effect size for the overall F test is defined as the square root of the ratio of the between-group variance to the within-group variance. For testing a mean contrast, the effect size is defined as the square root of the contrast variance to the error or within-group variance.

Using power oneway

`power oneway` computes sample size, power, or effect size and target variance of the effect for a one-way fixed-effects analysis of variance. All computations are performed assuming a significance level of 0.05. You may change the significance level by specifying the `alpha()` option.

By default, the computations are performed for an overall F test testing the equality of all group means. The within-group or error variance for this test is assumed to be 1 but may be changed by specifying the `varerror()` option.

To compute the total sample size, you must specify the alternative *meanspec* 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 total sample size in the `n()` option and the alternative *meanspec*.

Instead of the alternative group means, you can specify the number of groups in the `ngroups()` option and the variance of the group means (or the between-group variance) in the `varmeans()` option when computing sample size or power.

To compute effect size, the square root of the ratio of the between-group variance to the error variance, and the target between-group variance, you must specify the total sample size in the `n()` option, the power in the `power()` option, and the number of groups in the `ngroups()` option.

To compute sample size or power for a test of a mean contrast, in addition to the respective options `power()` or `n()` as described above, you must specify the alternative *meanspec* and the corresponding contrast coefficients in the `contrast()` option. A contrast coefficient must be specified for each of the group means, and the specified coefficients must sum to zero. The null value for the specified contrast is assumed to be zero but may be changed by specifying the `null()` suboption within `contrast()`. The default test is an F test. You can instead request a one-sided t test by specifying the `onesided` suboption within `contrast()`. Effect-size determination is not available when testing a mean contrast.

For all the above computations, the error or within-group variance is assumed to be 1. You can change this value by specifying the `varerror()` option.

By default, all computations assume a balanced- or equal-allocation design. You can use the `grweights()` option to specify an unbalanced design for power, sample-size, or effect-size computations. For power and effect-size computations, you can specify individual group sizes in options `n1()`, `n2()`, and so on instead of a combination of `n()` and `grweights()` to accommodate an unbalanced design. For a balanced design, you can also specify the `npergroup()` option to specify a group size instead of a total sample size in `n()`.

In a one-way analysis of variance, sample size and effect size depend on the noncentrality parameter of the F distribution, and their estimation requires iteration. The default initial values are obtained from a bisection search that brackets the solution. If you desire, you may change this by specifying your own value in the `init()` option. See [PSS-2] **power** for the descriptions of other options that control the iteration procedure.

Alternative ways of specifying effect

To compute power or sample size, you must specify the magnitude of the effect desired to be detected by the test. With `power oneway`, you can do this by specifying either the individual alternative *meanspec*, for example,

`power oneway` m_1 m_2 ... m_J, ...

or the variance of J group means (between-group variance) and the number of groups J:

`power oneway, varmeans(#) ngroups(#)` [...]

You can also specify multiple values for variances in `varmeans()`.

There are multiple ways in which you can supply the group means to `power oneway`.

As we showed above, you may specify group means following the command line as

`power oneway` m_1 m_2 ... m_J, ...

At least two means must be specified.

Alternatively, you can define a Stata matrix as a row or a column vector and use it with `power oneway`. The dimension of the Stata matrix must be at least 2. For example,

`matrix define` *meanmat* = (m_1, m_2,..., m_J)

`power oneway` *meanmat*, ...

You can also specify multiple values or *numlist* for each of the group means in parentheses:

`power oneway` ($m_{1,1}$ $m_{1,2}$ ... m_{1,K_1}) ($m_{2,1}$ $m_{2,2}$ ... m_{2,K_2}) ..., ...

Each of the *numlist*s may contain different numbers of values, $K_1 \neq K_2 \neq \ldots \neq K_J$. `power oneway` will produce results for all possible combinations of values across *numlist*s. If instead you would like to treat each specification as a separate scenario, you can specify the `parallel` option.

Similarly, you can accommodate multiple sets of group means in a matrix form by adding a row for each specification. The columns of a matrix with multiple rows correspond to J group means, and values within each column j correspond to multiple specifications of the jth group mean or *numlist* for the jth group mean.

For example, the following two specifications are the same:

`power oneway` ($m_{1,1}$ $m_{1,2}$ $m_{1,3}$) ($m_{2,1}$ $m_{2,2}$ $m_{2,3}$), ...

and

`matrix define` *meanmat* = ($m_{1,1}$, $m_{2,1}$ \ $m_{1,2}$, $m_{2,2}$ \ $m_{1,3}$, $m_{2,3}$)

`power oneway` *meanmat*, ...

In the above specification, if you wish to specify a *numlist* only for the first group, you may define your matrix as

`matrix define` *meanmat* = ($m_{1,1}$, $m_{2,1}$ \ $m_{1,2}$, . \ $m_{1,3}$, .)

and the results of

`power oneway` *meanmat*, ...

will be the same as the results of

`power oneway` ($m_{1,1}$ $m_{1,2}$ $m_{1,3}$) $m_{2,1}$, ...

If you wish to treat the rows of *meanmat* as separate scenarios, you must specify the `parallel` option.

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

Computing sample size

To compute sample size, you must specify the alternative group means or their variance and, optionally, the power of the test in the `power()` option. A power of 0.8 is assumed if `power()` is not specified.

▷ Example 1: Sample size for a one-way ANOVA

Consider an example from van Belle et al. (2004, 367), where the authors report the results of a study of the association between cholesterol level and the number of diseased blood vessels, which indicate the presence of coronary artery disease, in patients undergoing coronary bypass surgery. Suppose we wish to plan a similar new study in which a cholesterol level is considered a risk factor that may be associated with the number of diseased blood vessels, our grouping variable. We consider three groups of subjects with 1, 2, or 3 numbers of diseased blood vessels. We would like to know how many subjects we need to observe in each group to detect differences between cholesterol levels across groups. Our projected cholesterol levels in three groups are 260, 289, and 295 mg/dL, respectively. Suppose that we anticipate the within-group or error variance is 4,900. To compute the total sample size for `power oneway`'s default setting of a balanced design with 5% significance level and 80% power, we type

```
. power oneway 260 289 295, varerror(4900)
Performing iteration ...
Estimated sample size for one-way ANOVA
F test for group effect
H0: delta = 0  versus  Ha: delta != 0
Study parameters:
        alpha =    0.0500
        power =    0.8000
        delta =    0.2183
          N_g =         3
           m1 =  260.0000
           m2 =  289.0000
           m3 =  295.0000
        Var_m =  233.5556
        Var_e = 4900.0000
Estimated sample sizes:
            N =       207
  N per group =        69
```

We find that a total sample of 207 subjects, 69 subjects per group, is required to detect a change in average cholesterol levels in at least 1 of the 3 groups for this study.

In addition to the specified and implied study parameters, `power oneway` reports the value of the effect size, `delta` $= \sqrt{233.556/4900} = 0.2183$, computed as a square root of the ratio between the variance of the group means `Var_m` and the error variance `Var_e`. The two variances are also often referred to as the between-group variance and the within-group variance, respectively. The effect size δ provides a unitless measure of the magnitude of an effect with a lower bound of zero meaning no effect. It corresponds to Cohen's effect-size measure f (Cohen 1988). Cohen's convention is that $f = 0.1$ means small effect size, $f = 0.25$ means medium effect size, and $f = 0.4$ means large effect size. According to this convention, the effect size considered in our example is medium.

◁

▷ Example 2: Alternative ways of specifying effect

Instead of specifying the alternative means as in example 1, we can specify the variance between them and the number of groups. From example 1, the variance between the group means was computed to be 233.5556. We specify this value in `varmeans()` as well as the number of groups in `ngroups()`:

```
. power oneway, varmeans(233.5556) ngroups(3) varerror(4900)
Estimated sample size for one-way ANOVA
F test for group effect
H0: delta = 0  versus  Ha: delta != 0
Study parameters:
        alpha =    0.0500
        power =    0.8000
        delta =    0.2183
          N_g =         3
        Var_m =  233.5556
        Var_e = 4900.0000
Estimated sample sizes:
            N =       207
  N per group =        69
```

We obtain the exact same results as in example 1.

Instead of specifying alternative means directly following the command line, we can define a matrix, say, `means`, containing these means and use it with `power oneway`:

```
. matrix define means = (260,289,295)
. power oneway means, varerror(4900)
  (output omitted)
```

You can verify that the results are identical to the previous results.

◁

▷ Example 3: Computing sample size for a mean contrast

Continuing with example 1, suppose we would like to test whether the average of the first two cholesterol levels is different from the cholesterol level of the third group. We construct the following contrast $(0.5, 0.5, -1)$ to test this hypothesis. To compute sample size, we specify the contrast coefficients in the `contrast()` option:

```
. power oneway 260 289 295, varerror(4900) contrast(.5 .5 -1)
Performing iteration ...
Estimated sample size for one-way ANOVA
F test for contrast of means
H0: Cm = 0  versus  Ha: Cm != 0
Study parameters:
          alpha =    0.0500
          power =    0.8000
          delta =    0.1381
            N_g =         3
             m1 =  260.0000
             m2 =  289.0000
             m3 =  295.0000
            C*m =  -20.5000
             c0 =    0.0000
         Var_Cm =   93.3889
          Var_e = 4900.0000
Estimated sample sizes:
              N =       414
    N per group =       138
```

The required sample size to achieve this study objective is 414 with 138 subjects per group.

For a test of a mean contrast, we can also test a directional hypothesis by specifying the `onesided` option within `contrast()`. In this case, the computation is based on the t test instead of the F test.

For example, to test whether the average of the first two cholesterol levels is less than the cholesterol level of the third group, we type

```
. power oneway 260 289 295, varerror(4900) contrast(.5 .5 -1, onesided)
Performing iteration ...
Estimated sample size for one-way ANOVA
t test for contrast of means
H0: Cm = 0  versus  Ha: Cm < 0
Study parameters:
          alpha =    0.0500
          power =    0.8000
          delta =   -0.1381
            N_g =         3
             m1 =  260.0000
             m2 =  289.0000
             m3 =  295.0000
            C*m =  -20.5000
             c0 =    0.0000
         Var_Cm =   93.3889
          Var_e = 4900.0000
Estimated sample sizes:
              N =       327
    N per group =       109
```

The results show that the required sample size reduces to a total of 327 subjects for this lower one-sided hypothesis.

The default null value for the contrast is zero, but you can change this by specifying `contrast()`'s suboption `null()`.

◁

▷ Example 4: Unbalanced design

Continuing with example 1, suppose we anticipate that the first group will have twice as many subjects as the second and the third groups. We can accommodate this unbalanced design by specifying the corresponding group weights in the `grweights()` option:

```
. power oneway 260 289 295, varerror(4900) grweights(2 1 1)
Performing iteration ...
Estimated sample size for one-way ANOVA
F test for group effect
H0: delta = 0  versus  Ha: delta != 0
Study parameters:
        alpha =    0.0500
        power =    0.8000
        delta =    0.2306
          N_g =         3
           m1 =  260.0000
           m2 =  289.0000
           m3 =  295.0000
        Var_m =  260.5000
        Var_e = 4900.0000
Estimated sample sizes:
            N =       188
    Average N =   62.6667
           N1 =        94
           N2 =        47
           N3 =        47
```

The required total sample size for this unbalanced design is 188 with 94 subjects in the first group and 47 subjects in the second and third groups. The average number of subjects per group is 62.67.

We can compute results for multiple sets of group weights. The specification of group weights within `grweights()` is exactly the same as the specification of group means described in *Alternative ways of specifying effect*. Suppose that we would like to compute sample sizes for two unbalanced designs. The first design has twice as many subjects in the first group as the other two groups. The second design has the first two groups with twice as many subjects as the third group. We specify multiple group weights for the first and second groups in parentheses. We also specify the `parallel` option to treat multiple weight values in parallel instead of computing results for all possible combinations of these values, which would have been done by default.

```
. local tabcols alpha power N N1 N2 N3 grwgt1 grwgt2 grwgt3 Var_m Var_e
. power oneway 260 289 295, varerror(4900) grweights((2 2) (1 2) 1) parallel
> table(`tabcols', formats("%6.0g"))
Performing iteration ...
Estimated sample size for one-way ANOVA
F test for group effect
H0: delta = 0  versus  Ha: delta != 0
```

alpha	power	N	N1	N2	N3	grwgt1	grwgt2	grwgt3	Var_m	Var_e
.05	.8	188	94	47	47	2	1	1	260.5	4900
.05	.8	205	82	82	41	2	2	1	235.4	4900

The default table does not include group weights, so we request a table with custom columns containing group weights via `table()`. We also request a smaller format to make the table more compact.

◁

Computing power

To compute power, you must specify the total sample size in the `n()` option and the alternative group means or their variance.

▷ Example 5: Power for a one-way ANOVA

Continuing with example 1, suppose that we anticipate obtaining a total sample of 300 subjects. To compute the corresponding power, we specify the sample size of 300 in `n()`:

```
. power oneway 260 289 295, n(300) varerror(4900)
Estimated power for one-way ANOVA
F test for group effect
H0: delta = 0  versus  Ha: delta != 0
Study parameters:
        alpha =    0.0500
            N =       300
  N per group =       100
        delta =    0.2183
          N_g =         3
           m1 =  260.0000
           m2 =  289.0000
           m3 =  295.0000
        Var_m =  233.5556
        Var_e = 4900.0000
Estimated power:
        power =    0.9308
```

The power increases to 93.08% for a larger sample of 300 subjects.

◁

▷ Example 6: Multiple values of study parameters

Continuing with example 5, we may want to check powers for several sample sizes. We simply list multiple sample-size values in `n()`:

```
. power oneway 260 289 295, n(100 200 300) varerror(4900)
> table(, labels(N_per_group "N/N_g") formats("%6.2g"))
Estimated power for one-way ANOVA
F test for group effect
H0: delta = 0  versus  Ha: delta != 0
```

alpha	power	N	N/N_g	delta	N_g	m1	m2	m3	Var_m	Var_e
.05	.47	100	33	.22	3	260	289	295	234	4900
.05	.78	200	66	.22	3	260	289	295	234	4900
.05	.93	300	100	.22	3	260	289	295	234	4900

To shorten our default table, we specified a shorter label for the `N_per_group` column and reduced the default display format for all table columns by specifying the corresponding options within the `table()` option.

We can compute results for multiple values of group means. For example, to see how power changes when the first group mean takes values of 245, 260, and 280, we specify these values in parentheses:

```
. power oneway (245 260 280) 289 295, n(300) varerror(4900)
> table(, labels(N_per_group "N/N_g") formats("%6.2g"))
Estimated power for one-way ANOVA
F test for group effect
H0: delta = 0  versus  Ha: delta != 0
```

alpha	power	N	N/N_g	delta	N_g	m1	m2	m3	Var_m	Var_e
.05	1	300	100	.32	3	245	289	295	497	4900
.05	.93	300	100	.22	3	260	289	295	234	4900
.05	.25	300	100	.088	3	280	289	295	38	4900

We can compute results for a combination of multiple sample sizes and multiple mean values or a combination of multiple values of other study parameters.

For multiple values of parameters, the results are automatically displayed in a table, as we see above. For more examples of tables, see [PSS-2] **power, table**. If you wish to produce a power plot, see [PSS-2] **power, graph**.

◁

Computing effect size and between-group variance

Sometimes, we may be interested in determining the smallest effect that yields a statistically significant result for prespecified sample size and power. In this case, power, sample size, and the number of groups must be specified.

The effect size is defined as a square root of the ratio of the variance of the tested effect, for example, the between-group variance, to the error variance. Both the effect size and the target between-group variance are computed.

The effect-size determination is not available for testing a mean contrast.

▷ Example 7: Effect size for a one-way analysis of variance

Continuing with example 5, we now want to compute the effect size that can be detected for a sample of 300 subjects and a power of 80%. We specify both parameters in the respective options. For the effect-size determination, we must also specify the number of groups in `ngroups()`:

```
. power oneway, varerror(4900) n(300) power(0.80) ngroups(3)
Performing iteration ...
Estimated between-group variance for one-way ANOVA
F test for group effect
H0: delta = 0  versus  Ha: delta != 0
Study parameters:
        alpha =    0.0500
        power =    0.8000
            N =       300
  N per group =       100
          N_g =         3
        Var_e = 4900.0000
Estimated effect size and between-group variance:
        delta =    0.1801
        Var_m =  158.9648
```

For a larger sample size, given the same power, we can detect a smaller effect size, 0.18, compared with the effect size of 0.22 from example 1. The corresponding estimate of the between-group variance is 158.96, given the error variance of 4,900.

◁

Testing hypotheses about multiple group means

There are several ways in which you can compare group means of a single factor on the collected data. Two commonly used commands to do this are `oneway` (or `anova`) and `contrast`. We demonstrate a quick use of these commands here using the systolic blood pressure example; see [R] **oneway** and [R] **contrast** for more examples. Also see [R] **anova** for general ANOVA models.

▷ Example 8: One-way ANOVA

Consider `systolic.dta` containing 58 patients undergoing 4 different drug treatments for reducing systolic blood pressure. The `systolic` variable records the change in systolic blood pressure and the `drug` variable records four treatment levels. We would like to test whether the average change in systolic blood pressure is the same for all treatments. We use `oneway` to do this:

```
. use https://www.stata-press.com/data/r18/systolic
(Systolic blood pressure data)

. oneway systolic drug

                        Analysis of variance
    Source              SS         df      MS            F     Prob > F
------------------------------------------------------------------------
Between groups      3133.23851      3   1044.41284      9.09     0.0001
 Within groups      6206.91667     54   114.942901
------------------------------------------------------------------------
    Total           9340.15517     57   163.862371

Bartlett's equal-variances test: chi2(3) =   1.0063    Prob>chi2 = 0.800
```

We reject the null hypothesis that all treatment means are equal at the 5% significance level; the p-value is less than 0.0001.

Suppose we wish to design a new similar study. We use the estimates from this study to perform a sample-size analysis for our new study. First, we estimate the means of systolic blood pressure for different treatment levels:

```
. mean systolic, over(drug)

Mean estimation                               Number of obs = 58

-----------------------------------------------------------------
                |       Mean   Std. err.     [95% conf. interval]
----------------+------------------------------------------------
c.systolic@drug |
              1 |   26.06667   3.014989      20.02926    32.10408
              2 |   25.53333   2.999788      19.52636    31.54031
              3 |       8.75   2.892323      2.958224    14.54178
              4 |       13.5   2.330951      8.832351    18.16765
-----------------------------------------------------------------
```

From the `oneway` output, the estimate of the error variance is roughly 115. Second, we specify the means and the error variance with `power oneway` and compute the required sample size for a balanced design assuming 5% significance level and 90% power.

```
. power oneway 26.07 25.53 8.75 13.5, varerror(115) power(.9)
Performing iteration ...
Estimated sample size for one-way ANOVA
F test for group effect
H0: delta = 0  versus  Ha: delta != 0
Study parameters:
        alpha =    0.0500
        power =    0.9000
        delta =    0.7021
          N_g =         4
           m1 =   26.0700
           m2 =   25.5300
           m3 =    8.7500
           m4 =   13.5000
        Var_m =   56.6957
        Var_e =  115.0000
Estimated sample sizes:
            N =        36
  N per group =         9
```

The effect size for this ANOVA model specification is rather large, 0.7021. So we need only 36 subjects, 9 per group, to detect the effect of this magnitude with 90% power.

Suppose that in addition to testing the overall equality of treatment means, we are interested in testing a specific hypothesis of whether the average of the first two treatment means is equal to the average of the last two treatment means.

To perform this test on the collected data, we can use the `contrast` command. The `contrast` command is not available after `oneway`, so we repeat our one-way ANOVA analysis using the `anova` command before using `contrast`; see [R] **contrast** for details.

```
. anova systolic i.drug

                         Number of obs =         58    R-squared     =  0.3355
                         Root MSE      =    10.7211    Adj R-squared =  0.2985

                  Source | Partial SS         df         MS        F    Prob>F
              -----------+----------------------------------------------------
                   Model |  3133.2385          3   1044.4128      9.09  0.0001
                         |
                    drug |  3133.2385          3   1044.4128      9.09  0.0001
                         |
                Residual |  6206.9167         54    114.9429
              -----------+----------------------------------------------------
                   Total |  9340.1552         57   163.86237
```

```
. contrast {drug .5 .5 -.5 -.5}

Contrasts of marginal linear predictions

Margins: asbalanced

------------------------------------------------
             |         df           F        P>F
-------------+----------------------------------
        drug |          1       26.85     0.0000
             |
 Denominator |         54
------------------------------------------------

--------------------------------------------------------------
             |   Contrast   Std. err.     [95% conf. interval]
-------------+------------------------------------------------
        drug |
        (1)  |     14.675   2.832324      8.996533    20.35347
--------------------------------------------------------------
```

As with the overall equality test, we find statistical evidence to reject this hypothesis as well.

To compute the required sample size for this hypothesis, we specify the contrast coefficients in the `contrast()` option of `power oneway`:

```
. power oneway 26.07 25.53 8.75 13.5, varerror(115) power(.9)
> contrast(.5 .5 -.5 -.5)

Performing iteration ...

Estimated sample size for one-way ANOVA
F test for contrast of means
H0: Cm = 0  versus  Ha: Cm != 0

Study parameters:

        alpha =    0.0500
        power =    0.9000
        delta =    0.6842
          N_g =         4
           m1 =   26.0700
           m2 =   25.5300
           m3 =    8.7500
           m4 =   13.5000
          C*m =   14.6750
           c0 =    0.0000
       Var_Cm =   53.8389
        Var_e =  115.0000

Estimated sample sizes:

            N =        28
  N per group =         7
```

The required sample size is 28 subjects with 7 subjects per group, which is smaller than the required sample size computed earlier for the overall test of the equality of means.

◁

Video examples

Sample-size calculation for one-way analysis of variance

Power calculation for one-way analysis of variance

Minimum detectable effect size for one-way analysis of variance

Stored results

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

Scalars

`r(alpha)`	significance level
`r(power)`	power
`r(beta)`	probability of a type II error
`r(delta)`	effect size
`r(N)`	total sample size
`r(N_a)`	actual sample size
`r(N_avg)`	average sample size
`r(N#)`	number of subjects in group #
`r(N_per_group)`	number of subjects per group
`r(N_g)`	number of groups
`r(nfractional)`	1 if `nfractional` is specified, 0 otherwise
`r(balanced)`	1 for a balanced design, 0 otherwise
`r(grwgt#)`	group weight #
`r(onesided)`	1 for a one-sided test of a mean contrast, 0 otherwise
`r(m#)`	group mean #
`r(Cm)`	mean contrast
`r(c0)`	null mean contrast
`r(Var_m)`	group-means (between-group) variance
`r(Var_Cm)`	contrast variance
`r(Var_e)`	error (within-group) variance
`r(separator)`	number of lines between separator lines in the table
`r(divider)`	1 if `divider` is requested in the table, 0 otherwise
`r(init)`	initial value for the sample size or effect size
`r(maxiter)`	maximum number of iterations
`r(iter)`	number of iterations performed
`r(tolerance)`	requested parameter tolerance
`r(deltax)`	final parameter tolerance achieved
`r(ftolerance)`	requested distance of the objective function from zero
`r(function)`	final distance of the objective function from zero
`r(converged)`	1 if iteration algorithm converged, 0 otherwise

Macros

`r(type)`	`test`
`r(method)`	`oneway`
`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 single factor A with J groups or levels, where each level comprises n_j observations for $j = 1, \ldots, J$. The total number of observations is $n = \sum_{j=1}^{J} n_j$. Let Y_{ij} denote the response for the jth level of the ith individual and μ_j denote the factor-level or group means. Individual responses from J populations are assumed to be normally distributed with mean μ_j and a constant variance σ_e^2.

The hypothesis of an overall F test of the equality of group means is

$$H_0\colon \mu_1 = \cdots = \mu_J$$

versus

$$H_a\colon \mu_j\text{'s are not all equal} \qquad (1)$$

The hypothesis for a test of a mean contrast is

$$H_0\colon \sum_{j=1}^{J} c_j \mu_j = c_0 \tag{2}$$

versus a two-sided $H_a\colon \sum_{j=1}^{J} c_j \mu_j \neq c_0$, upper one-sided $H_a\colon \sum_{j=1}^{J} c_j \mu_j > c_0$, and lower one-sided $H_a\colon \sum_{j=1}^{J} c_j \mu_j < c_0$, where c_j's are contrast coefficients such that $\sum_{i=1}^{J} c_j = 0$ and c_0 is a null value of a mean constant. The two-sided hypothesis is tested using an F test, and one-sided hypotheses are tested using a t test.

Hypotheses (1) and (2) can be tested in a general linear model framework. Consider a linear model

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \boldsymbol{\epsilon}$$

where $\mathbf{y}$ is an $n \times 1$ vector of observations, $\mathbf{X}$ is an $n \times p$ matrix of predictors, $\mathbf{b}$ is a $p \times 1$ vector of unknown and fixed coefficients, and $\boldsymbol{\epsilon}$ is an $n \times 1$ vector of error terms that are independent and identically distributed as $N(0, \sigma_e^2)$.

For the one-way model, $p = J$ and the contents of $\mathbf{b} = (b_1, \ldots, b_J)'$ are the μ_j, $j = 1, \ldots, J$.

A general linear hypothesis in this framework is given by

$$\mathbf{C}\mathbf{b} = \mathbf{c_0}$$

where $\mathbf{C}$ is a $\nu \times p$ matrix with $\text{rank}(\mathbf{C}) = \nu \leq p$, and $\mathbf{c_0}$ is a vector of constants. For an overall test of the means in (1), $\mathbf{c_0} = 0$. The estimates of $\mathbf{b}$ and σ_e^2, respectively, are given by

$$\widehat{\mathbf{b}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

$$\widehat{\sigma}_e^2 = (\mathbf{y} - \mathbf{X}\widehat{\mathbf{b}})'(\mathbf{y} - \mathbf{X}\widehat{\mathbf{b}})/(n - p)$$

A general test statistic for testing hypotheses (1) and (2) is given by

$$F_C = \frac{SS_C}{(p-1)\widehat{\sigma}_e^2} \tag{3}$$

where $SS_C = (\mathbf{C}\widehat{\mathbf{b}} - \mathbf{c_0})'\{\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}'\}^{-1}(\mathbf{C}\widehat{\mathbf{b}} - \mathbf{c_0})$. Let α be the significance level, β be the probability of a type II error, and $F_{p-1,n-p,1-\alpha}$ be the $(1-\alpha)$th quantile of an F distribution with $p-1$ numerator and $n-p$ denominator degrees of freedom. We reject the null hypothesis if we observe a statistic $F_C > F_{p-1,n-p,1-\alpha}$.

The test statistic in (3) under the alternative hypothesis is distributed as a noncentral F distribution with $p-1$ numerator and $n-p$ denominator degrees of freedom with a noncentrality parameter λ given by

$$\begin{aligned}
\lambda &= (\mathbf{C}\mathbf{b} - \mathbf{c_0})'\{\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}'\}^{-1}(\mathbf{C}\mathbf{b} - \mathbf{c_0})/\sigma_e^2 \\
&= n(\mathbf{C}\mathbf{b} - \mathbf{c_0})'\{\mathbf{C}(\ddot{\mathbf{X}}'\mathbf{W}\ddot{\mathbf{X}})^{-1}\mathbf{C}'\}^{-1}(\mathbf{C}\mathbf{b} - \mathbf{c_0})/\sigma_e^2 \\
&= n\delta^2
\end{aligned}$$

where the matrix $\ddot{\mathbf{X}}$ contains the unique rows of $\mathbf{X}$ and $\mathbf{W} = \text{diag}(w_1, \ldots, w_p)$. We define δ as the effect size.

For the one-way design, the dimension of $\ddot{\mathbf{X}}$ is $J \times J$. The weights are formally $w_j = n_j/n$ but can also be expressed in terms of the group weights (specified in `grweights()`), normalized by the sum of the group weights, making w_j independent of n. Specifically, let the group weights be denoted $\widetilde{w}_j$, then define a cell sample size multiplier as $n_c = n/\sum_j \widetilde{w}_j$ so that $n_j = n_c\widetilde{w}_j$. The cell-means parameterization simplifies $\ddot{\mathbf{X}}$ to the identity matrix, $\mathbf{I}_J$.

See O'Brien and Muller (1993) for details.

The power of the overall F test in (1) is given by

$$1 - \beta = F_{\nu,n-p,\lambda}\left(F_{\nu,n-p,1-\alpha}\right) \tag{4}$$

where $F_{\cdot,\cdot,\lambda}\left(\cdot\right)$ is the cdf of a noncentral F distribution.

Total sample size and effect size are obtained by iteratively solving the nonlinear equation (4). When the `grweights()` option is specified, a constant multiplier n_c is computed and rounded to an integer unless the `nfractional` option is specified. The group sizes are then computed as $\widetilde{w}_j n_c$. The actual sample size, `N_a`, is the sum of the group sizes.

See Kutner et al. (2005) for details.

The power of the test (2) for a mean contrast is given by

$$1 - \beta = \begin{cases} 1 - T_{n-1,\widetilde{\lambda}}\left(t_{n-1,1-\alpha}\right) & \text{for an upper one-sided test} \\ T_{n-1,\widetilde{\lambda}}\left(-t_{n-1,1-\alpha}\right) & \text{for a lower one-sided test} \\ F_{1,n-p,\lambda}\left(F_{1,n-p,1-\alpha}\right) & \text{for a two-sided test} \end{cases} \tag{5}$$

where $T_{\cdot,\widetilde{\lambda}}\left(\cdot\right)$ is the cumulative of a noncentral Student's t distribution with the noncentrality parameter $\widetilde{\lambda}$ given by

$$\begin{aligned} \widetilde{\lambda} &= \sqrt{n}(\mathbf{Cb} - \mathbf{c_0})'\sqrt{\left\{\mathbf{C}(\ddot{\mathbf{X}}'\mathbf{W}\ddot{\mathbf{X}})^{-1}\mathbf{C}'\right\}^{-1}/\sigma_e^2} \\ &= \sqrt{n}\widetilde{\delta} \end{aligned}$$

Sample size is obtained by iteratively solving the nonlinear equation (5).

References

Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: Erlbaum.

Kutner, M. H., C. J. Nachtsheim, J. Neter, and W. Li. 2005. *Applied Linear Statistical Models*. 5th ed. New York: McGraw–Hill/Irwin.

O'Brien, R. G., and K. E. Muller. 1993. Unified power analysis for t-tests through multivariate hypotheses. In *Applied Analysis of Variance in Behavioral Science*, ed. L. K. Edwards, 297–344. New York: Dekker.

van Belle, G., L. D. Fisher, P. J. Heagerty, and T. S. Lumley. 2004. *Biostatistics: A Methodology for the Health Sciences*. 2nd ed. New York: Wiley.

Also see

[PSS-2] **power** — Power and sample-size analysis for hypothesis tests

[PSS-2] **power repeated** — Power analysis for repeated-measures analysis of variance

[PSS-2] **power twomeans** — Power analysis for a two-sample means test

[PSS-2] **power twoway** — Power analysis for two-way analysis of variance

[PSS-2] **power, graph** — Graph results from the power command

[PSS-2] **power, table** — Produce table of results from the power command

[PSS-5] **Glossary**

[R] **anova** — Analysis of variance and covariance

[R] **contrast** — Contrasts and linear hypothesis tests after estimation

[R] **oneway** — One-way analysis of variance

For suggested citations, see the FAQ on citing Stata documentation.