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:

```
m1 m2 [m3 ... mJ] 
```

$m_j$, where $j = 1, 2, \ldots, J$, is the alternative group mean or the group mean under the alternative hypothesis for the $j$th 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 a `numlist` for the $j$th group mean.
## options

<table>
<thead>
<tr>
<th>Main</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>*alpha(numlist)</td>
<td>significance level; default is alpha(0.05)</td>
</tr>
<tr>
<td>*power(numlist)</td>
<td>power; default is power(0.8)</td>
</tr>
<tr>
<td>*beta(numlist)</td>
<td>probability of type II error; default is beta(0.2)</td>
</tr>
<tr>
<td>*n(numlist)</td>
<td>total sample size; required to compute power or effect size</td>
</tr>
<tr>
<td></td>
<td>allow fractional sample sizes</td>
</tr>
<tr>
<td>nfractional</td>
<td></td>
</tr>
<tr>
<td>*npergroup(numlist)</td>
<td>number of subjects per group; implies balanced design</td>
</tr>
<tr>
<td>*n#(numlist)</td>
<td>number of subjects in group #</td>
</tr>
<tr>
<td>grweights(wgtspec)</td>
<td>group weights; default is one for each group, meaning equal-group sizes</td>
</tr>
<tr>
<td>ngroups(#)</td>
<td>number of groups</td>
</tr>
<tr>
<td>*varmeans(numlist)</td>
<td>variance of the group means or between-group variance</td>
</tr>
<tr>
<td>*varerror(numlist)</td>
<td>error (within-group) variance; default is varerror(1)</td>
</tr>
<tr>
<td>contrast(contrastspec)</td>
<td>contrast specification for group means</td>
</tr>
<tr>
<td>parallel</td>
<td>treat number lists in starred options as parallel when multiple values per option are specified (do not enumerate all possible combinations of values)</td>
</tr>
</tbody>
</table>

### Table

- `[no]table [tablespec]` suppress table or display results as a table; see [PSS] power, table
- `saving(filename [, replace])` save the table data to `filename`; use replace to overwrite existing `filename`

### Graph

- `graph [graphopts]` graph results; see [PSS] power, graph

### Iteration

- `init(#)` initial value for the sample size or the 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

*Starred options may be specified either as one number or as a list of values; see [U] 11.1.8 numlist.*  
`notitle` does not appear in the dialog box.


wgtspec  

Description

\#_1 \#_2 \ldots \#_J  

J group weights. Weights must be positive and must be integers unless option \texttt{n\textunderscore fractiona\textunderscore l} is specified. Multiple values for each group weight \#_j can be specified as a \texttt{numlist} enclosed in parentheses.

\texttt{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 \texttt{numlist} for the jth weight.

where \texttt{tablespec} is

\texttt{column[ :label] [ column[ :label] [ . . . ] ] [ , tableopts ]}

\texttt{column} is one of the columns defined below, and \texttt{label} is a column label (may contain quotes and compound quotes).

\begin{tabular}{lll}
\texttt{column} & Description & Symbol \\
\hline
\texttt{alpha} & significance level & \(\alpha\) \\
\texttt{power} & power & \(1 - \beta\) \\
\texttt{beta} & type II error probability & \(\beta\) \\
\texttt{N} & total number of subjects & \(N\) \\
\texttt{N\_per\_group} & number of subjects per group & \(N/N_g\) \\
\texttt{N\_avg} & average number of subjects per group & \(N_{\text{avg}}\) \\
\texttt{N\#\_1,\#\_2} & number of subjects in group (#_1, #_2) & \(N_{\#_1,\#_2}\) \\
\texttt{delta} & effect size & \(\delta\) \\
\texttt{N\_g} & number of groups & \(N_g\) \\
\texttt{m\#} & group mean \# & \(\mu_\#\) \\
\texttt{Cm} & mean contrast & \(C \cdot \mu\) \\
\texttt{c0} & null mean contrast & \(c_0\) \\
\texttt{Var\_m} & group means (between-group) variance & \(\sigma^2_m\) \\
\texttt{Var\_Cm} & contrast variance & \(\sigma^2_{C\mu}\) \\
\texttt{Var\_e} & error (within-group) variance & \(\sigma^2_e\) \\
\texttt{grwgt\#} & group weight \# & \(w_\#\) \\
\texttt{target} & target parameter; synonym for Var\_m or Var\_Cm & \\
\texttt{\_all} & display all supported columns & \\
\end{tabular}

Column \texttt{beta} is shown in the default table in place of column \texttt{power} if specified.

Column \texttt{N\_per\_group} is available and is shown in the default table only for balanced designs.

Columns \texttt{N\_avg} and \texttt{N\#} are shown in the default table only for unbalanced designs.

Columns \texttt{m\#} are shown in the default table only if group means are specified.

Column \texttt{Var\_m} is not shown in the default table if the \texttt{contrast()} option is specified.

Columns \texttt{Cm}, \texttt{c0}, and \texttt{Var\_Cm} are shown in the default table only if the \texttt{contrast()} option is specified.

Columns \texttt{grwgt\#} are not shown in the default table.

\section*{Menu}

Statistics \(\rightarrow\) Power and sample size
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] power for a general introduction to the power command using hypothesis tests.

Options

alpha(), power(), beta(), n(), nfractional; see [PSS] 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). If varerror() is not specified with effect-size determination, only effect size $\delta$ is reported; otherwise, the estimated between-group variance is also reported.

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

$#_1$ $#_2$ $[$ $#_3$ . . . $#_J$ $]$ [, null(numlist) onesided]

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] power.

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

saving(); see [PSS] power.

graph, graph(); see [PSS] power, graph. Also see the column table for a list of symbols used by the graphs.
power oneway — Power analysis for one-way analysis of variance

init(#) specifies the initial value of the sample size for the sample-size determination or the initial value of the effect size \( \delta \) for the effect-size determination. The default uses a bisection algorithm to bracket the solution.

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

The following option is available with power oneway 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 oneway
Alternative ways of specifying effect
Computing sample size
Computing power
Computing effect size and between-group variance
Testing hypotheses about multiple group means

This entry describes the power oneway command and the methodology for power and sample-size analysis for one-way ANOVA. 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

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] 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, 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. If the `varerror()` option is also specified, the target between-group variance is reported in addition to the effect size.

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] 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, varemeans(#) ngroups(#) ...
```

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

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 a `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 `numlists` may contain different numbers of values, $K_1 \neq K_2 \neq ... \neq K_J$. `power oneway` will produce results for all possible combinations of values across `numlists`. 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 $j$th group mean or a `numlist` for the $j$th 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 = ($m_{1,1}$, $m_{1,2}$ \ $m_{1,3}$, $m_{2,1}$ \ $m_{2,2}$ \ $m_{2,3}$)
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 = ($m_{1,1}$, $m_{1,2}$ \ $m_{1,3}$, . \ $m_{1,3}$, .)
```
and the results of

```
power oneway meanmat, ...
```

will be the same as the results of

```
power oneway (m1,1 m1,2 m1,3) m2,1, ...
```

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, Fisher, Heagerty, and Lumley (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
Ho: 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 $\text{Var}_m$ and the error variance $\text{Var}_e$. The two variances are also often referred to as the between-group variance and the within-group variance, respectively. The effect size $\delta$ provides a unit-less 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
Ho: $\delta = 0$ versus Ha: $\delta \neq 0$

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)
```

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

### Example 3: Computing sample size for a mean contrast

Continuing 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:

```
. matrix define means = (260, 289, 295)
. power oneway means, varerror(4900) contrast((0.5, 0.5, -1))
```

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

Estimated sample size for one-way ANOVA
F test for contrast of means
Ho: 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

Performing iteration ...

Estimated sample size for one-way ANOVA
t test for contrast of means
Ho: 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 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:

```plaintext
. 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
Ho: 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.

```plaintext
. 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
Ho: delta = 0 versus Ha: delta != 0

<table>
<thead>
<tr>
<th>alpha</th>
<th>power</th>
<th>N</th>
<th>N1</th>
<th>N2</th>
<th>N3</th>
<th>grwgt1</th>
<th>grwgt2</th>
<th>grwgt3</th>
<th>Var_m</th>
<th>Var_e</th>
</tr>
</thead>
<tbody>
<tr>
<td>.05</td>
<td>.8</td>
<td>188</td>
<td>94</td>
<td>47</td>
<td>47</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>260.5</td>
<td>4900</td>
</tr>
<tr>
<td>.05</td>
<td>.8</td>
<td>205</td>
<td>82</td>
<td>82</td>
<td>41</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>235.4</td>
<td>4900</td>
</tr>
</tbody>
</table>
```

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 \texttt{n()} option and the alternative group means or their variance.

\section*{Example 5: Power for a one-way ANOVA}

Continuing 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 \texttt{n()}:\[
.\text{power oneway 260 289 295, n(300) varerror(4900)}
\]
Estimated power for one-way ANOVA
F test for group effect
\texttt{Ho: delta = 0} versus \texttt{Ha: delta != 0}
\begin{verbatim}
  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
\end{verbatim}
Estimated power:
\begin{verbatim}
  power = 0.9308
\end{verbatim}
The power increases to 93.08\% for a larger sample of 300 subjects.

\section*{Example 6: Multiple values of study parameters}

Continuing example 5, we may want to check powers for several sample sizes. We simply list multiple sample-size values in \texttt{n()}:\[
.\text{power oneway 260 289 295, n(100 200 300) varerror(4900)}
\]
\texttt{> table(, labels(N_per_group "N/N_g") formats("%6.2g"))}
Estimated power for one-way ANOVA
F test for group effect
\texttt{Ho: delta = 0} versus \texttt{Ha: delta != 0}
\begin{verbatim}
\begin{tabular}{cccccccccc}
  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 \\
\end{tabular}
\end{verbatim}
To shorten our default table, we specified a shorter label for the \texttt{N_per_group} column and reduced the default display format for all table columns by specifying the corresponding options within the \texttt{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
Ho: delta = 0 versus Ha: delta ! = 0

<table>
<thead>
<tr>
<th>alpha</th>
<th>power</th>
<th>N</th>
<th>N/N_g</th>
<th>delta</th>
<th>N_g</th>
<th>m1</th>
<th>m2</th>
<th>m3</th>
<th>Var_m</th>
<th>Var_e</th>
</tr>
</thead>
<tbody>
<tr>
<td>.05</td>
<td>1</td>
<td>300</td>
<td>100</td>
<td>.32</td>
<td>3</td>
<td>245</td>
<td>289</td>
<td>295</td>
<td>497</td>
<td>4900</td>
</tr>
<tr>
<td>.05</td>
<td>.93</td>
<td>300</td>
<td>100</td>
<td>.22</td>
<td>3</td>
<td>260</td>
<td>289</td>
<td>295</td>
<td>234</td>
<td>4900</td>
</tr>
<tr>
<td>.05</td>
<td>.25</td>
<td>300</td>
<td>100</td>
<td>.088</td>
<td>3</td>
<td>280</td>
<td>289</td>
<td>295</td>
<td>38</td>
<td>4900</td>
</tr>
</tbody>
</table>

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] power, table. If you wish to produce a power plot, see [PSS] 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. The effect size is computed by default. If the error variance is specified in the varerror() option, the between-group variance is also reported.

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

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

Continuing 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
Ho: delta = 0 versus Ha: delta ! = 0

Study parameters:

<table>
<thead>
<tr>
<th>alpha</th>
<th>power</th>
<th>N</th>
<th>N per group</th>
<th>Var_e</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0500</td>
<td>0.8000</td>
<td>300</td>
<td>100</td>
<td>4900.0000</td>
</tr>
</tbody>
</table>

Estimated effect size and between-group variance:

<table>
<thead>
<tr>
<th>delta</th>
<th>Var_m</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1801</td>
<td>158.9648</td>
</tr>
</tbody>
</table>
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 http://www.stata-press.com/data/r13/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 test for equal variances: 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
1: drug = 1
2: drug = 2
3: drug = 3
4: drug = 4
```

<table>
<thead>
<tr>
<th>Over</th>
<th>Mean</th>
<th>Std. Err.</th>
<th>[95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>systolic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>26.06667</td>
<td>3.014989</td>
<td>20.02926 32.10408</td>
</tr>
<tr>
<td>2</td>
<td>25.53333</td>
<td>2.999788</td>
<td>19.52636 31.54031</td>
</tr>
<tr>
<td>3</td>
<td>8.75</td>
<td>2.892323</td>
<td>2.958224 14.54178</td>
</tr>
<tr>
<td>4</td>
<td>13.5</td>
<td>2.330951</td>
<td>8.832351 18.16765</td>
</tr>
</tbody>
</table>

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.
Performing iteration ...

Estimated sample size for one-way ANOVA
F test for group effect
Ho: 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           3133.23851   3     1044.41284   9.09  0.0001
drug            3133.23851   3     1044.41284   9.09  0.0001
Residual        6206.91667  54     114.942901
Total           9340.15517  57    163.862371

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

Contrasts of marginal linear predictions
Margins         : asbalanced

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>F</th>
<th>P&gt;F</th>
</tr>
</thead>
<tbody>
<tr>
<td>drug</td>
<td>1</td>
<td>26.85</td>
<td>0.0000</td>
</tr>
<tr>
<td>Denominator</td>
<td>54</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Contrast</th>
<th>Std. Err.</th>
<th>[95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>drug (1)</td>
<td>14.675</td>
<td>2.832324</td>
<td>8.996533 20.35347</td>
</tr>
</tbody>
</table>
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
Ho: Cm = 0 versus Ha: Cm != 0

Study parameters:

\[
\begin{align*}
\alpha &= 0.0500 \\
\text{power} &= 0.9000 \\
\delta &= 0.6842 \\
N_g &= 4 \\
m_1 &= 26.0700 \\
m_2 &= 25.5300 \\
m_3 &= 8.7500 \\
m_4 &= 13.5000 \\
Cm &= 14.6750 \\
c_0 &= 0.0000 \\
\text{Var}_Cm &= 53.8389 \\
\text{Var}_e &= 115.0000
\end{align*}
\]

Estimated sample sizes:

- \( N = 28 \)
- \( N \text{ per group} = 7 \)

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.

---

**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_{#1,#2})` number of subjects in group \((#_1, #_2)\)
- `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(grwt_{#1,#2})` group weight \((#_1, #_2)\)
- `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
Computing power

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

\[ y = Xb + \epsilon \]

where \( y \) is an \( n \times 1 \) vector of observations, \( X \) is an \( n \times p \) matrix of predictors, \( b \) is a \( p \times 1 \) vector of unknown and fixed coefficients, and \( \epsilon \) is an \( n \times 1 \) vector of error terms that are independent and identically distributed as \( N(0, \sigma^2_e) \).
For the one-way model, $p = J$ and the contents of $b = (b_1, \ldots, b_J)'$ are the $\mu_j, j = 1, \ldots, J$.

A general linear hypothesis in this framework is given by

$$Cb = c_0$$

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

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

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

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

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

where $SS_C = (C\hat{b} - c_0)'\{C(X'X)^{-1}C'\}^{-1}(C\hat{b} - c_0)$. Let $\alpha$ be the significance level, $\beta$ 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 $\lambda$ given by

$$\lambda = (Cb - c_0)'\{C(X'X)^{-1}C'\}^{-1}(Cb - c_0)/\sigma^2_e$$

$$= n(Cb - c_0)'\{\bar{X}'W\bar{X}\}^{-1}C'\{C(X'X)^{-1}C'\}^{-1}(Cb - c_0)/\sigma^2_e$$

$$= n\delta^2$$

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

For the one-way design, the dimension of $\bar{X}$ 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 $\bar{w}_j$, then define a cell sample size multiplier as $n_c = n/\sum \bar{w}_j$ so that $n_j = n_c\bar{w}_j$. The cell-means parameterization simplifies $\bar{X}$ to the identity matrix, $I_J$.

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

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

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

where $F_{\nu,n-p,\lambda}(\cdot)$ 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 $\bar{w}_jn_c$. 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

\[
1 - \beta = \begin{cases} 
1 - T_{n-1,\tilde{\lambda}}(t_{n-1,1-\alpha}) & \text{for an upper one-sided test} \\
T_{n-1,\tilde{\lambda}}(-t_{n-1,1-\alpha}) & \text{for a lower one-sided test} \\
F_{1,n-p,\lambda}(F_{1,n-p,1-\alpha}) & \text{for a two-sided test}
\end{cases}
\]  

(5)

where \( T_{\cdot,\tilde{\lambda}}(\cdot) \) is the cumulative of a noncentral Student’s \( t \) distribution with the noncentrality parameter \( \tilde{\lambda} \) given by

\[
\tilde{\lambda} = \sqrt{n}(Cb - c_0)' \sqrt{\left\{ C(\bar{X}'W\bar{X})^{-1}C' \right\}^{-1}} / \sigma_e^2 = \sqrt{n}\tilde{\delta}
\]

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

References


Also see

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

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

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

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

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

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

[PSS] Glossary

[R] anova — Analysis of variance and covariance

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

[R] oneway — One-way analysis of variance