Title

stata.com

power twoway — Power analysis for two-way analysis of variance

Syntax

Compute sample size

`power twoway` *meanspec* [, `power(`*numlist*`)` *options*]

Compute power

`power twoway` *meanspec* `, n(`*numlist*`)` [*options*]

Compute effect size and target effect variance

`power twoway, n(`*numlist*`) power(`*numlist*`) nrows(`*#*`) ncols(`*#*`)` [*options*]

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

$m_{1,1}\ m_{1,2}\ \left[\ldots\right]\ \backslash\ m_{2,1}\ m_{2,2}\ \left[\ldots\right]\ \left[\backslash\ \ldots \backslash\ m_{J,1}\ \ldots m_{J,K}\right]$

m_{jk}, where $j = 1, 2, \ldots, J$ and $k = 1, 2, \ldots, K$, is the alternative cell mean or the cell mean of the jth row and kth column under the alternative hypothesis.

matname is the name of a Stata matrix with J rows and K columns containing values of alternative cell means.

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
* npercell(*numlist*)	number of subjects per cell; implies balanced design
cellweights(*wgtspec*)	cell weights; default is one for each cell, meaning equal-cell sizes
nrows(#)	number of rows
ncols(#)	number of columns
factor(row | column | rowcol)	tested effect
* vareffect(*numlist*)	variance explained by the tested effect in factor()
* varrow(*numlist*)	variance explained by the row effect; synonym for factor(row) and vareffect(*numlist*)
* varcolumn(*numlist*)	variance explained by the column effect; synonym for factor(column) and vareffect(*numlist*)
* varrowcolumn(*numlist*)	variance explained by the row–column interaction effect; synonym for factor(rowcol) and vareffect(*numlist*)
* varerror(*numlist*)	error variance; default is varerror(1)
showmatrices	display cell means and sample sizes as matrices
showmeans	display cell means
showcellsizes	display cell sizes
parallel	treat number lists in starred options as parallel when multiple values per option are specified (do not enumerate all possible combinations of values)
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
$\#_{1,1}\ \ldots\ \#_{1,K}\ \backslash \ldots \backslash\ \#_{J,1}\ \ldots\ \#_{J,K}$	$J \times K$ cell weights; weights must be positive and must be integers unless option nfractional is specified
matname	$J \times K$ matrix containing cell weights

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_cell	number of subjects per cell	N/N_{rc}
N_avg	average number of subjects per cell	$N_{\rm avg}$
N$\#_1$_$\#_2$	number of subjects in cell ($\#_1$, $\#_2$)	$N_{\#_1,\#_2}$
delta	effect size	δ
N_rc	number of cells	N_{rc}
N_r	number of rows	N_r
N_c	number of columns	N_c
m$\#_1$_$\#_2$	cell mean ($\#_1$, $\#_2$)	$\mu_{\#_1,\#_2}$
Var_r	variance explained by the row effect	σ^2_r
Var_c	variance explained by the column effect	σ^2_c
Var_rc	variance explained by the row–column interaction	σ^2_{rc}
Var_e	error variance	σ^2_e
cwgt$\#_1$_$\#_2$	cell weight ($\#_1$, $\#_2$)	$w_{\#_1,\#_2}$
target	target parameter; synonym for target effect variance	
_all	display all supported columns	

Column beta is shown in the default table in place of column power if specified.
Column N_per_cell is available and is shown in the default table only for balanced designs.
Column N_avg is shown in the default table only for unbalanced designs.
Columns N$\#_1$_$\#_2$, N_rc, m$\#_1$_$\#_2$, and cwgt$\#_1$_$\#_2$ are not shown in the default table.

Menu

Statistics > Power and sample size

Description

power twoway computes sample size, power, or effect size for two-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 cells. You can choose between testing for main row or column effect or their interaction. Also see [PSS] **power** for a general introduction to the power command using hypothesis tests.

Options

Main

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

npercell(*numlist*) specifies the cell size. Only positive integers are allowed. This option implies a balanced design. npercell() cannot be specified with n() or cellweights().

cellweights(*wgtspec*) specifies $J \times K$ cell weights for an unbalanced design. The weights must be positive and must also be integers unless the nfractional option is specified. cellweights() cannot be specified with npercell().

nrows(*#*) specifies the number of rows or the number of levels of the row factor in a two-way ANOVA. At least two rows must be specified. This option is required if *meanspec* is not specified. This option is also required for effect-size determination unless cellweights() is specified.

ncols(*#*) specifies the number of columns or the number of levels of the column factor in a two-way ANOVA. At least two columns must be specified. This option is required if *meanspec* is not specified. This option is also required for effect-size determination unless cellweights() is specified.

factor(row | column | rowcol) specifies the effect of interest for which power and sample-size analysis is to be performed. In a two-way ANOVA, the tested effects include the main effects of a row factor (row effect), the main effects of a column factor (column effect), or the interaction effects between the row and column factors (row–column effect). The default is factor(row).

vareffect(*numlist*) specifies the variance explained by the tested effect specified in factor(). For example, if factor(row) is specified, vareffect() specifies the variance explained by the row factor. This option is required if the factor() option is specified and cell means are not specified. This option is not allowed with the effect-size determination. Only one of vareffect(), varrow(), varcolumn(), or varrowcolumn() may be specified.

varrow(*numlist*) specifies the variance explained by the row factor. This option is equivalent to specifying factor(row) and vareffect(*numlist*) and thus cannot be combined with factor(). This option is not allowed with the effect-size determination. Only one of vareffect(), varrow(), varcolumn(), or varrowcolumn() may be specified.

varcolumn(*numlist*) specifies the variance explained by the column factor. This option is equivalent to specifying factor(column) and vareffect(*numlist*) and thus cannot be combined with factor(). This option is not allowed with the effect-size determination. Only one of vareffect(), varrow(), varcolumn(), or varrowcolumn() may be specified.

varrowcolumn(*numlist*) specifies the variance explained by the interaction between row and column factors. This option is equivalent to specifying factor(rowcol) and vareffect(*numlist*) and thus cannot be combined with factor(). This option is not allowed with the effect-size determination. Only one of vareffect(), varrow(), varcolumn(), or varrowcolumn() may be specified.

`varerror(`*numlist*`)` specifies the error (within-cell) variance. The default is `varerror(1)`. If `varerror()` is not specified with effect-size determination, only effect size δ is reported; otherwise, the estimated variance of the tested effect is also reported.

`showmatrices` specifies that the matrices of cell means and cell sizes be displayed, when applicable. The cell means will be displayed only if specified. The cell sizes will be displayed only for an unbalanced design.

`showmeans` specifies that the cell means be reported. For a text or graphical output, this option is equivalent to `showmatrices` except only the cell-mean matrix will be reported. For a tabular output, the columns containing cell means will be included in the default table.

`showcellsizes` specifies that the cell sizes be reported. For a text or graphical output, this option is equivalent to `showmatrices` except only the cell-sizes matrix will be reported. For a tabular output, the columns containing cell sizes will be included in the default table.

`parallel`; see [PSS] **power**.

Table

`table`, `table(`*tablespec*`)`, `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 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] **power**.

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

`notitle`; see [PSS] **power**.

Remarks and examples

stata.com

Remarks are presented under the following headings:

This entry describes the `power twoway` command and the methodology for power and sample-size analysis for two-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

ANOVA has been one of the most widely used statistical tools in many scientific applications. Two-way ANOVA models allow analysts to study the effects of two factors simultaneously. The term "two way" refers to two factors each containing an arbitrary number of groups or levels.

For example, consider a type of drug with three levels of dosage in reducing blood pressure for males and females. In this case, three interesting hypotheses arise: an investigator may wish to test whether the average change in blood pressure is the same for both genders, whether the average change in blood pressure is the same across all levels of dosage regardless of gender, or whether there is any interaction between dosage levels and gender.

This entry describes power and sample-size analysis for the inference about main and interaction effects of two factors based on hypothesis testing. Let μ_{jk} be the cell mean of the jth row and the kth column in a two-way cell-means ANOVA model, $\mu_{j\cdot}$ be the marginal mean of the jth row, $\mu_{\cdot k}$ be the marginal mean of the kth column, and $\mu_{\cdot\cdot}$ be the grand mean. The jth-row-by-kth-column interaction effect is then $(ab)_{jk} = \mu_{jk} - \mu_{j\cdot} - \mu_{\cdot k} + \mu_{\cdot\cdot}$. We consider the null hypotheses 1) H_0: $\mu_{1\cdot} = \ldots = \mu_{J\cdot}$, for testing the main row effect; 2) H_0: $\mu_{\cdot 1} = \ldots = \mu_{\cdot K}$, for testing the main column effect; and 3) H_0: all $(ab)_{jk} = 0$, for testing the row-by-column interaction effect.

The test statistic for each of the three hypotheses is based on the ratio of the variance explained by the tested effect to the error variance. Under the null hypothesis, the test statistics used for items 1, 2, and 3 above have an F distribution. We will refer to the corresponding tests as F tests for row, column, and row-by-column effects. For power analysis, we consider the distribution of the test statistic under the alternative hypothesis. This distribution is a noncentral F distribution for all 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 each of the F tests is defined as the square root of the ratio of the variance explained by the tested effect to the error variance.

`power twoway` performs power and sample-size computation for a two-way fixed-effects ANOVA model based on an F test of the effect of interest.

Using power twoway

`power twoway` computes sample size, power, or effect size and target variance of the effect for a two-way fixed-effects ANOVA. 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 F test of the main row effects; `factor(row)` is assumed. You can instead request a test of the main column effects by specifying `factor(column)` or a test of the row-by-column interaction effects by specifying `factor(rowcol)`. The error variance for all tests 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 cell means, you can specify the number of rows in the `nrows()` option, the number of columns in the `ncols()` option, and the variance explained by the tested effect in the `vareffect()` option when computing sample size or power. See *Alternative ways of specifying effect*.

To compute effect size, the square root of the ratio of the variance explained by the tested factor 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 rows and columns in nrows() and ncols(), respectively. If the varerror() option is also specified, the target variance explained by the tested effect is reported in addition to the effect size.

By default, all computations assume a balanced or an equal-allocation design. You can use the cellweights() option to specify an unbalanced design for power, sample-size, or effect-size computations. For power and effect-size computations of a balanced design, you can also specify the npercell() option to specify a cell size instead of a total sample size in n().

In a two-way ANOVA, 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.

In a two-way ANOVA, all computations depend on the noncentrality parameter of the F distribution. The sample-size and effect-size computation requires a nonlinear search algorithm where the default initial value is obtained using a bisection search algorithm that brackets the solution.

power twoway provides several ways of specifying the study parameters that we discuss next.

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 twoway, you can do this in several ways. For example, consider a two-way model with $J = 2 \geq 2$ row-factor levels and $K = 3 \geq 2$ column-factor levels. You can specify either the individual alternative *meanspec*,

power twoway $m_{1,1}$ $m_{1,2}$ $m_{1,3}$ \ $m_{2,1}$ $m_{2,2}$ $m_{2,3}$ [, factor() ...]

or the variance of the tested effect and the number of rows J and columns K:

power twoway, factor() vareffect(#) nrows(2) ncols(3) [...]

You can also replace vareffect() and factor() in the above with the varrow(), varcolumn(), or varrowcolumn() option. And you can specify multiple values of the variances in these options.

As an alternative to directly specifying alternative cell means following the command name, you can define a Stata matrix containing these means and use it with power twoway. For example,

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

The matrix must have at least two rows and two columns.

power twoway *meanmat*, ...

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

Computing sample size

To compute sample size, you must specify the alternative cell means or the variance of the tested effect 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 two-way ANOVA—row effect

van Belle et al. (2004, 376) provide an example of an experimental study that investigates the effect of an automobile emission pollutant, nitrogen dioxide (NO_2). The experiment considers the effect of NO_2 exposure on protein leakage in the lungs of mice. In the experimental group, mice were exposed to 0.5 part per million (ppm) NO_2 for 10, 12, and 14 days. Measurements on the response variable, serum fluorescence, were taken on mice in the experimental (exposed) and control (unexposed) groups. The analysis of these data used a two-way ANOVA model with the exposure status as a row factor and the number of days of exposure to NO_2 as a column factor. The row factor has two levels, exposed or unexposed, and the column factor has three levels: 10, 12, and 14 days.

Suppose that investigators are planning to conduct another similar study. They would like to know how many subjects, mice, they need for the experiment. We will use the estimates of parameters from the above study to answer this question.

The estimated cell means from this study over the number of days are 134.4, 143, and 91.3 in the control group and 106.4, 173.2, and 145.5 in the experimental group. From the ANOVA table on page 379 (van Belle et al. 2004), the estimated residual variance is 1417.35. For convenience, we round these numbers down to the nearest integers in our computations.

We begin by computing sample size for testing the main treatment (exposure) effects using `power twoway`'s defaults for other aspects of the study: a balanced design, a 5% significance level, and 80% power. (Option `factor(row)` is assumed by default.)

```
. power twoway 134 143 91 \ 106 173 145, varerror(1417)
Performing iteration ...
Estimated sample size for two-way ANOVA
F test for row effect
Ho: delta = 0  versus  Ha: delta != 0
Study parameters:
        alpha =    0.0500
        power =    0.8000
        delta =    0.2479
          N_r =         2
          N_c =         3
        means =  <matrix>
        Var_r =   87.1111
        Var_e = 1417.0000
Estimated sample sizes:
            N =       132
   N per cell =        22
```

Assuming a balanced design, we need a total of 132 mice with 22 mice per cell to detect the effect of exposure to NO_2 on the protein leakage of mice.

Like all other `power` methods, `power twoway` reports study parameters first and the estimated parameters next. The reported study parameters include the specified and implied parameters such as significance level, power, number of rows, number of columns, and so on. `power twoway` does not display the specified cell means by default but indicates in the output that the means are specified. You can specify the `showmeans` option to display cell means as a matrix.

In addition to the specified and implied study parameters, `power twoway` reports the value of the effect size, `delta` $= \sqrt{87.1111/1417} = 0.2479$, computed as a square root of the ratio between the variance of the row effect `Var_r` and the error variance `Var_e`. As for the one-way ANOVA models, the effect size δ 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: Sample size for a two-way ANOVA—column effect

Continuing example 1, we can compute the required sample size for the main column effects by specifying the `factor(column)` option:

```
. power twoway 134 143 91 \ 106 173 145, varerror(1417) factor(column)
Performing iteration ...
Estimated sample size for two-way ANOVA
F test for column effect
Ho: delta = 0  versus  Ha: delta != 0
Study parameters:
          alpha =    0.0500
          power =    0.8000
          delta =    0.4889
            N_r =         2
            N_c =         3
          means =   <matrix>
          Var_c =  338.6667
          Var_e = 1417.0000
Estimated sample sizes:
              N =        48
     N per cell =         8
```

Assuming a balanced design, we need a total of 48 mice with 8 mice per cell to detect the effect of the length of exposure to NO_2 on the protein leakage of mice.

Similarly to the row effect, the effect size for the column effect, `delta` $= \sqrt{338.6667/1417} = 0.4889$, is computed as a square root of the ratio between the variance of the column effect `Var_c` and the error variance `Var_e`. The interpretation remains the same but with respect to the main column effects. According to Cohen's scale, the effect size corresponding to the test of the main column effects is large, so we need fewer subjects to detect the column effect than we need to detect the previous row effect.

◁

▷ Example 3: Sample size for a two-way ANOVA—row-by-column effect

Continuing example 2, we can also compute the required sample size for the row-by-column effects interaction by specifying the `factor(rowcol)` option:

```
. power twoway 134 143 91 \ 106 173 145, varerror(1417) factor(rowcol)
Performing iteration ...
Estimated sample size for two-way ANOVA
F test for row-by-column effect
Ho: delta = 0  versus  Ha: delta != 0
Study parameters:
        alpha =    0.0500
        power =    0.8000
        delta =    0.4572
          N_r =         2
          N_c =         3
        means =  <matrix>
       Var_rc =  296.2222
        Var_e = 1417.0000
Estimated sample sizes:
            N =        54
   N per cell =         9
```

For a balanced design, we need a total of 54 mice with 9 mice per cell to detect the joint effects of exposure and the length of exposure to NO_2 on the protein leakage of mice.

Similarly to the row-by-column effects, the effect size for the row-by-column effect, `delta` $= \sqrt{296.2222/1417} = 0.4572$, is computed as a square root of the ratio between the variance of the row-by-column effect `Var_rc` and the error variance `Var_e`. The interpretation is again the same but with respect to the interaction of row-by-column effects. According to Cohen's scale, the effect size corresponding to the test of the row-by-column effects is also large, so we need fewer subjects to detect this effect than we need to detect the row effect. The effect size is similar to the column effect size, so the required numbers of subjects are comparable for the two tests. As a final sample size, we would choose the largest of the three sizes to ensure that we have enough subjects to detect any of the considered effects.

◁

▷ Example 4: Alternative ways of specifying effect

Instead of specifying the alternative cell means as in previous examples, we can specify the variance explained by the corresponding tested effect and the numbers of rows and columns.

For instance, from example 2, the variance explained by the column effect was computed to be 338.6667. We specify this value in `vareffect()` as well as the number of rows in `nrows()` and the number of columns in `ncols()`:

```
. power twoway, varerror(1417) factor(column) vareffect(338.6667) nrows(2) ncols(3)
Performing iteration ...
Estimated sample size for two-way ANOVA
F test for column effect
Ho: delta = 0  versus  Ha: delta != 0
Study parameters:
        alpha =    0.0500
        power =    0.8000
        delta =    0.4889
          N_r =         2
          N_c =         3
        Var_c =  338.6667
        Var_e = 1417.0000
Estimated sample sizes:
            N =        48
   N per cell =         8
```

We obtain the exact same results as in example 2.

A shorthand for the specification of `factor(column)` and `vareffect()` is the `varcolumn()` option. You can verify that the specification

```
. power twoway, varerror(1417) varcolumn(338.6667) nrows(2) ncols(3)
  (output omitted )
```

produces results identical to the results above.

You can also use similar alternative specifications for the tests of row and row-by-column effects with intuitive modifications to the syntax.

`power twoway` also provides another alternative specification of the cell means. Instead of specifying alternative cell means directly following the command line, as in example 2, we can define a matrix, say, `means`, containing these means and use it with `power twoway`:

```
. matrix define means = (134, 143, 91 \ 106, 173, 145)
. power twoway means, varerror(1417) factor(column)
  (output omitted )
```

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

◁

▷ Example 5: Unbalanced design

Continuing example 1, let's compute the required sample size for an unbalanced design. For instance, consider a design in which the control group (the first row) contains twice as many subjects as the experimental group (the second row) for each level of the other factor. We use the `cellweights()` option to specify weights for each cell.

```
. power twoway 134 143 91 \ 106 173 145, varerror(1417) cellweights(2 2 2 \ 1 1 1)
> showcellsizes
Performing iteration ...
Estimated sample size for two-way ANOVA
F test for row effect
Ho: delta = 0  versus  Ha: delta != 0
Study parameters:
        alpha =    0.0500
        power =    0.8000
        delta =    0.2338
          N_r =         2
          N_c =         3
        means =  <matrix>
        Var_r =   77.4321
        Var_e = 1417.0000
Estimated sample sizes:
            N =       153
    Average N =   25.5000
  Cell sample sizes

                 columns
                        1           2           3

  rows
                1      34          34          34
                2      17          17          17
```

The required total sample size for this unbalanced design is 153 with the average number of subjects in a cell of 25.5. We specified the `showcellsizes` option to display the number of subjects for each cell along with the total and average sample sizes that are displayed by default for an unbalanced design.

You can alternatively specify cell weights as a matrix.

◁

Computing power

To compute power, you must specify the total sample size in the `n()` option and the desired effect size, expressed using alternative cell means, for example. See *Alternative ways of specifying effect*.

▷ Example 6: Power for a two-way ANOVA

Continuing example 1, suppose that we anticipate a sample of 90 mice. To compute the corresponding power, we specify the sample size of 90 in `n()`.

```
. power twoway 134 143 91 \ 106 173 145, varerror(1417) n(90)
Estimated power for two-way ANOVA
F test for row effect
Ho: delta = 0  versus  Ha: delta != 0
Study parameters:
        alpha =    0.0500
            N =        90
   N per cell =        15
        delta =    0.2479
          N_r =         2
          N_c =         3
        means =   <matrix>
        Var_r =   87.1111
        Var_e = 1417.0000
Estimated power:
        power =    0.6426
```

For this smaller sample size, the power for detecting the effect size of 0.25 is only 64%.

◁

▷ Example 7: Multiple values of study parameters

Continuing example 6, we may want to check powers for several sample sizes. We simply list multiple sample-size values in `n()`:

```
. power twoway 134 143 91 \ 106 173 145, varerror(1417) n(90 114 126)
Estimated power for two-way ANOVA
F test for row effect
Ho: delta = 0  versus  Ha: delta != 0
   means = <matrix>
```

alpha	power	N	N_per_cell	delta	N_r	N_c	Var_r	Var_e
.05	.6426	90	15	.2479	2	3	87.11	1417
.05	.7466	114	19	.2479	2	3	87.11	1417
.05	.7884	126	21	.2479	2	3	87.11	1417

The larger the sample size, the larger the power.

We can even compute results for multiple sample sizes and, for example, multiple values of error variances.

```
. power twoway 134 143 91 \ 106 173 145, varerror(1000 1800) n(90 114 126)
> table(, sep(3))
Estimated power for two-way ANOVA
F test for row effect
Ho: delta = 0  versus  Ha: delta != 0
   means = <matrix>
```

alpha	power	N	N_per_cell	delta	N_r	N_c	Var_r	Var_e
.05	.7904	90	15	.2951	2	3	87.11	1000
.05	.8776	114	19	.2951	2	3	87.11	1000
.05	.9076	126	21	.2951	2	3	87.11	1000
.05	.5411	90	15	.22	2	3	87.11	1800
.05	.6436	114	19	.22	2	3	87.11	1800
.05	.6878	126	21	.22	2	3	87.11	1800

We specified `table()`'s suboption `separator()`, abbreviated to `sep()`, to improve readability of the table.

We can also compute results for combinations 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 target variance explained by the tested effect

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 numbers of rows and columns must be specified.

The effect size is defined as a square root of the ratio of the variance explained by the tested effect to the error variance. The effect size is computed by default. If the error variance is specified in the `varerror()` option, the target variance of the effect is also reported.

▷ Example 8: Effect size for a two-way ANOVA—row effect

Continuing example 6, we now want to compute the effect size that can be detected for a sample of 90 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 rows in `nrows()` and the number of columns in `ncols()`:

```
. power twoway, varerror(1417) n(90) power(0.8) nrows(2) ncols(3)
Performing iteration ...
Estimated row variance for two-way ANOVA
F test for row effect
Ho: delta = 0  versus  Ha: delta != 0
Study parameters:
        alpha =    0.0500
        power =    0.8000
            N =        90
   N per cell =        15
          N_r =         2
          N_c =         3
        Var_e = 1417.0000
Estimated effect size and row variance:
        delta =    0.2987
        Var_r =  126.4634
```

With a smaller sample size, given the same power, we can only detect a larger effect size of 0.2987, compared with the effect size of 0.2479 from example 1. The corresponding estimate of the variance explained by the row effect is 126.46, given the error variance of 1417.

◁

Testing hypotheses about means from multiple populations

▷ Example 9: Two-way ANOVA

After the initial power and sample-size planning, we can use Stata's `anova` command to perform inference for two-way ANOVA based on the collected sample. We show a quick example of how to do this here; see [R] **anova** for more examples and details.

We use data on systolic blood pressure. Consider a sample of 58 patients, each suffering from 1 of 3 different diseases, who were randomly assigned to 1 of 4 different drug treatments and whose change in systolic blood pressure was recorded. To test for the effects of the drug and the disease and their interaction, we type the following:

```
. use http://www.stata-press.com/data/r13/systolic
(Systolic Blood Pressure Data)
. anova systolic drug disease drug#disease
                           Number of obs =      58     R-squared     =  0.4560
                           Root MSE      = 10.5096     Adj R-squared =  0.3259

                  Source   Partial SS    df       MS           F     Prob > F

                   Model   4259.33851    11   387.212591       3.51     0.0013

                    drug   2997.47186     3   999.157287       9.05     0.0001
                 disease   415.873046     2   207.936523       1.88     0.1637
            drug#disease   707.266259     6    117.87771       1.07     0.3958

                Residual   5080.81667    46   110.452536

                   Total   9340.15517    57   163.862371
```

We find that only the main effect of the drug is significant.

Suppose that we would like to conduct a similar study. We use the estimates from the study above to compute the required sample size for our new study. We are particularly interested in testing the interaction between the drug and disease, so we would like to compute the sample size for this test.

First, we estimate the cell means of systolic blood pressure for different treatment and disease levels:

```
. table drug disease, contents(mean systolic) format(%9.0f)

──────────┬─────────────────
          │ Patient's Disease
Drug Used │     1      2      3
──────────┼─────────────────
        1 │    29     28     20
        2 │    28     34     18
        3 │    16      4      8
        4 │    14     13     14
──────────┴─────────────────
```

From the twoway output, the estimate of the error variance is roughly 110. Second, we specify the means and the error variance with power twoway and compute the required sample size for a balanced design assuming 5% significance level and 80% power for the test of interaction effects.

```
. power twoway 29 28 20 \ 28 34 18 \ 16 4 8 \ 14 13 14, varerror(110) f(rowcol)
Performing iteration ...
Estimated sample size for two-way ANOVA
F test for row-by-column effect
Ho: delta = 0  versus  Ha: delta != 0
Study parameters:
        alpha =    0.0500
        power =    0.8000
        delta =    0.3465
          N_r =         4
          N_c =         3
        means =   <matrix>
       Var_rc =   13.2083
        Var_e =  110.0000
Estimated sample sizes:
            N =       132
   N per cell =        11
```

We need a total of 132 subjects with 11 subjects per cell to detect the drug-by-disease effect size of 0.3465 for this design.

To determine the final sample size, you may want to repeat the same computations for the tests of the main effects of drug and the main effects of disease and select the sample size based on the three tests.

◁

Stored results

`power twoway` 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 cell ($\#_1$, $\#_2$)
`r(N_per_cell)`	number of subjects per cell
`r(N_rc)`	number of cells
`r(nfractional)`	`1` if `nfractional` is specified; `0` otherwise
`r(balanced)`	`1` for a balanced design; `0` otherwise
`r(cwgt`$\#_1$`_`$\#_2$`)`	cell weight ($\#_1$, $\#_2$)
`r(onesided)`	`1` for a one-sided test of a mean contrast; `0` otherwise
`r(N_r)`	number of rows
`r(N_c)`	number of columns
`r(m`$\#_1$`_`$\#_2$`)`	cell mean ($\#_1$, $\#_2$)
`r(Var_r)`	row variance
`r(Var_c)`	column variance
`r(Var_rc)`	row-by-column variance
`r(Var_e)`	error 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 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)`	`twoway`
`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
`r(Nij)`	cell-sizes matrix
`r(means)`	cell-means matrix
`r(cwgt)`	cell-weights matrix

Methods and formulas

Consider factor A with J levels and factor B with K levels. Let μ_{jk} be the mean of cell (j,k) in a table formed by the levels of factors A and B. For example, let $J = 3$ and $K = 3$; then the following cell-means table summarizes the experiment.

	Factor B			
Factor A	$k=1$	$k=2$	$k=3$	Total
$j=1$	μ_{11}	μ_{12}	μ_{13}	$\mu_{1\cdot}$
$j=2$	μ_{21}	μ_{22}	μ_{23}	$\mu_{2\cdot}$
$j=3$	μ_{31}	μ_{32}	μ_{33}	$\mu_{3\cdot}$
Total	$\mu_{\cdot 1}$	$\mu_{\cdot 2}$	$\mu_{\cdot 3}$	$\mu_{\cdot\cdot}$

Methods and formulas are presented under the following headings:

Main effects
Interaction effects
Hypothesis testing

Main effects

Main effects measure the deviation of the factor-level means from the overall or grand mean. The larger the main effect, the more likely you can detect the effect. From the above table, the main effect of factor A at the jth level is $a_j = \mu_{j\cdot} - \mu_{\cdot\cdot}$. Similarly, the main effect of factor B at the kth level is $b_k = \mu_{\cdot k} - \mu_{\cdot\cdot}$. The overall mean can be expressed as

$$\mu_{\cdot\cdot} = \frac{\sum_{j=1}^{J}\sum_{k=1}^{K}\mu_{jk}}{JK} = \frac{\sum_{j=1}^{J}\mu_{j\cdot}}{J} = \frac{\sum_{k=1}^{K}\mu_{\cdot k}}{K}$$

This implies that

$$\sum_{j=1}^{J} a_j = 0 \qquad \text{and} \qquad \sum_{k=1}^{K} b_k = 0$$

Interaction effects

Unlike main effects that measure the effect of individual factors on the dependent variable, interaction effects measure the effect of the two factors jointly on the dependent variable. For example, the interaction effect of factor A at the jth level and factor B at the kth level is $(ab)_{jk} = \mu_{jk} - \mu_{j\cdot} - \mu_{\cdot k} + \mu_{\cdot\cdot}$.

The sum of interaction effects is zero:

$$\sum_{j=1}^{J}(ab)_{jk} = 0 \text{ at each level of } k = 1, \dots, K$$

$$\sum_{k=1}^{K}(ab)_{jk} = 0 \text{ at each level of } j = 1, \dots, J$$

This implies

$$\sum_{j=1}^{J}\sum_{k=1}^{K}(ab)_{jk} = 0$$

Hypothesis testing

Let n denote the total sample size and y_{ijk} denote the response of ith individual at the jth level of factor A and kth level of factor B for $i = 1, \ldots, n_{jk}$ such that $n = \sum_{j,k} n_{jk}$. ANOVA models assume that responses y_{ijk} within each cell are independent and identically distributed random normal with constant variance σ_e^2. Using the definitions of a_j, b_k, and $(ab)_{jk}$ from the previous sections, our linear model is expressed as

$$\begin{aligned} y_{ijk} &= \mu_{..} + a_j + b_k + (ab)_{jk} + e_{ijk} \\ &= \mu_{jk} + e_{ijk} \end{aligned}$$

where $\mu_{..}$ is the overall mean, and e_{ijk}'s are the independent error terms that have the standard normal distribution. The first equation corresponds to the formulation of an ANOVA model using effects, and the second formulation corresponds to the cell-means formulation.

The variance explained by the row effects is $\sigma_a^2 = \sum_j a_j^2/J$, by the column effects is $\sigma_b^2 = \sum_k b_k^2/K$, and by the row-by-column effects is $\sigma_{(ab)}^2 = \sum_{j,k} (ab)_{jk}^2/JK$.

The following sets of hypotheses are of interest in a two-way ANOVA:

$$\begin{aligned} H_0\text{: all } a_j &= 0 \quad \text{versus} \quad H_a\text{: at least one } a_j \neq 0 & (1) \\ H_0\text{: all } b_k &= 0 \quad \text{versus} \quad H_a\text{: at least one } b_k \neq 0 & (2) \\ H_0\text{: all } (ab)_{jk} &= 0 \quad \text{versus} \quad H_a\text{: at least one } (ab)_{jk} \neq 0 & (3) \end{aligned}$$

Hypotheses (1) and (2) test the main effects of factors A and B, respectively, and hypothesis (3) tests the interaction effects between A and B.

To test the above hypotheses, we can use the general linear model framework discussed in *Methods and formulas* of [PSS] **power oneway**. We recapitulate it here with application to the two-way model.

A general test statistic for testing hypotheses like (1), (2), and (3) is given by

$$F_C = \frac{SS_C}{\nu \widehat{\sigma}_e^2} \tag{4}$$

where $SS_C = (\mathbf{C}\widehat{\mathbf{b}})'\{\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}'\}^{-1}(\mathbf{C}\widehat{\mathbf{b}})$. The $n \times JK$ matrix $\mathbf{X}$ specifies the coding for the two-way design. Matrix $\mathbf{C}$ is $\nu \times JK$ and contains the contrasts for the means that are used to test each of the three hypotheses. For the two-way design, v is $J-1$, $K-1$, or $(J-1)(K-1)$ for hypotheses (1), (2), and (3), respectively.

Let α be the significance level and $F_{\nu,n-JK,1-\alpha}$ be the $(1-\alpha)$th quantile of an F distribution with ν numerator and $n - JK$ denominator degrees of freedom. We reject the null hypothesis if we observe a statistic $F_C > F_{\nu,n-JK,1-\alpha}$.

Under the alternative hypothesis, the test statistic (4) is distributed as a noncentral F distribution with ν numerator and $n - JK$ denominator degrees of freedom and a noncentrality parameter λ given by

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

where the matrix $\ddot{\mathbf{X}}$ contains the unique rows of $\mathbf{X}$ such that $\boldsymbol{\mu} = \ddot{\mathbf{X}}\mathbf{b}$, $\mathbf{W} = \text{diag}(w_1, \ldots, w_{JK})$, and δ is the effect size. For a two-way design, the dimension of $\ddot{\mathbf{X}}$ is $JK \times JK$, and the weights are $w_i = n_{jk}/n$, $i = (k-1)J + j$. The cell-means parameterization simplifies $\ddot{\mathbf{X}}$ to the identity matrix, $\mathbf{I}_{JK}$.

The rows of $\ddot{\mathbf{X}}$ and weights w_i are associated with the column-major, `vec()`, order of the two-way table with factor A indexed on the rows and factor B indexed on the columns. (See the 3×3 table in the beginning of this section, and scan each column $k = 1, 2, 3$.) The weight w_i can be reexpressed as a cell weight $\widetilde{w}_i$, which is independent of the sample size n; see *Methods and formulas* of [PSS] **power oneway** for details.

When the `cellweights()` option is specified, a constant cell-size multiplier n_c is computed and rounded to an integer unless the `nfractional` option is specified. The cell sizes are then computed as $\widetilde{w}_j n_c$. The actual sample size, `N_a`, is the sum of the cell sizes.

References

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

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

[PSS] **power oneway** — Power analysis for one-way analysis of variance

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