power pairedproportions — Power analysis for a two-sample paired-proportions test

Syntax

Compute sample size

   Specify discordant proportions

   power pairedproportions p_{12} p_{21} [, power(*numlist*) *discordopts*]

   Specify marginal proportions

   power pairedproportions p_{1+} p_{+1} , corr(*numlist*) [power(*numlist*) *margopts*]

Compute power

   Specify discordant proportions

   power pairedproportions p_{12} p_{21} , n(*numlist*) [*discordopts*]

   Specify marginal proportions

   power pairedproportions p_{1+} p_{+1} , corr(*numlist*) n(*numlist*) [*margopts*]

Compute effect size and target discordant proportions

   power pairedproportions, n(*numlist*) power(*numlist*) prdiscordant(*numlist*)

   [*discordopts*]

where p_{12} is the probability of a success at occasion 1 and a failure at occasion 2, and p_{21} is the probability of a failure at occasion 1 and a success at occasion 2. Each represents the probability of a discordant pair. p_{1+} is the marginal probability of a success for occasion 1, and p_{+1} is the marginal probability of a success for occasion 2. Each may be specified either as one number or as a list of values in parentheses; see [U] **11.1.8 numlist**.

discordopts	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*)*	sample size; required to compute power or effect size
nfractional	allow fractional sample size
*prdiscordant(*numlist*)*	sum of the discordant proportions, $p_{12} + p_{21}$
*sum(*numlist*)*	synonym for prdiscordant()
*diff(*numlist*)*	difference between the discordant proportions, $p_{21} - p_{12}$
*ratio(*numlist*)*	ratio of the discordant proportions, p_{21}/p_{12}
effect(*effect*)	specify the type of effect to display; default is effect(diff)
direction(upper|lower)	direction of the effect for effect-size determination; default is direction(upper), which means that the postulated value of the parameter is larger than the hypothesized value
onesided	one-sided test; default is two sided
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 difference between discordant proportions
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.

margopts	Description
[Main]	
* <u>alpha</u>(*numlist*)	significance level; default is `alpha(0.05)`
* <u>pow</u>er(*numlist*)	power; default is `power(0.8)`
* <u>beta</u>(*numlist*)	probability of type II error; default is `beta(0.2)`
* n(*numlist*)	sample size; required to compute power or effect size
<u>nfrac</u>tional	allow fractional sample size
* <u>corr</u>(*numlist*)	correlation between the paired observations
* <u>diff</u>(*numlist*)	difference between the marginal proportions, $p_{+1} - p_{1+}$
* <u>ratio</u>(*numlist*)	ratio of the marginal proportions, p_{+1}/p_{1+}
* <u>rr</u>isk(*numlist*)	relative risk, p_{+1}/p_{1+}
* <u>or</u>atio(*numlist*)	odds ratio, $\{p_{+1}(1-p_{+1})\}/\{p_{1+}(1-p_{+1})\}$
<u>eff</u>ect(*effect*)	specify the type of effect to display; default is `effect(diff)`
<u>dir</u>ection(<u>upper</u>|<u>lower</u>)	direction of the effect for effect-size determination; default is `direction(upper)`, which means that the postulated value of the parameter is larger than the hypothesized value
<u>onesided</u>	one-sided test; default is two sided
<u>parallel</u>	treat number lists in starred options as parallel when multiple values per option are specified (do not enumerate all possible combinations of values)
[Table]	
[<u>no</u>]<u>table</u>[(*tablespec*)]	suppress table or display results as a table; see [PSS] **power, table**
<u>saving</u>(*filename* [, replace])	save the table data to *filename*; use replace to overwrite existing *filename*
[Graph]	
<u>graph</u>[(*graphopts*)]	graph results; see [PSS] **power, graph**
<u>notitle</u>	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.

effect	Description
diff	difference between the discordant proportions, $p_{21} - p_{12}$, or marginal proportions, $p_{+1} - p_{1+}$; the default
ratio	ratio of the discordant proportions, p_{21}/p_{12}, or of the marginal proportions, p_{+1}/p_{1+}
<u>rr</u>isk	relative risk, p_{+1}/p_{1+}; may only be specified with marginal proportions
<u>or</u>atio	odds ratio, $\{p_{+1}(1-p_{+1})\}/\{p_{1+}(1-p_{+1})\}$; may only be specified with marginal proportions

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	number of subjects	N
delta	effect size	δ
p12	success–failure proportion	p_{12}
p21	failure–success proportion	p_{21}
pmarg1	success proportion in occasion 1	p_{1+}
pmarg2	success proportion in occasion 2	p_{+1}
corr	correlation between paired observations	ρ
prdiscordant	proportion of discordant pairs	$p_{12} + p_{21}$
sum	sum of discordant proportions	$p_{12} + p_{21}$
diff	difference between discordant proportions	$p_{21} - p_{12}$
	difference between marginal proportions	$p_{+1} - p_{1+}$
ratio	ratio of discordant proportions	p_{21}/p_{12}
	ratio of marginal proportions	p_{+1}/p_{1+}
rrisk	relative risk for marginal proportions	p_{+1}/p_{1+}
oratio	odds ratio for marginal proportions	θ
target	target parameter; synonym for p12	
_all	display all supported columns	

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

Columns p12 and p21 are shown in the default table if discordant proportions are specified.

Columns pmarg1, pmarg2, and corr are shown in the default table if marginal proportions are specified.

Columns pmarg1, pmarg2, corr, rrisk, and oratio are available only if marginal proportions are specified.

Columns diff, ratio, prdiscordant, sum, rrisk, and oratio are shown in the default table if specified.

Menu

Statistics > Power and sample size

Description

power pairedproportions computes sample size, power, or target discordant proportions for a two-sample paired-proportions test, also known as McNemar's test. By default, it computes sample size for given power and values of the discordant or marginal proportions. Alternatively, it can compute power for given sample size and values of the discordant or marginal proportions, or it can compute the target discordant proportions for given sample size and power. 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**. The nfractional option is allowed only for sample-size determination.

corr(*numlist*) specifies the correlation between paired observations. This option is required if marginal proportions are specified.

diff(*numlist*) specifies the difference between the discordant proportions, $p_{21} - p_{12}$, or the marginal proportions, $p_{+1} - p_{1+}$. See *Alternative ways of specifying effect* for details about the specification of this option.

ratio(*numlist*) specifies the ratio of the discordant proportions, p_{21}/p_{12}, or the marginal proportions, p_{+1}/p_{1+}. See *Alternative ways of specifying effect* for details about the specification of this option.

prdiscordant(*numlist*) specifies the proportion of discordant pairs or the sum of the discordant proportions, $p_{12} + p_{21}$. See *Alternative ways of specifying effect* for details about the specification of this option.

sum(*numlist*) is a synonym for prdiscordant(). See *Alternative ways of specifying effect* for details about the specification of this option.

rrisk(*numlist*) specifies the relative risk of the marginal proportions, p_{+1}/p_{1+}. See *Alternative ways of specifying effect* for details about the specification of this option.

oratio(*numlist*) specifies the odds ratio of the marginal proportions, $\{p_{+1}(1-p_{+1})\}/\{p_{1+}(1-p_{+1})\}$. See *Alternative ways of specifying effect* for details about the specification of this option.

effect(*effect*) specifies the type of the effect size to be reported in the output as delta. *effect* is one of diff or ratio for discordant proportions and one of diff, ratio, rrisk, or oratio for marginal proportions. By default, the effect size delta is the difference between proportions. If diff(), ratio(), rrisk(), or oratio() is specified, the effect size delta will contain the effect corresponding to the specified option. For example, if ratio() is specified, delta will contain the ratio of the proportions. See *Alternative ways of specifying effect* for details about the specification of this option.

direction(), onesided, 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 for the estimated parameter. The estimated parameter is sample size for sample-size determination or the difference between the discordant proportions for the effect-size determination.

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

The following option is available with `power pairedproportions` 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 pairedproportions` command and the methodology for power and sample-size analysis for a two-sample paired-proportions test. See [PSS] **intro** for a general introduction to power and sample-size analysis and [PSS] **power** for a general introduction to the `power` command using hypothesis tests.

Introduction

The analysis of paired proportions is used to compare two dependent binomial populations. Dependent binomial data arise from matched case–control studies, where the cases are matched to the controls on the basis of similar demographic characteristics, or from longitudinal studies, where the same cases serve as their own controls over time or for different treatments. In all cases, each observation represents a pair of correlated binary outcomes.

There are many examples of studies where a researcher would like to compare two dependent proportions. For example, a state highway department might be concerned that tollbooth workers may experience hearing loss because of chronic exposure to traffic noise. It wants to test whether the proportion of workers with moderate to severe hearing loss is the same between a sample of workers exposed to traffic noise and a sample of workers sheltered by the quieter interior of the booth. Or a pediatrician might conduct a study to compare the proportions of males and females with a particular food allergy in a study of male/female fraternal twins.

This entry describes power and sample-size analysis for correlated binary outcomes in a two-way contingency table. Consider a 2×2 table from a study where the outcome of interest is a pair of results from "occasion 1" and "occasion 2", each either a "success" or a "failure".

	Occasion 2		
Occasion 1	Success	Failure	Total
Success	n_{11}	n_{12}	n_{1+}
Failure	n_{21}	n_{22}	n_{2+}
Total	n_{+1}	n_{+2}	n

n is the total number of pairs; n_{11} is the number of pairs for which the response is a success for both occasions; n_{12} is the number of *success–failure pairs* for which the response is a success on occasion 1 and a failure on occasion 2; n_{21} is the number of *failure–success pairs* for which the response is a failure on occasion 1 and a success on occasion 2; and n_{22} is the number of pairs for which the response is a failure for both occasions. The success–failure and failure–success pairs form *discordant pairs*, and the remaining pairs form *concordant pairs*.

The above table can also be expressed in terms of the proportions.

| | Occasion 2 | | |
Occasion 1	Success	Failure	Total
Success	p_{11}	p_{12}	p_{1+}
Failure	p_{21}	p_{22}	$1 - p_{1+}$
Total	p_{+1}	$1 - p_{+1}$	1

p_{1+} is the success probability for occasion 1, and p_{+1} is the success probability for occasion 2. The marginal probabilities, p_{1+} and p_{+1}, are used to compare the outcomes between occasion 1 and occasion 2.

The null hypothesis for the test of equality of marginal proportions, also known as the test of *marginal homogeneity*, is H_0: $p_{+1} = p_{1+}$. The null hypothesis can be formulated in terms of the discordant probabilities, the failure–success probability, p_{21}, and the success–failure probability, p_{12}, using the relationships $p_{+1} = p_{11} + p_{21}$ and $p_{1+} = p_{11} + p_{12}$. The considered null hypothesis is then H_0: $p_{21} = p_{12}$ versus the two-sided alternative hypothesis H_a: $p_{21} \neq p_{12}$, the upper one-sided alternative H_a: $p_{21} > p_{12}$, or the lower one-sided alternative H_a: $p_{21} < p_{12}$. For a 2×2 table, the test of marginal homogeneity is also called a *test of symmetry*.

A large-sample McNemar's test is commonly used for testing the above hypotheses. Under the null hypothesis, the test statistic is distributed as a χ_1^2 distribution with 1 degree of freedom.

`power pairedproportions` provides power and sample-size analysis for McNemar's test of two correlated proportions.

Using power pairedproportions

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

For sample-size and power determinations, `power pairedproportions` provides a number of ways of specifying the magnitude of an effect desired to be detected by the test. Below we describe the use of the command, assuming that the desired effect is expressed by the values of the two discordant proportions; see *Alternative ways of specifying effect* for other specifications.

To compute sample size, you must specify the discordant proportions, p_{12} and p_{21}, and, optionally, the power of the test in option `power()`. The default power is set to 0.8.

To compute power, you must specify the sample size in option `n()` and the discordant proportions, p_{12} and p_{21}.

The effect-size determination is available only for discordant proportions. To compute effect size and target discordant proportions, you must specify the sample size in option `n()`, the power in option `power()`, the sum of the discordant proportions in option `prdiscordant()`, and, optionally, the direction of the effect. The direction is upper by default, `direction(upper)`, which means that the failure–success proportion, p_{21}, is assumed to be larger than the specified success–failure proportion, p_{12}. You can change the direction to lower, which means that p_{21} is assumed to be smaller than p_{12}, by specifying the `direction(lower)` option.

There are multiple definitions of effect size for a two-sample paired-proportions test. The `effect()` option specifies what definition `power pairedproportions` should use when reporting the effect size, which is labeled as `delta` in the output of the `power` command.

When you specify the discordant proportions, the available definitions are the difference $p_{21} - p_{12}$ between the discordant proportions, effect(diff), or the ratio p_{21}/p_{12} of the discordant proportions, effect(ratio).

When you specify the marginal proportions, the available definitions are the difference $p_{+1} - p_{1+}$ between the marginal proportions, effect(diff); the relative risk or ratio p_{+1}/p_{1+} of the marginal proportions, effect(rrisk) or effect(ratio); or the odds ratio $\{p_{+1}(1-p_{1+})\}/\{p_{1+}(1-p_{+1})\}$ of the marginal proportions, effect(oratio).

When effect() is specified, the effect size delta in the output of the power command contains the estimate of the corresponding effect and is labeled accordingly. By default, delta corresponds to the difference between proportions. If any one of the options diff(), ratio(), rrisk(), or oratio() is specified and effect() is not specified, delta will contain the effect size corresponding to the specified option.

Some of power pairedproportions's computations require iteration. For example, a sample size for a two-sided test is obtained by iteratively solving a nonlinear power equation. The default initial value for the sample size for the iteration procedure is obtained using a closed-form one-sided formula. If you desire, you may change it by specifying 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 also specify the magnitude of the effect that is desired to be detected by the test. You can do this by specifying either the discordant proportions, p_{12} and p_{21},

> power pairedproportions p_{12} p_{21} , ...

or the marginal proportions, p_{1+} and p_{+1}:

> power pairedproportions p_{1+} p_{+1} , corr(*numlist*) ...

When you specify marginal proportions, you must also specify the correlation between paired observations in option corr().

Below we describe other alternative specifications separately for discordant proportions and marginal proportions.

Effect specifications for discordant proportions

Instead of specifying p_{21}, you may specify the discordant proportion p_{12} as the argument to the command and the sum of the discordant proportions, $p_{12} + p_{21}$, in option prdiscordant() or option sum(),

> power pairedproportions p_{12} , prdiscordant(*numlist*) ...

> power pairedproportions p_{12} , sum(*numlist*) ...

the difference between the discordant proportions, $p_{21} - p_{12}$, in option diff(),

> power pairedproportions p_{12} , diff(*numlist*) ...

or the ratio of the discordant proportions, p_{21}/p_{12}, in option ratio():

> power pairedproportions p_{12} , ratio(*numlist*) ...

You may omit both command arguments p_{12} and p_{21} altogether and specify options prdiscordant() or sum(), diff(), and ratio() in pairs.

For example, you can specify the sum $p_{12} + p_{21}$ and difference $p_{21} - p_{12}$ of the discordant proportions:

      `power pairedproportions , prdiscordant(`*numlist*`) diff(`*numlist*`)` ...

Or you can specify the sum $p_{12} + p_{21}$ and ratio p_{21}/p_{12} of the discordant proportions:

      `power pairedproportions , sum(`*numlist*`) ratio(`*numlist*`)` ...

Or you can specify the difference $p_{21} - p_{12}$ and ratio p_{21}/p_{12} of the discordant proportions:

      `power pairedproportions , diff(`*numlist*`) ratio(`*numlist*`)` ...

When discordant proportions are specified, the effect size may be expressed as the difference between discordant proportions, $p_{21} - p_{12}$, or the ratio of discordant proportions, p_{21}/p_{12}. You may choose what effect to compute by specifying the `effect()` option.

By default, effect size is the difference between the discordant proportions. For example, for the specification below, the effect size δ is the difference between the discordant proportions.

      `power pairedproportions` p_{12} p_{21} `,` ...

The above specification is equivalent to

      `power pairedproportions` p_{12} p_{21} `, effect(diff)` ...

Alternatively, you may request the effect size to be the ratio instead of the difference.

      `power pairedproportions` p_{12} p_{21} `, effect(ratio)` ...

Likewise, if you specify the `ratio()` option, the effect size is the ratio of the proportions.

Effect specifications for marginal proportions

Instead of specifying p_{+1}, you may specify the marginal proportion, p_{1+}, as the argument to the command and the difference between the marginal proportions, $p_{+1} - p_{1+}$, in option `diff()`,

      `power pairedproportions` p_{1+} `, corr(`*numlist*`) diff(`*numlist*`)` ...

the ratio of the marginal proportions or relative risk, p_{+1}/p_{1+}, in option `ratio()` or option `rrisk()`,

      `power pairedproportions` p_{1+} `, corr(`*numlist*`) ratio(`*numlist*`)` ...

      `power pairedproportions` p_{1+} `, corr(`*numlist*`) rrisk(`*numlist*`)` ...

or the odds ratio, $\{p_{+1}(1 - p_{1+})\}/\{p_{1+}(1 - p_{+1})\}$, in option `oratio()`:

      `power pairedproportions` p_{1+} `, corr(`*numlist*`) oratio(`*numlist*`)` ...

Alternatively, you may omit both command arguments p_{1+} and p_{+1} and specify one of the combinations of `diff()` and `rrisk()` or `ratio()`, or `oratio()` and `rrisk()` or `ratio()`. You may not combine `diff()` and `oratio()`, because marginal proportions cannot be identified uniquely from this combination.

For example, you can specify the difference $p_{+1} - p_{1+}$ and ratio p_{+1}/p_{1+} of the marginal proportions,

      `power pairedproportions , corr(`*numlist*`) diff(`*numlist*`) ratio(`*numlist*`)` ...

or the odds ratio, $\{p_{+1}(1 - p_{1+})\}/\{p_{1+}(1 - p_{+1})\}$, and the relative risk, p_{+1}/p_{1+}:

      `power pairedproportions , corr(`*numlist*`) oratio(`*numlist*`) rrisk(`*numlist*`)` ...

The effect size for marginal proportions may be defined as one of the difference between marginal proportions, $p_{+1} - p_{1+}$, the odds ratio, $\{p_{+1}(1 - p_{1+})\}/\{p_{1+}(1 - p_{+1})\}$, or the relative risk or, equivalently, the ratio p_{+1}/p_{1+}. The `effect()` option for marginal proportions may contain one of `diff`, `oratio`, `rrisk`, or `ratio`.

By default, effect size is defined as the difference between the marginal proportions. For example, the following specification,

> `power pairedproportions` p_{1+} p_{+1} `, corr(`*numlist*`)` ...

is equivalent to

> `power pairedproportions` p_{1+} p_{+1} `, corr(`*numlist*`) effect(diff)` ...

You may request other measures of effect size for marginal proportions such as the risk ratio,

> `power pairedproportions` p_{1+} p_{+1} `, corr(`*numlist*`) effect(rrisk)` ...

or odds ratio:

> `power pairedproportions` p_{1+} p_{+1} `, corr(`*numlist*`) effect(oratio)` ...

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

Computing sample size

To compute sample size, you must specify the discordant proportions, p_{12} and p_{21}, and, optionally, the power of the test in option `power()`. The default power is set to 0.8. Instead of the discordant proportions, you can specify an effect of interest as shown in *Alternative ways of specifying effect*.

▷ Example 1: Sample size for a two-sample paired-proportions test

Consider a study from Agresti (2013, 413) where the same group of subjects was asked who they voted for in the 2004 and 2008 presidential elections. In 2008, we witnessed a shift from the Republican President George W. Bush, who was finishing his second term, to Democratic President Barack Obama, who was beginning his first term. Suppose that we would like to conduct another survey for the 2012 and 2016 elections to see whether a similar shift—this time from a Democrat to a Republican—would occur in 2016, when President Obama finishes his second term. We are interested in testing the hypothesis whether the proportions of votes for a Democratic president in 2012 will be the same as in 2016.

Consider the following 2×2 table:

| | 2016 Election | | |
2012 Election	Democratic	Republican	Total
Democratic	p_{11}	p_{12}	p_{1+}
Republican	p_{21}	p_{22}	$1 - p_{1+}$
Total	p_{+1}	$1 - p_{+1}$	1

The test of marginal homogeneity that the proportion of Democratic votes in 2012 will be the same in 2016 is given by the null hypothesis $H_0: p_{1+} = p_{+1}$. This is equivalent to testing whether the proportion p_{12} of voters who changed parties from Democratic to Republican is the same as the proportion p_{21} of voters who changed parties from Republican to Democratic between 2012 and 2016. The corresponding null hypothesis that tests these discordant proportions is $H_0: p_{12} = p_{21}$.

Suppose that the previous survey reported that the proportion of respondents who voted for a Democratic president in 2012 is $p_{1+} = 0.53$. Using this empirical evidence, a political expert believes that the odds of a candidate being elected as the president when his or her party has already served two consecutive terms as the president are low. According to the expert's opinion, the odds of the population voting Democrat in 2016 to the odds of the population voting Democrat in 2012 are 2:3; that is, the corresponding odds ratio is $\theta = 2/3 = 0.667$. Using the relationship between the odds ratio and marginal probabilities, we compute the marginal probability p_{+1} to be $= [0.667 \times \{0.53/(1 - 0.53)\}]/(1 + [(0.667 \times \{0.53/(1 - 0.53)\}]) = 0.4293$.

It is generally believed that voting behavior is positively correlated, which means that a person voting for one party in an election year is very likely to vote for the same party in the next election year. Suppose the expert posits a correlation of 0.8.

We wish to compute the sample size required for our survey to detect the difference between the considered marginal proportions. To compute the minimum sample size, we specify the marginal proportions after the command name and the correlation in option `corr()`:

```
. power pairedproportions 0.53 0.4293, corr(0.8)

Performing iteration ...

Estimated sample size for a two-sample paired-proportions test
Large-sample McNemar's test
Ho: p+1 = p1+  versus  Ha: p+1 != p1+

Study parameters:
        alpha =    0.0500
        power =    0.8000
        delta =   -0.1007  (difference)
          p1+ =    0.5300
          p+1 =    0.4293
         corr =    0.8000

Estimated sample size:
            N =        82
```

We find that 82 subjects are required in our survey for a 5%-level two-sided McNemar's test to detect a change in the proportion voting Democrat from 0.53 in 2012 to 0.4293 in 2016, which corresponds to the difference of $\delta = 0.4293 - 0.53 = -0.1007$, with 80% power.

◁

▷ Example 2: Reporting odds ratio

By default, as in example 1, the effect size δ is the difference between the marginal proportions. Alternatively, we can request that the effect size be reported as the odds ratio by specifying option `effect(oratio)`.

```
. power pairedproportions 0.53 0.4293, corr(0.8) effect(oratio)

Performing iteration ...

Estimated sample size for a two-sample paired-proportions test
Large-sample McNemar's test
Ho: p+1 = p1+  versus  Ha: p+1 != p1+

Study parameters:
        alpha =    0.0500
        power =    0.8000
        delta =    0.6671  (odds ratio)
          p1+ =    0.5300
          p+1 =    0.4293
         corr =    0.8000

Estimated sample size:
            N =        82
```

The effect size `delta` now contains the odds ratio estimated from the specified marginal proportions.
Also see *Alternative ways of specifying effect* for other available measures of effect.

◁

▷ Example 3: Specifying odds ratio

In example 1, we computed the second marginal proportion, p_{+1}, using the postulated values of the
first marginal proportion, p_{1+}, and the odds ratio, θ, and we specified the two marginal proportions
with power pairedproportions. We can instead specify the first marginal proportion and the odds
ratio directly:

```
. power pairedproportions 0.53, corr(0.8) oratio(0.667)

Performing iteration ...

Estimated sample size for a two-sample paired-proportions test
Large-sample McNemar's test
Ho: p+1 = p1+  versus  Ha: p+1 != p1+

Study parameters:
        alpha =    0.0500
        power =    0.8000
        delta =    0.6670  (odds ratio)
          p1+ =    0.5300
          p+1 =    0.4293
         corr =    0.8000
   odds ratio =    0.6670

Estimated sample size:
            N =        82
```

When the `oratio()` option is specified, the reported effect size `delta` corresponds to the odds ratio,
and the value of the odds ratio specified in `oratio()` is also reported in the output.

Also see *Alternative ways of specifying effect* for other ways of specifying an effect.

◁

▷ Example 4: Specifying discordant proportions

Instead of marginal proportions, as in example 1, we can specify discordant proportions. We compute discordant proportions using (1) and the estimates of marginal proportions and correlation in this study: $p_{12} = 0.53(1 - 0.4293) - 0.8\sqrt{0.53 \times (1 - 0.53) \times 0.4293 \times (1 - 0.4293)} = 0.105$ and $p_{21} = 0.105 + 0.4293 - 0.53 = 0.004$.

```
. power pairedproportions 0.105 0.004
Performing iteration ...
Estimated sample size for a two-sample paired-proportions test
Large-sample McNemar's test
Ho: p21 = p12  versus  Ha: p21 != p12
Study parameters:
        alpha =    0.0500
        power =    0.8000
        delta =   -0.1010  (difference)
          p12 =    0.1050
          p21 =    0.0040
Estimated sample size:
            N =        82
```

We obtain the same sample size of 82 as in example 1. The reported effect size `delta` corresponds to the difference of -0.1010 between the discordant proportions.

Also see *Alternative ways of specifying effect* for other ways of specifying the effect desired to be detected by the test.

◁

Computing power

To compute power, you must specify the sample size in option `n()` and the discordant proportions, p_{12} and p_{21}. Instead of the discordant proportions, you can specify an effect of interest as shown in *Alternative ways of specifying effect*.

▷ Example 5: Power of a two-sample paired-proportions test

Continuing with example 4, we will suppose that we anticipate to obtain a sample of 100 subjects and want to compute the power corresponding to this sample size.

In addition to the discordant proportions, we specify the sample size of 100 in option `n()`:

```
. power pairedproportions 0.105 0.004, n(100)
Estimated power for a two-sample paired-proportions test
Large-sample McNemar's test
Ho: p21 = p12  versus  Ha: p21 != p12
Study parameters:
        alpha =    0.0500
            N =       100
        delta =   -0.1010  (difference)
          p12 =    0.1050
          p21 =    0.0040
Estimated power:
        power =    0.8759
```

As expected, with a larger sample size, this example achieves a larger power, about 88%, compared with example 4.

◁

▷ Example 6: Alternative specification of an effect for power determination

Continuing with example 5, we can specify the first discordant proportion, p_{12}, and one of the options prdiscordant(), diff(), or ratio() instead of specifying both discordant proportions.

For example, let's specify the sum of the discordant proportions instead of the discordant proportion p_{21}:

```
. power pairedproportions 0.105, n(100) prdiscordant(0.109)
Estimated power for a two-sample paired-proportions test
Large-sample McNemar's test
Ho: p21 = p12   versus  Ha: p21 != p12

Study parameters:

        alpha =     0.0500
            N =        100
        delta =    -0.1010   (difference)
          p12 =     0.1050
          p21 =     0.0040
    p12 + p21 =     0.1090

Estimated power:

        power =     0.8759
```

We obtain results identical to those in example 5.

◁

▷ Example 7: Power determination with marginal proportions

We can compute the power corresponding to the sample of 100 subjects for the specification using marginal proportions from example 1 by additionally specifying option n(100):

```
. power pairedproportions 0.53 0.4293, corr(0.8) n(100)
Estimated power for a two-sample paired-proportions test
Large-sample McNemar's test
Ho: p+1 = p1+   versus  Ha: p+1 != p1+

Study parameters:

        alpha =     0.0500
            N =        100
        delta =    -0.1007   (difference)
          p1+ =     0.5300
          p+1 =     0.4293
         corr =     0.8000

Estimated power:

        power =     0.8739
```

As expected, the estimated power of 0.8739 is very close to the estimated power of 0.8759 in example 5. If we had used input values for discordant proportions with more precision in example 5, we would have obtained nearly identical results.

◁

▷ Example 8: Multiple values of study parameters

Continuing with example 7, we would like to assess the effect of varying correlation on the power of our study.

We specify a range of correlations between 0.2 and 0.8 with a step size of 0.1 in option `corr()`:

```
. power pairedproportions 0.53 0.4293, corr(0.2(0.1)0.8) n(100)
Estimated power for a two-sample paired-proportions test
Large-sample McNemar's test
Ho: p+1 = p1+  versus  Ha: p+1 != p1+
```

alpha	power	N	delta	pmarg1	pmarg2	corr
.05	.3509	100	−.1007	.53	.4293	.2
.05	.3913	100	−.1007	.53	.4293	.3
.05	.4429	100	−.1007	.53	.4293	.4
.05	.5105	100	−.1007	.53	.4293	.5
.05	.6008	100	−.1007	.53	.4293	.6
.05	.7223	100	−.1007	.53	.4293	.7
.05	.8739	100	−.1007	.53	.4293	.8

For a given sample size, the power increases as the correlation increases, which means that for a given power, the required sample size decreases as the correlation increases. This demonstrates that a paired design can improve the precision of the statistical inference compared with an independent design, which corresponds to a correlation of zero.

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 discordant proportions

The effect-size determination is available only for discordant proportions. As we describe in detail in *Alternative ways of specifying effect*, there are multiple definitions of the effect size for a paired-proportions test. The default is the difference between the failure–success proportion, p_{21}, and the success–failure proportion, p_{12}.

Sometimes, we may be interested in determining the smallest effect and the corresponding discordant proportions that yield a statistically significant result for prespecified sample size and power. In this case, power, sample size, and the sum of the discordant proportions must be specified. In addition, you must also decide on the direction of the effect: upper, meaning $p_{21} > p_{12}$, or lower, meaning $p_{21} < p_{12}$. The direction may be specified in the `direction()` option; `direction(upper)` is the default.

▷ Example 9: Compute effect size and target proportions

Suppose that we want to compute the corresponding discordant proportions for given sample size and power. To compute the discordant proportions, we must specify the proportion of discordant pairs (the sum of the discordant proportions) in addition to sample size and power.

Continuing with example 4, we will compute the corresponding effect size and target proportions for the sample size of 82 and the power of 0.8 using 0.109 for the proportion of discordant pairs. We also specify the `direction(lower)` option because $p_{21} < p_{12}$ in our example.

```
. power pairedproportions, prdiscordant(.109) n(82) power(0.8) direction(lower)
Performing iteration ...
Estimated discordant proportions for a two-sample paired-proportions test
Large-sample McNemar's test
Ho: p21 = p12  versus  Ha: p21 != p12; p21 < p12
Study parameters:
        alpha =    0.0500
        power =    0.8000
            N =        82
    p12 + p21 =    0.1090
Estimated effect size and discordant proportions:
        delta =   -0.1007  (difference)
          p12 =    0.1048
          p21 =    0.0042
```

The estimated discordant proportions of 0.1048 and 0.0042 are very close to the respective original estimates of the discordant proportions of 0.105 and 0.004 from example 4.

◁

Testing a hypothesis about two correlated proportions

Suppose we collected data from two paired binomial samples and wish to test whether the two proportions of an outcome of interest are the same. We wish to use McNemar's test to test this hypothesis. We can use the `mcc` command to perform McNemar's test; see [ST] **epitab** for details.

▷ Example 10: Testing for paired proportions

We use data provided in table 11.1 of Agresti (2013, 414) that present the results of a General Social Survey, which asked males who they voted for in the 2004 and 2008 presidential elections.

	2008 Election		
2004 Election	Democrat	Republican	Total
Democrat	175	16	191
Republican	54	188	42
Total	229	204	433

We wish to test whether there was a change in the voting behavior of males in 2008 compared with 2004 using McNemar's test. We use `mcci`, the immediate form of `mcc`, to perform this test.

```
. mcci 175 16 54 188
                    │    Controls
Cases               │  Exposed   Unexposed  │       Total
────────────────────┼──────────────────────┼─────────────
            Exposed │      175          16  │         191
          Unexposed │       54         188  │         242
────────────────────┼──────────────────────┼─────────────
              Total │      229         204  │         433
McNemar's chi2(1) =      20.63    Prob > chi2 = 0.0000
Exact McNemar significance probability       = 0.0000
Proportion with factor
          Cases        .4411085
          Controls     .5288684      [95% Conf. Interval]
                     ─────────────  ─────────────────────
          difference -.0877598       -.1270274  -.0484923
          ratio        .8340611       .7711619   .9020907
          rel. diff. -.1862745       -.2738252  -.0987238

          odds ratio  .2962963        .1582882   .5254949   (exact)
```

McNemar's test statistic is 20.63 with the corresponding two-sided p-value less than 10^{-4}, which provides strong evidence of a shift in the Democratic direction among male voters in 2008.

We use the estimates of this study to perform a sample-size analysis we would have conducted before a new study. The discordant proportions are $p_{12} = 16/433 = 0.037$ and $p_{21} = 54/433 = 0.125$.

```
. power pairedproportions 0.037 0.125
Performing iteration ...
Estimated sample size for a two-sample paired-proportions test
Large-sample McNemar's test
Ho: p21 = p12  versus  Ha: p21 != p12
Study parameters:
        alpha =    0.0500
        power =    0.8000
        delta =    0.0880  (difference)
          p12 =    0.0370
          p21 =    0.1250
Estimated sample size:
            N =       162
```

We find that we need a sample of 162 respondents to detect a difference of 0.0880 between discordant proportions of 0.037 and 0.125 with 80% power using a 5%-level two-sided test.

◁

Stored results

power pairedproportions 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)	sample size
r(nfractional)	1 if nfractional is specified; 0 otherwise
r(onesided)	1 for a one-sided test; 0 otherwise
r(p12)	success–failure proportion (first discordant proportion)
r(p21)	failure–success proportion (second discordant proportion)
r(pmarg1)	success proportion for occasion 1 (first marginal proportion)
r(pmarg2)	success proportion for occasion 2 (second marginal proportion)
r(corr)	correlation between paired observations
r(diff)	difference between proportions
r(ratio)	ratio of proportions
r(prdiscordant)	proportion of discordant pairs
r(sum)	sum of discordant proportions
r(rrisk)	relative risk
r(oratio)	odds ratio
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 the difference between discordant proportions
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(effect)	diff, ratio, oratio, or rrisk
r(type)	test
r(method)	pairedproportions
r(direction)	upper or lower
r(columns)	displayed table columns
r(labels)	table column labels
r(widths)	table column widths
r(formats)	table column formats

Matrix
r(pss_table)	table of results

Methods and formulas

Consider a 2×2 contingency table formed with n pairs of observations. The first subscript $i = 1, 2$ denotes the success of failure in occasion 1, and the second subscript $j = 1, 2$ denotes the success of failure in occasion 2.

	Occasion 2		
Occasion 1	Success	Failure	Total
Success	p_{11}	p_{12}	p_{1+}
Failure	p_{21}	p_{22}	$1 - p_{1+}$
Total	p_{+1}	$1 - p_{+1}$	1

Each element in the above table denotes the probability of observing a specific pair. For example, p_{11} is the probability of jointly observing a success on occasion 1 and occasion 2, and p_{12} is the probability of observing a success on occasion 1 and a failure on occasion 2. p_{1+} is the marginal probability of a success on occasion 1, and p_{+1} is the marginal probability of a success on occasion 2. The off-diagonal proportions p_{12} and p_{21} are referred to as *discordant proportions*. The relationship between the discordant proportions and the marginal proportions is given by

$$p_{12} = p_{1+}(1 - p_{+1}) - \rho\sqrt{p_{1+}(1 - p_{1+})\ p_{+1}(1 - p_{+1})}$$

$$p_{21} = p_{12} + p_{+1} - p_{1+} \tag{1}$$

where ρ is the correlation between the paired observations.

A two-sample paired proportions test involves testing the null hypothesis H_0: $p_{+1} = p_{1+}$ versus the two-sided alternative hypothesis H_a: $p_{+1} \neq p_{1+}$, the upper one-sided alternative H_a: $p_{+1} > p_{1+}$, or the lower one-sided alternative H_a: $p_{+1} < p_{1+}$. Using the relationship $p_{1+} = p_{11} + p_{12}$ and $p_{+1} = p_{11} + p_{21}$, test hypotheses may be stated in terms of the discordant proportions, for example, H_0: $p_{21} = p_{12}$ versus H_a: $p_{21} \neq p_{12}$.

McNemar's test statistic is:

$$\chi^2 = (n_{12} - n_{21})^2 / (n_{12} + n_{21})$$

where n_{ij} is the number of successes ($i = 1$) or failures ($i = 2$) on occasion 1 and the number of successes ($j = 1$) or failures ($j = 2$) on occasion 2; see Lachin (2011, chap. 5) for details. This test statistic has an approximately χ^2 distribution with 1 degree of freedom under the null hypothesis. The square root of the χ^2 test statistic is approximately normal with zero mean and variance of one.

Let α be the significance level, β be the probability of a type II error, and $z_{1-\alpha/k}$ and z_{β} be the $(1 - \alpha/k)$th and the βth quantiles of the standard normal distribution.

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

$$
\pi = \begin{cases}
\Phi\left\{\dfrac{p_{\text{diff}}\sqrt{n} - z_{1-\alpha}\sqrt{p_{\text{disc}}}}{\sqrt{p_{\text{disc}} - p_{\text{diff}}^2}}\right\} & \text{for an upper one-sided test} \\[3ex]
\Phi\left\{\dfrac{-p_{\text{diff}}\sqrt{n} - z_{1-\alpha}\sqrt{p_{\text{disc}}}}{\sqrt{p_{\text{disc}} - p_{\text{diff}}^2}}\right\} & \text{for a lower one-sided test} \\[3ex]
\Phi\left\{\dfrac{p_{\text{diff}}\sqrt{n} - z_{1-\alpha/2}\sqrt{p_{\text{disc}}}}{\sqrt{p_{\text{disc}} - p_{\text{diff}}^2}}\right\} + \Phi\left\{\dfrac{-p_{\text{diff}}\sqrt{n} - z_{1-\alpha/2}\sqrt{p_{\text{disc}}}}{\sqrt{p_{\text{disc}} - p_{\text{diff}}^2}}\right\} & \text{for a two-sided test}
\end{cases}
\tag{2}
$$

where $\Phi(\cdot)$ is the cdf of a standard normal distribution, $p_{\text{diff}} = p_{21} - p_{12}$, and $p_{\text{disc}} = p_{12} + p_{21}$.

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

$$n = \left(\frac{z_\alpha\sqrt{p_{\text{disc}}} + z_{1-\beta}\sqrt{p_{\text{disc}} - p_{\text{diff}}^2}}{p_{\text{diff}}}\right)^2 \tag{3}$$

See Connor (1987) for details.

For a two-sided test, sample size is computed iteratively from the two-sided power equation in (2). The default initial value is obtained from the corresponding one-sided formula (3) with the significance level $\alpha/2$.

The effect size $\delta = p_{21} - p_{12}$ is computed iteratively from the corresponding power equation in (2). The default initial value is $p_{\text{disc}}/2$.

References

Agresti, A. 2013. *Categorical Data Analysis*. 3rd ed. Hoboken, NJ: Wiley.

Connor, R. J. 1987. Sample size for testing differences in proportions for the paired-sample design. *Biometrics* 43: 207–211.

Lachin, J. M. 2011. *Biostatistical Methods: The Assessment of Relative Risks*. 2nd ed. Hoboken, NJ: Wiley.

Also see

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

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

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

[PSS] **Glossary**

[ST] **epitab** — Tables for epidemiologists