Title

> **power twoproportions, cluster** — Power analysis for a two-sample proportions test, CRD

Description	Quick start	Menu	Syntax
Options	Remarks and examples	Stored results	Methods and formulas
References	Also see		

Description

power twoproportions, cluster computes group-specific numbers of clusters, group-specific cluster sizes, power, or the experimental-group proportion for a two-sample proportions test in a cluster randomized design (CRD). It computes group-specific numbers of clusters given cluster sizes, power, and the values of the control-group and experimental-group proportions. It also computes group-specific cluster sizes given numbers of clusters, power, and the values of the control-group and experimental-group proportions. Alternatively, it computes power given numbers of clusters, cluster sizes, and the values of the control-group and experimental-group proportions, or it computes the experimental-group proportion given numbers of clusters, cluster sizes, power, and the control-group proportion. See [PSS-2] **power twoproportions** for a general discussion of power and sample-size analysis for a two-sample proportions test. Also see [PSS-2] **power** for a general introduction to the power command using hypothesis tests.

Quick start

Number of clusters for a test of $H_0: \pi_1 = \pi_2$ versus $H_a: \pi_1 \neq \pi_2$ given alternative control- and experimental-group proportions $p_1 = 0.8$ and $p_2 = 0.6$ with common cluster size of 5 using default intraclass correlation of 0.5, power of 0.8, and significance level $\alpha = 0.05$

    power twoproportions 0.8 0.6, m1(5) m2(5)

Same as above, but assume the intraclass correlation is 0.4

    power twoproportions 0.8 0.6, m1(5) m2(5) rho(0.4)

Same as above, but with an average cluster size of 5 and a coefficient of variation of 0.1

    power twoproportions 0.8 0.6, m1(5) m2(5) rho(0.4) cvcluster(0.1)

Group-specific numbers of clusters where the ratio of experimental to control group clusters is 0.5

    power twoproportions 0.8 0.6, m1(5) m2(5) rho(0.4) kratio(0.5)

Cluster sizes for a test of $H_0: \pi_1 = \pi_2$ versus $H_a: \pi_1 \neq \pi_2$ given alternative control- and experimental-group proportions $p_1 = 0.4$ and $p_2 = 0.2$ for 60 clusters in the control group and 40 clusters in the experimental group using default intraclass correlation of 0.5, power of 0.8, and significance level $\alpha = 0.05$

    power twoproportions 0.4 0.2, k1(60) k2(40)

Same as above, but compute experimental-group cluster size given the control-group cluster size of 10

    power twoproportions 0.4 0.2, k1(60) k2(40) m1(10) compute(M2)

Power given control-group proportion $p_1 = 0.1$, experimental-group proportion $p_2 = 0.2$, and 20 clusters with cluster sizes of 5 in the control and experimental groups

    power twoproportions 0.1 0.2, k1(20) k2(20) m1(5) m2(5)

Power for multiple numbers of clusters in the experimental group
```
power twoproportions 0.1 0.2, k1(20) k2(20(5)50) m1(5) m2(5)
```

Same as above, but display results in a graph of power versus the number of clusters in the experimental group
```
power twoproportions 0.1 0.2, k1(20) k2(20(5)50) m1(5) m2(5) graph
```

Effect size and experimental-group proportion with power of 0.8
```
power twoproportions 0.1, k1(20) k2(20) m1(5) m2(5) power(0.8)
```

Menu

Statistics > Power, precision, and sample size

Syntax

Compute numbers of clusters

> power <u>two</u>proportions p_1 p_2 , {*mspec* | *nspec* cluster} [*options*]

Compute cluster sizes

> power <u>two</u>proportions p_1 p_2 , *kspec* [*options*]

Compute power

> power <u>two</u>proportions p_1 p_2 , *kspec* { *mspec* | *nspec* } [*options*]

Compute effect size and experimental-group proportion

> power <u>two</u>proportions p_1 , *kspec* { *mspec* | *nspec* } <u>p</u>ower(*numlist*) [*options*]

where p_1 is the proportion in the control (reference) group and p_2 is the proportion in the experimental (comparison) group. p_1 and p_2 may each be specified either as one number or as a list of values in parentheses (see [U] **11.1.8 numlist**).

kspec is one of

> k1() k2()
> k1() [<u>krat</u>io()]
> k2() [<u>krat</u>io()]

mspec is one of

> m1() m2()
> m1() [<u>mrat</u>io()]
> m2() [<u>mrat</u>io()]

nspec is one of

```
n1()  n2()
n1()  [ nratio() ]
n2()  [ nratio() ]
```

options	Description
Main	
cluster	perform computations for a CRD; implied by k1(), k2(), m1(), or m2()
* 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)
* k1(*numlist*)	number of clusters in the control group
* k2(*numlist*)	number of clusters in the experimental group
* kratio(*numlist*)	cluster ratio, K2/K1; default is kratio(1)
* m1(*numlist*)	cluster size of the control group
* m2(*numlist*)	cluster size of the experimental group
* mratio(*numlist*)	cluster-size ratio, M2/M1; default is mratio(1)
* n1(*numlist*)	sample size of the control group
* n2(*numlist*)	sample size of the experimental group
* nratio(*numlist*)	sample-size ratio, N2/N1; default is nratio(1)
compute(K1 | K2 | M1 | M2)	solve for the number of clusters or cluster size in one group given the other group
nfractional	allow fractional numbers of clusters, cluster sizes, and sample sizes
* diff(*numlist*)	difference between the experimental-group and control-group proportions, $p_2 - p_1$; specify instead of the experimental-group proportion p_2
* ratio(*numlist*)	ratio of the experimental-group proportion to the control-group proportion, p_2/p_1; specify instead of the experimental-group proportion p_2
* rdiff(*numlist*)	risk difference, $p_2 - p_1$; synonym for diff()
* rrisk(*numlist*)	relative risk, p_2/p_1; synonym for ratio()
* oratio(*numlist*)	odds ratio, $\{p_2(1 - p_1)\}/\{p_1(1 - p_2)\}$
effect(*effect*)	specify the type of effect to display; default is effect(diff)
* rho(*numlist*)	intraclass correlation; default is rho(0.5)
* cvcluster(*numlist*)	coefficient of variation for cluster sizes
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 or in command arguments as parallel when multiple values per option or argument are specified (do not enumerate all possible combinations of values)

Table

$\lceil$ <u>no</u> $\rceil$ <u>table</u> $\lceil$ (*tablespec*) $\rceil$	suppress table or display results as a table; see [PSS-2] **power, table**
<u>saving</u>(*filename* $\lceil$, replace $\rceil$)	save the table data to *filename*; use replace to overwrite existing *filename*

Graph

<u>graph</u> $\lceil$ (*graphopts*) $\rceil$	graph results; see [PSS-2] **power, graph**

Iteration

init(*#*)	initial value for numbers of clusters, cluster sizes, or experimental-group proportion
<u>iterate</u>(*#*)	maximum number of iterations; default is iterate(500)
tolerance(*#*)	parameter tolerance; default is tolerance(1e-12)
<u>ftol</u>erance(*#*)	function tolerance; default is ftolerance(1e-12)
$\lceil$ no $\rceil$ log	suppress or display iteration log
$\lceil$ no $\rceil$ dots	suppress or display iterations as dots
<u>notitle</u>	suppress the title

*Specifying a list of values in at least two starred options, or at least two command arguments, or at least one starred option and one argument results in computations for all possible combinations of the values; see [U] **11.1.8 numlist**. Also see the parallel option.

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

notitle does not appear in the dialog box.

effect	Description
diff	difference between proportions, $p_2 - p_1$; the default
ratio	ratio of proportions, p_2/p_1
<u>rdiff</u>	risk difference, $p_2 - p_1$
<u>rrisk</u>	relative risk, p_2/p_1
<u>oratio</u>	odds ratio, $\{p_2(1 - p_1)\}/\{p_1(1 - p_2)\}$

where *tablespec* is

   column $\lceil$:*label* $\rceil$ $\lceil$ *column* $\lceil$:*label* $\rceil$ $\lceil$ ... $\rceil$ $\rceil$ $\lceil$, *tableopts* $\rceil$

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
K1	number of clusters in the control group	K_1
K2	number of clusters in the experimental group	K_2
kratio	ratio of numbers of clusters, experimental to control	K_2/K_1
M1	cluster size of the control group	M_1
M2	cluster size of the experimental group	M_2
mratio	ratio of cluster sizes, experimental to control	M_2/M_1
N	total number of observations	N
N1	number of observations in the control group	N_1
N2	number of observations in the experimental group	N_2
nratio	ratio of sample sizes, experimental to control	N_2/N_1
delta	effect size	δ
p1	control-group proportion	p_1
p2	experimental-group proportion	p_2
diff	difference between the experimental-group proportion and the control-group proportion	$p_2 - p_1$
ratio	ratio of the experimental-group proportion to the control-group proportion	p_2/p_1
rdiff	risk difference	$p_2 - p_1$
rrisk	relative risk	p_2/p_1
oratio	odds ratio	θ
rho	intraclass correlation	ρ
CV_cluster	coefficient of variation for cluster sizes	CV_{cl}
target	target parameter; synonym for p2	
_all	display all supported columns	

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

Column N is shown in the table if specified.

Columns N1 and N2 are shown in the default table if n1() or n2() is specified.

Columns nratio, diff, ratio, rdiff, rrisk, oratio, and CV_cluster are shown in the default table if specified.

Options

> Main

cluster specifies that computations should be performed for a CRD. This option is implied when the k1(), k2(), m1(), or m2() option is specified. cluster is required to compute the numbers of clusters when *nspec* is used to specify sample sizes instead of *mspec* for cluster sizes.

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

k1(*numlist*) specifies the number of clusters in the control group.

k2(*numlist*) specifies the number of clusters in the experimental group.

kratio(*numlist*) specifies the ratio of the numbers of clusters of the experimental group relative to the control group, K2/K1. The default is kratio(1), meaning equal numbers of clusters in the two groups.

m1(*numlist*) specifies the cluster size of the control group. m1() may contain noninteger values.

m2(*numlist*) specifies the cluster size of the experimental group. m2() may contain noninteger values.

mratio(*numlist*) specifies the ratio of cluster sizes of the experimental group relative to the control group, M2/M1. The default is mratio(1), meaning equal cluster sizes in the two groups.

n1(), n2(), nratio(); see [PSS-2] **power**.

compute(K1 | K2 | M1 | M2) solve for the number of clusters or cluster size of one group given the other group.

nfractional; see [PSS-2] **power**. The nfractional option displays fractional (without rounding) values of the numbers of clusters, cluster sizes, and sample sizes.

diff(), ratio(), rdiff(), rrisk(), oratio(), effect(); see [PSS-2] **power twoproportions**.

rho(*numlist*) specifies the intraclass correlation. The default is rho(0.5).

cvcluster(*numlist*) specifies the coefficient of variation for cluster sizes. This option is used with varying cluster sizes.

direction(), onesided, parallel; see [PSS-2] **power**.

_____ Table _____

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

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

_____ Graph _____

graph, graph(); see [PSS-2] **power, graph**. Also see the *column* table for a list of symbols used by the graphs.

_____ Iteration _____

init(*#*) specifies the initial value for the numbers of clusters or cluster sizes for sample-size determination or the initial value for the experimental-group proportion for the effect-size determination. The default is to use a closed-form normal approximation to compute an initial value for the estimated parameter.

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

The following option is available with power twoproportions, cluster but is not shown in the dialog box:

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

Remarks and examples

Remarks are presented under the following headings:

> *Using power twoproportions, cluster*
> *Computing numbers of clusters*
> *Computing number of clusters in one group*
> *Computing cluster sizes*
> *Computing power*
> *Computing effect size and experimental-group proportion*
> *Testing hypotheses about two proportions in a CRD*

power twoproportions, cluster requests that computations for the power twoproportions command be done for a CRD. In a CRD, groups of subjects or clusters are randomized instead of individual subjects, so the sample size is determined by the numbers of clusters and the cluster sizes. The sample-size determination thus consists of the determination of the numbers of clusters given cluster sizes or the determination of cluster sizes given the numbers of clusters. For a general discussion of using power twoproportions, see [PSS-2] **power twoproportions**. The discussion below is specific to the CRD.

Using power twoproportions, cluster

If you specify the cluster option, include k1() or k2() to specify the number of clusters or include m1() or m2() to specify the cluster size, the power twoproportions command will perform computations for a two-sample proportions test in a CRD. The computations for a CRD are based on the large-sample Pearson's χ^2 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. By default, all computations assume a balanced or equal-allocation design, meaning equal numbers of clusters and cluster sizes in both groups; see [PSS-4] **Unbalanced designs** for a description of how to specify an unbalanced design.

To compute the number of clusters in both groups, you must provide cluster sizes for both groups. There are multiple ways to supply cluster sizes, but the most common is to specify the cluster size of the control group in the m1() option and the cluster size of the experimental group in the m2() option. See *mspec* and *nspec* under *Syntax* for other specifications. When *nspec* is specified, the cluster option is also required to request that power twoproportions perform computations for a CRD. The number of clusters is assumed to be equal in the two groups, but you can change this by specifying the ratio of the numbers of clusters in the experimental to the control group in the kratio() option. Other parameters are specified as described in *Using power twoproportions* in [PSS-2] **power twoproportions**.

To compute the cluster sizes in both groups, you must provide the numbers of clusters in both groups. There are several ways to supply the numbers of clusters; see *kspec* under *Syntax*. The most common is to specify the numbers of clusters in the control group and the experimental group in the k1() and k2() options, respectively. Equal cluster sizes are assumed in the two groups, but you can change this by specifying the ratio of the cluster sizes in the experimental to that of the control group in the mratio() option. Other parameters are specified as described in *Using power twoproportions* in [PSS-2] **power twoproportions**.

You can also compute the number of clusters or the cluster size in one of the groups given the number of clusters or the cluster size in the other group by specifying the compute() option. For example, to compute the number of clusters in the control group, you specify compute(K1) and provide the number of clusters in the experimental group in k2(). Likewise, to compute the cluster

size in the control group, you specify compute(M1) and provide the cluster size of the experimental group in m2(). You can compute the number of clusters or cluster size for the experimental group in a similar manner.

Instead of the experimental-group proportion p_2, you can specify other alternative measures of effect when computing numbers of clusters, cluster sizes, or power; see *Alternative ways of specifying effect* in [PSS-2] **power twoproportions**.

The power and effect-size determination is the same as described in *Using power twoproportions* in [PSS-2] **power twoproportions**, but the sample-size information is supplied as the numbers of clusters *kspec* and either cluster sizes using *mspec* or, less commonly, sample sizes using *nspec*.

All computations assume an intraclass correlation of 0.5. You can change this by specifying the rho() option. Also, all clusters are assumed to be of the same size unless the coefficient of variation for cluster sizes is specified in the cvcluster() option.

By default, the computed numbers of clusters, cluster sizes, and sample sizes are rounded up. However, you can specify the nfractional option to see the corresponding fractional values; see *Fractional sample sizes* in [PSS-4] **Unbalanced designs** for an example. If the cvcluster() option is specified when computing cluster sizes, then cluster sizes represent average cluster sizes and are thus not rounded. When sample sizes are specified using *nspec*, fractional cluster sizes may be reported to accommodate the specified numbers of clusters and sample sizes.

Some of power twoproportions, cluster's computations require iteration, such as to compute the numbers of clusters for a two-sided test; see *Methods and formulas* for details and [PSS-2] **power** for the descriptions of options that control the iteration procedure.

Computing numbers of clusters

To compute the numbers of clusters in each group, you must either provide the cluster size for each group using *mspec* or specify the cluster option and provide the sample sizes of both groups using *nspec*. The most common method is to use *mspec* of m1() and m2(). In addition, the control- and experimental-group proportions must be specified.

▷ Example 1: Numbers of clusters for a two-sample proportions test in a CRD, specify cluster sizes

Consider a study investigating the effectiveness of a program to promote after-school activities in increasing the rate of students participating in the after-school club. Schools that are involved in the study will be randomly assigned either to the experimental group that participates in the program or to the control group that does not. A researcher plans to recruit 50 students from each school and assumes an intraclass correlation of 0.2. The researcher wants to be able to detect an increase of 0.2 in the anticipated control-group rate of 0.4, which corresponds to the experimental-group rate of 0.6.

To compute the number of schools in each group required to detect the desired rate with 80% power using a 5%-level two-sided test, we type

```
. power twoproportions 0.4 0.6, m1(50) m2(50) rho(0.2)

Performing iteration ...

Estimated numbers of clusters for a two-sample proportions test
Cluster randomized design, Pearson's chi-squared test
H0: p2 = p1  versus  Ha: p2 != p1

Study parameters:

        alpha =    0.0500
        power =    0.8000
        delta =    0.2000  (difference)
           p1 =    0.4000
           p2 =    0.6000

Cluster design:

           M1 =        50
           M2 =        50
          rho =    0.2000

Estimated numbers of clusters and sample sizes:

           K1 =        21
           K2 =        21
           N1 =     1,050
           N2 =     1,050
```

We find that for 50 students per school, 21 schools per group and thus a total of 1,050 students per group are required to detect a 0.2 difference, from 0.4 to 0.6, in participation rates in the after-school club with 80% power using a 5%-level two-sided test.

◁

▷ Example 2: Numbers of clusters for a two-sample proportions test in a CRD, specify sample sizes

Suppose that for our study, we can recruit only 1,000 students per group because of limited funding. We need to know the number of schools in each group and how many students to recruit from each school. In this case, we specify the n1(1000) and n2(1000) options. Because none of the k1(), k2(), m1(), or m2() options are specified, we also need to specify the cluster option to request the test for a CRD instead of the conventional individual-level design.

```
. power twoproportions 0.4 0.6, cluster n1(1000) n2(1000) rho(0.2)

Performing iteration ...

Estimated numbers of clusters for a two-sample proportions test
Cluster randomized design, Pearson's chi-squared test
H0: p2 = p1  versus  Ha: p2 != p1

Study parameters:

        alpha =    0.0500
        power =    0.8000
        delta =    0.2000  (difference)
           p1 =    0.4000
           p2 =    0.6000

Cluster design:

           N1 =     1,000
           N2 =     1,000
          rho =    0.2000

Estimated numbers of clusters and cluster sizes:

           K1 =        22
           K2 =        22
   Average M1 =   45.4545
   Average M2 =   45.4545
```

To achieve the desired power, we need to recruit an average of about 45 students per school from 22 schools for each of the control and experimental groups. power twoproportions, cluster did not round the cluster sizes of 45.45 to meet our required total of 1,000 students per group. In practice, you can decide to recruit 45 members from some schools and 46 from other schools to have roughly constant cluster sizes or decide to change the total number of students you want to recruit.

◁

Computing number of clusters in one group

To compute the number of clusters in one of the groups, you must specify the compute() option and the number of clusters in the other group. For example, to compute the number of clusters in the experimental group, you must specify the compute(K2) option and provide the number of clusters in the control group in the k1() option. Similarly, you can compute the number of clusters for the control group. In addition, you must provide cluster sizes *mspec* or sample sizes *nspec* of both groups and the control- and experimental-group proportions.

▷ Example 3: Number of clusters in the experimental group for a two-sample proportions test in a CRD

Continuing with example 1, suppose that we are designing a new study and we are planning to recruit 30 schools for the control group. We want to know the minimum number of schools we need to recruit to the experimental group. Given other study parameters from example 1, we compute the number of schools in the experimental group by specifying the compute(K2) option and the number of clusters in the control group of 30 in the k1() option:

```
. power twoproportions 0.4 0.6, compute(K2) k1(30) m1(50) m2(50) rho(0.2)

Performing iteration ...

Estimated treatment-group number of clusters for a two-sample proportions test
Cluster randomized design, Pearson's chi-squared test
H0: p2 = p1  versus  Ha: p2 != p1

Study parameters:

        alpha =    0.0500
        power =    0.8000
        delta =    0.2000  (difference)
           p1 =    0.4000
           p2 =    0.6000

Cluster design:

           K1 =        30
           M1 =        50
           M2 =        50
           N1 =     1,500
          rho =    0.2000

Estimated number of clusters and sample size:

           K2 =        17
           N2 =       850
```

With 30 schools in the control group, we need to recruit 17 schools for the experimental group.

◁

Computing cluster sizes

To compute cluster sizes in both groups, you must provide the numbers of clusters in both groups by using *kspec*. The most common method is to specify the numbers of clusters in the control and experimental groups in the k1() and k2() options, respectively. In addition, the control- and experimental-group proportions must be specified.

▷ Example 4: Cluster sizes for a two-sample proportions test in a CRD

Continuing with example 1, suppose that we are designing a new study and we are planning to recruit 40 schools, with 20 schools in each group. Given other study parameters from example 1, we compute the number of students to recruit from each school by specifying 20 clusters in the k1() and k2() options:

```
. power twoproportions 0.4 0.6, k1(20) k2(20) rho(0.2)

Performing iteration ...

Estimated cluster sizes for a two-sample proportions test
Cluster randomized design, Pearson's chi-squared test
H0: p2 = p1   versus   Ha: p2 != p1

Study parameters:

        alpha =    0.0500
        power =    0.8000
        delta =    0.2000   (difference)
           p1 =    0.4000
           p2 =    0.6000

Cluster design:

           K1 =        20
           K2 =        20
          rho =    0.2000

Estimated cluster sizes and sample sizes:

           M1 =       127
           M2 =       127
           N1 =     2,540
           N2 =     2,540
```

With 20 schools per group, we need to recruit 127 students per school for a total of 2,540 students per group.

◁

Computing power

To compute power in a CRD, you supply the sample-size information as the numbers of clusters by using *kspec* along with either the cluster sizes by using *mspec* or, less commonly, the sample sizes by using *nspec*. The most common method is to specify the k1(), k2(), m1(), and m2() options. In addition, the control- and experimental-group proportions must be specified.

▷ Example 5: Power for a two-sample proportions test in a CRD

Continuing with example 1, suppose that we can recruit 50 students from each of 40 schools (20 schools per group) and we want to compute power for this design. Given other study parameters from example 1, we compute the power by specifying 20 clusters in the k1() and k2() options and the cluster size of 50 in the m1() and m2() options:

```
. power twoproportions 0.4 0.6, k1(20) k2(20) m1(50) m2(50) rho(0.2)

Estimated power for a two-sample proportions test
Cluster randomized design, Pearson's chi-squared test
H0: p2 = p1  versus  Ha: p2 != p1

Study parameters:
        alpha =    0.0500
        delta =    0.2000  (difference)
           p1 =    0.4000
           p2 =    0.6000

Cluster design:
           K1 =        20
           K2 =        20
           M1 =        50
           M2 =        50
           N1 =     1,000
           N2 =     1,000
          rho =    0.2000

Estimated power:
        power =    0.7815
```

The computed power is about 78%.

◁

▷ Example 6: Multiple values of study parameters

To investigate the effect of the number of clusters in the experimental group on power, we can specify a list of numbers in the k2() option:

```
. power twoproportions 0.4 0.6, k1(20) k2(5(10)45) m1(50) m2(50) rho(0.2)
> table(power K2)

Estimated power for a two-sample proportions test
Cluster randomized design, Pearson's chi-squared test
H0: p2 = p1  versus  Ha: p2 != p1
```

power	K2
.4095	5
.7164	15
.8233	25
.8721	35
.8987	45

In this example, we also specified the table(power K2) option to list the only two columns that vary. As expected, as the number of clusters increases, the power tends to get closer to 1.

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

◁

Computing effect size and experimental-group proportion

There are multiple definitions of the effect size for a two-sample proportions test; see *Computing effect size and experimental-group proportion* in [PSS-2] **power twoproportions** for details. To compute effect size in a CRD, you supply the sample-size information as the numbers of clusters by using *kspec* along with either the cluster sizes by using *mspec* or, less commonly, the sample sizes by using *nspec*. The most common method is to specify the k1(), k2(), m1(), and m2() options. In addition, power and control-group proportion must be specified. You must also decide on the direction of the effect, which is specified in the direction() option. For the default, upper, meaning $p_2 > p_1$, power twoproportions, cluster uses direction(upper). For lower, meaning $p_2 < p_1$, specify direction(lower).

▷ Example 7: Effect size for a two-sample proportions test in a CRD

Continuing with example 5, we may also be interested in finding the minimum value of the participation rate in the after-school club that can be detected with a sample of 20 schools per group, 50 students per school, and 80% power. To compute this, we specify the control-group rate of 0.4 as the command argument and the required options k1(20), k2(20), m1(50), m2(50), and power(0.8) and continue to use rho(0.2).

```
. power twoproportions 0.4, k1(20) k2(20) m1(50) m2(50) power(0.8) rho(0.2)

Performing iteration ...

Estimated experimental-group proportion for a two-sample proportions test
Cluster randomized design, Pearson's chi-squared test
H0: p2 = p1  versus  Ha: p2 != p1; p2 > p1

Study parameters:

        alpha =    0.0500
        power =    0.8000
           p1 =    0.4000

Cluster design:

           K1 =        20
           K2 =        20
           M1 =        50
           M2 =        50
           N1 =     1,000
           N2 =     1,000
          rho =    0.2000

Estimated effect size and experimental-group proportion:

        delta =    0.2046  (difference)
           p2 =    0.6046
```

Given the proportion of 0.4 in the control group, the smallest (in absolute value) difference that can be detected is about 0.2, corresponding to a proportion of about 0.6 in the experimental group.

◁

Testing hypotheses about two proportions in a CRD

There are different ways to account for a CRD or for clustered data when performing hypothesis tests that compare proportions in two groups. With large samples or when you know intraclass correlation, the simplest way is to use a large-sample test that accounts for clustered data; see [R] **prtest** for details. More commonly, two-level binary models such as melogit are used because they also allow adjusting for covariates.

In this section, we briefly demonstrate the prtest command for comparing proportions of two groups with clustered data.

▷ Example 8: Two-sample proportions test with clustered data

Consider example 4 from [R] prtest that compared the proportions of bacterial pneumonia episodes in the two groups of infants from 18 clusters in each group. Each group received a different type of vaccine: the control group received a meningococcal C conjugate vaccine (MnCC) and the experimental group received the seven-valent pneumococcal conjugate vaccince (PnCRM7). The two groups are identified by the vaccine variable, and the pneumonia variable records 1 if an infant had at least one bacterial pneumonia episode and 0 otherwise. See example 4 in [R] prtest for other details. We replicate the analysis from that example below.

For clustered data, prtest requires that we specify the cluster identifier in the cluster() option and population intraclass correlation in the rho() option.

```
. use https://www.stata-press.com/data/r18/pneumoniacrt
(Bacterial pneumonia episodes data from CRT (Hayes and Moulton 2009))

. prtest pneumonia, by(vaccine) cluster(cluster) rho(0.02)

Two-sample test of proportions
Cluster variable: cluster
```

Group: MnCC			Group: PnCRM7		
Number of obs	=	238	Number of obs	=	211
Number of clusters	=	18	Number of clusters	=	18
Avg. cluster size	=	13.22	Avg. cluster size	=	11.72
CV cluster size	=	0.9605	CV cluster size	=	0.7976
Intraclass corr.	=	0.0200	Intraclass corr.	=	0.0200

Group	Mean	Std. err.	z	P>|z|	[95% conf. interval]	
MnCC	.2226891	.0329017			.1582029	.2871753
PnCRM7	.1658768	.0299027			.1072686	.224485
diff	.0568123	.04446			−.0303278	.1439524
	under H0:	.0447641	1.27	0.204		

```
        diff = prop(MnCC) - prop(PnCRM7)                         z =    1.2691
    H0: diff = 0

    Ha: diff < 0                    Ha: diff != 0                   Ha: diff > 0
 Pr(Z < z) = 0.8978        Pr(|Z| > |z|) = 0.2044          Pr(Z > z) = 0.1022
```

We do not have statistical evidence to reject the null hypothesis that the two group proportions are the same.

Suppose that we want to use the results of this study to design another study that compares the two vaccines in the same population. Specifically, we want to compute the required number of clusters given the average cluster sizes of 13.22 and 11.72 in the two groups, the intraclass correlation of 0.02, and the coefficient of variation of cluster sizes of 0.96, as shown in the output above. The coefficients of variation of cluster sizes are slightly different between the two groups, so we use the larger value to obtain conservative results from power twoproportions, which assumes a common coefficient of variation of cluster sizes. We also use the observed proportion estimates of 0.22 and 0.17 in the computation.

```
. power twoproportions 0.22 0.17, m1(13.22) m2(11.72) rho(0.02) cvcluster(0.96)

Performing iteration ...

Estimated numbers of clusters for a two-sample proportions test
Cluster randomized design, Pearson's chi-squared test
H0: p2 = p1  versus  Ha: p2 != p1

Study parameters:

        alpha =    0.0500
        power =    0.8000
        delta =   -0.0500  (difference)
           p1 =    0.2200
           p2 =    0.1700

Cluster design:

   Average M1 =   13.2200
   Average M2 =   11.7200
          rho =    0.0200
        CV_cl =    0.9600

Estimated numbers of clusters and sample sizes:

           K1 =       115
           K2 =       115
           N1 =     1,521
           N2 =     1,348
```

The required number of clusters for each group is 115. Given varying cluster sizes, we need to have a total of 1,521 infants in the control group and a total of 1,348 infants in the experimental group.

The observed difference of roughly 0.05 between the two proportions may be too small for the purpose of PSS computations. Suppose that we want to use a larger difference (in absolute value), for example, from 0.22 to 0.1, between the two proportions.

```
. power twoproportions 0.22 0.1, m1(13.22) m2(11.72) rho(0.02) cvcluster(0.96)

Performing iteration ...

Estimated numbers of clusters for a two-sample proportions test
Cluster randomized design, Pearson's chi-squared test
H0: p2 = p1  versus  Ha: p2 != p1

Study parameters:

        alpha =    0.0500
        power =    0.8000
        delta =   -0.1200  (difference)
           p1 =    0.2200
           p2 =    0.1000

Cluster design:

   Average M1 =   13.2200
   Average M2 =   11.7200
          rho =    0.0200
        CV_cl =    0.9600

Estimated numbers of clusters and sample sizes:

           K1 =        17
           K2 =        17
           N1 =       225
           N2 =       200
```

In this case, we need only 17 clusters and a total of 225 and 200 infants in the control and experimental groups, respectively.

◁

Stored results

`power twoproportions, cluster` 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(K1)	number of clusters in the control group
r(K2)	number of clusters in the experimental group
r(kratio)	ratio of numbers of clusters, K2/K1
r(M1)	cluster size of the control group
r(M2)	cluster size of the experimental group
r(mratio)	ratio of cluster sizes, M2/M1
r(N)	total sample size
r(N1)	sample size of the control group
r(N2)	sample size of the experimental group
r(nratio)	ratio of sample sizes, N2/N1
r(nfractional)	1 if nfractional is specified, 0 otherwise
r(onesided)	1 for a one-sided test, 0 otherwise
r(p1)	control-group proportion
r(p2)	experimental-group proportion
r(diff)	difference between the experimental- and control-group proportions
r(ratio)	ratio of the experimental-group proportion to the control-group proportion
r(rdiff)	risk difference
r(rrisk)	relative risk
r(oratio)	odds ratio
r(rho)	intraclass correlation
r(CV_cluster)	coefficient of variation for cluster sizes
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 estimated parameter
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)	twoproportions
r(design)	CRD
r(test)	chi2
r(effect)	specified effect: diff, ratio, etc.
r(direction)	upper or lower
r(columns)	displayed table columns
r(labels)	table column labels
r(widths)	table column widths
r(formats)	table column formats

Matrices
r(pss_table)	table of results

Methods and formulas

The computation in a CRD is based on the Pearson's χ^2 test under the large-sample normal approximation, adjusted by the cluster design; see *Pearson's χ^2 test* under *Methods and formulas* in [PSS-2] **power twoproportions** for the common notation for a two-sample proportions test.

In a CRD, let K_1 and K_2 be the numbers of clusters in the control and experimental groups, respectively, and M_1 and M_2 be the cluster sizes of the control and experimental groups, respectively. We have $n_1 = K_1 M_1$ and $n_2 = K_2 M_2$. Let R_k be the ratio of the numbers of clusters, K_2/K_1, and R_m be the ratio of the cluster sizes, M_2/M_1. Let ρ be the intraclass correlation coefficient and DE_1 and DE_2 be the design effect in the control and experimental groups, with

$$\text{DE}_1 = 1 + \rho(M_1 - 1) \qquad \text{and} \qquad \text{DE}_2 = 1 + \rho(M_2 - 1)$$

For unequal cluster sizes, we assume that the cluster sizes are independent and identically distributed and are small relative to the number of clusters; see Ahn, Heo, and Zhang (2015) for details. Let the coefficient of variation of the cluster sizes be CV_{cl}. According to van Breukelen, Candel, and Berger (2007) and Campbell and Walters (2014), to adjust for varying cluster sizes, define the relative efficiency of unequal versus equal cluster sizes as

$$\text{RE}_i = 1 - \lambda_i(1 - \lambda_i)\text{CV}_{\text{cl}}^2$$

where $\lambda_i = \rho M_i/(\rho M_i + 1 - \rho)$, where $i = 1$ corresponds to the control group and $i = 2$ corresponds to the experimental group.

When cluster sizes are equal, we use $\text{RE}_1 = \text{RE}_2 = 1$ in the formulas.

For a large sample size, a binomial process can be approximated by a normal distribution. Similarly to the discussion for the two-sample proportions test in the individual-level design, the asymptotic sampling distribution of the test statistic

$$z = \frac{(\widehat{p}_2 - \widehat{p}_1) - (p_2 - p_1)}{\sqrt{\overline{p}(1 - \overline{p})\left(\frac{\text{DE}_1}{n_1 \text{RE}_1} + \frac{\text{DE}_2}{n_2 \text{RE}_2}\right)}}$$

is standard normal, where the pooled proportion $\overline{p}$ is defined as

$$\overline{p} = \frac{n_1 p_1 \text{RE}_1/\text{DE}_1 + n_2 p_2 \text{RE}_2/\text{DE}_2}{n_1 \text{RE}_1/\text{DE}_1 + n_2 \text{RE}_2/\text{DE}_2}$$

The square of this statistic, z^2, has an approximate χ^2 distribution with one degree of freedom, and the corresponding test is known as Pearson's χ^2 test.

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

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

$$\pi = \begin{cases} \Phi\left\{\frac{(p_2 - p_1) - z_{1-\alpha}\sigma_p}{\sigma_D}\right\} & \text{for an upper one-sided test} \\[2mm] \Phi\left\{\frac{-(p_2 - p_1) - z_{1-\alpha}\sigma_p}{\sigma_D}\right\} & \text{for a lower one-sided test} \\[2mm] \Phi\left\{\frac{(p_2 - p_1) - z_{1-\alpha/2}\sigma_p}{\sigma_D}\right\} + \Phi\left\{\frac{-(p_2 - p_1) - z_{1-\alpha/2}\sigma_p}{\sigma_D}\right\} & \text{for a two-sided test} \end{cases} \qquad (1)$$

where $\Phi(\cdot)$ is the c.d.f. of the standard normal distribution, σ_p is the pooled standard deviation, and σ_D is the standard deviation of the difference between proportions, both defined below.

$$\sigma_p = \sqrt{\overline{p}(1 - \overline{p})\{\text{DE}_1/(n_1 \text{RE}_1) + \text{DE}_2/(n_2 \text{RE}_2)\}}$$

$$\sigma_D = \sqrt{p_1(1-p_1)\text{DE}_1/(n_1\text{RE}_1) + p_2(1-p_2)\text{DE}_2/(n_2\text{RE}_2)}$$

The above definitions of σ_p and σ_D are based on the "minimum variance weights" method in Ahn, Heo, and Zhang (2015, 32). The relative efficiencies RE_1 and RE_2 are used to adjust for varying cluster sizes approximately.

Given the cluster sizes M_1 and M_2, and intraclass correlation ρ, we can compute R_m, DE_1, DE_2, RE_1, and RE_2. With these parameters and the ratio of the numbers of clusters R_k, the numbers of clusters K_1 and K_2 for a one-sided test are computed as follows. K_1 is computed by inverting a one-sided power equation from (1),

$$K_1 = \frac{\left\{z_{1-\alpha}\sqrt{\overline{p}(1-\overline{p})} - z_\beta\sqrt{w_1 p_1(1-p_1) + w_2 p_2(1-p_2)}\right\}^2}{w_1 w_2 (p_2 - p_1)^2 (\text{RE}_1/\text{DE}_1 + R_k R_m \text{RE}_2/\text{DE}_2)}$$

where $w_1 = 1 - 1/\{1 + R_k R_m \text{DE}_1\text{RE}_2/(\text{DE}_2\text{RE}_1)\}$ and $w_2 = 1 - w_1$. Then, K_2 is computed using $K_2 = R_k K_1$.

In all other cases, parameters are computed iteratively using the power equations in (1).

References

Ahn, C., M. Heo, and S. Zhang. 2015. *Sample Size Calculations for Clustered and Longitudinal Outcomes in Clinical Research*. Boca Raton, FL: CRC Press.

Campbell, M. J., and S. J. Walters. 2014. *How to Design, Analyse and Report Cluster Randomised Trials in Medicine and Health Related Research*. Chichester, UK: Wiley.

Gallis, J. A., F. Li, H. Yu, and E. L. Turner. 2018. cvcrand and cptest: Commands for efficient design and analysis of cluster randomized trials using constrained randomization and permutation tests. *Stata Journal* 18: 357–378.

van Breukelen, G. J. P., M. J. J. M. Candel, and M. P. F. Berger. 2007. Relative efficiency of unequal versus equal cluster sizes in cluster randomized and multicentre trials. *Statistics in Medicine* 26: 2589–2603. https://doi.org/10.1002/sim.2740.

Also see

[PSS-2] **power twoproportions** — Power analysis for a two-sample proportions test

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

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

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

[PSS-5] **Glossary**

[ME] **melogit** — Multilevel mixed-effects logistic regression

[R] **prtest** — Tests of proportions