

ciwidth pairedmeans — Precision analysis for a paired-means-difference CI

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

Description

`ciwidth pairedmeans` computes sample size, CI width, and probability of CI width for a CI for the difference between two means from paired samples. It can compute sample size for a given CI width and probability of CI width. Alternatively, it can compute CI width for a given sample size and probability of CI width. It can also compute probability of CI width for a given sample size and CI width. Also see [\[PSS-3\] ciwidth](#) for PrSS analysis for other CI methods.

For power and sample-size analysis for a two-sample paired-means test, see [\[PSS-2\] power pairedmeans](#).

Quick start

Sample size required for a two-sided 95% CI for the difference between paired means to have a width no larger than 12 with a probability of 90%, assuming the standard deviation for the differences of 36

```
ciwidth pairedmeans, width(12) probwidth(0.9) sddiff(36)
```

As above, but instead of standard deviation of the differences, specify correlation between paired observations of 0.5 with pretreatment standard deviation of 29 and posttreatment standard deviation of 40

```
ciwidth pairedmeans, width(12) probwidth(0.9) corr(.5) sd1(29) sd2(40)
```

CI width for sample sizes of 20, 40, 60, and 80, given a 90% probability that the CI width will be no larger than the estimated value

```
ciwidth pairedmeans, n(20(20)80) probwidth(0.9) sddiff(36)
```

As above, but display results as a graph of CI width versus sample size

```
ciwidth pairedmeans, n(20(20)80) probwidth(0.9) sddiff(36) graph
```

Probability that the CI width is no larger than 12 for a sample size of 50

```
ciwidth pairedmeans, width(12) n(50) sddiff(36)
```

Menu

Statistics > Power, precision, and sample size

Syntax

Compute sample size

```
ciwidth pairedmeans, corrspec width(numlist) probwidth(numlist)  
[options]
```

Compute CI width

```
ciwidth pairedmeans, corrspec probwidth(numlist) n(numlist) [options]
```

Compute probability of CI width

```
ciwidth pairedmeans, corrspec width(numlist) n(numlist) [options]
```

where *corrspec* is one of

```
sddiff()
```

```
corr() [sd()]
```

```
corr() [sd1() sd2()]
```

<i>options</i>	Description
Main	
* <u>l</u> evel(<i>numlist</i>)	confidence level; default is level(95)
* <u>a</u> lpha(<i>numlist</i>)	significance level; default is alpha(0.05)
* <u>p</u> ro w idth(<i>numlist</i>)	probability of CI width; required to compute sample size and CI width
* <u>w</u> idth(<i>numlist</i>)	CI width; required to compute sample size and probability of CI width
* <u>n</u> (<i>numlist</i>)	sample size; required to compute CI width and probability of CI width
<u>n</u> fractional	allow fractional sample sizes
* <u>s</u> ddiff(<i>numlist</i>)	standard deviation σ_d of the differences; may not be combined with corr()
* <u>c</u> orr(<i>numlist</i>)	correlation between paired observations; required unless sddiff() is specified
* <u>s</u> d(<i>numlist</i>)	common standard deviation; default is sd(1) and requires corr()
* <u>s</u> d1(<i>numlist</i>)	standard deviation of the pretreatment group; requires corr()
* <u>s</u> d2(<i>numlist</i>)	standard deviation of the posttreatment group; requires corr()
<u>k</u> nownsd	request computation assuming a known standard deviation σ_d ; default is to assume an unknown standard deviation
* <u>f</u> pc(<i>numlist</i>)	finite population correction (FPC) as a sampling rate or as a population size
<u>l</u> ower	lower one-sided CI; default is two-sided CI
<u>u</u> pper	upper one-sided CI; default is two-sided CI
<u>o</u> nesided	synonym for option upper
<u>p</u> arallel	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	
[<u>n</u> o] <u>t</u> able [(<i>tablespec</i>)]	suppress table or display results as a table; see [PSS-3] ciwidth, table
<u>s</u> aving(<i>filename</i> [, replace])	save the table data to <i>filename</i> ; use replace to overwrite existing <i>filename</i>
Graph	
<u>g</u> raph [(<i>graphopts</i>)]	graph results; see [PSS-3] ciwidth, graph

<code>iteration</code>	
<code>init(#)</code>	initial value for sample size; default is to use a closed-form normal approximation
<code>iterate(#)</code>	maximum number of iterations; default is <code>iterate(500)</code>
<code>tolerance(#)</code>	parameter tolerance; default is <code>tolerance(1e-12)</code>
<code>ftolerance(#)</code>	function tolerance; default is <code>ftolerance(1e-12)</code>
<code>[no]log</code>	suppress or display iteration log
<code>[no]dots</code>	suppress or display iterations as dots
<code>notitle</code>	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.

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).

<i>column</i>	Description	Symbol
<code>level</code>	confidence level	$100(1 - \alpha)$
<code>alpha</code>	significance level	α
<code>N</code>	number of subjects	N
<code>Pr_width</code>	probability of CI width	p_{width}
<code>width</code>	CI width	w
<code>sd_d</code>	standard deviation of the differences	σ_d
<code>sd</code>	common standard deviation	σ
<code>sd1</code>	standard deviation of the pretreatment group	σ_1
<code>sd2</code>	standard deviation of the posttreatment group	σ_2
<code>corr</code>	correlation between paired observations	ρ
<code>fpc</code>	FPC as a population size	N_{pop}
	FPC as a sampling rate	γ
<code>_all</code>	display all supported columns	

Column `alpha` is shown in the default table in place of column `level` if `alpha()` is specified.

Columns `sd`, `sd1`, `sd2`, `corr`, and `fpc` are shown in the default table if the corresponding options are specified.

Options

Main

`level()`, `alpha()`, `probwidth()`, `width()`, `n()`, `nfractional`; see [PSS-3] [ciwidth](#). `probwidth()` may not be combined with `knownsd`. The `nfractional` option is allowed only for sample-size determination.

`sddiff(numlist)` specifies the standard deviation σ_d of the differences. Either `sddiff()` or `corr()` must be specified.

`corr(numlist)` specifies the correlation between paired, pretreatment and posttreatment, observations. This option along with `sd1()` and `sd2()` or `sd()` is used to compute the standard deviation of the differences unless that standard deviation is supplied directly in the `sddiff()` option. Either `corr()` or `sddiff()` must be specified.

`sd(numlist)` specifies the common standard deviation of the pretreatment and posttreatment groups. Specifying `sd(#)` implies that both `sd1()` and `sd2()` are equal to `#`. Options `corr()` and `sd()` are used to compute the standard deviation of the differences unless that standard deviation is supplied directly with the `sddiff()` option. The default is `sd1()`.

`sd1(numlist)` specifies the standard deviation of the pretreatment group. Options `corr()`, `sd1()`, and `sd2()` are used to compute the standard deviation of the differences unless that standard deviation is supplied directly with the `sddiff()` option.

`sd2(numlist)` specifies the standard deviation of the posttreatment group. Options `corr()`, `sd1()`, and `sd2()` are used to compute the standard deviation of the differences unless that standard deviation is supplied directly with the `sddiff()` option.

`knownsd` requests that the standard deviation of the differences σ_d be treated as known in the computation. By default, the standard deviation is treated as unknown, and the computation is performed for a Student's t -based CI. If `knownsd` is specified, the computation is performed for a normal-based CI. `knownsd` may not be combined with `probwidth()` and is not allowed when computing the probability of CI width.

`fpc(numlist)` requests that a finite population correction be used in the computation. If `fpc()` has values between 0 and 1, it is interpreted as a sampling rate, n/N , where N is the total number of units in the population. When sample size n is specified, if `fpc()` has values greater than n , it is interpreted as a population size, but it is an error to have values between 1 and n . For sample-size determination, `fpc()` with a value greater than 1 is interpreted as a population size. It is an error for `fpc()` to have a mixture of sampling rates and population sizes.

`lower`, `upper`, `onesided`, `parallel`; see [PSS-3] [ciwidth](#).

Table

`table`, `table()`, `notable`; see [PSS-3] [ciwidth](#), [table](#).

`saving()`; see [PSS-3] [ciwidth](#).

Graph

`graph`, `graph()`; see [PSS-3] [ciwidth](#), [graph](#). Also see the [column](#) table for a list of symbols used by the graphs.

Iteration

`init(#)` specifies an initial value for the sample size when iteration is used to compute the sample size. The default is to use a closed-form normal approximation to compute an initial sample size.

`iterate()`, `tolerance()`, `ftolerance()`, `log`, `nolog`, `dots`, `nodots`; see [PSS-3] **ciwidth**.

The following option is available with `ciwidth pairedmeans` but is not shown in the dialog box: `notitle`; see [PSS-3] **ciwidth**.

Remarks and examples

[stata.com](http://www.stata.com)

Remarks are presented under the following headings:

- Introduction*
- Using ciwidth pairedmeans*
- Computing sample size*
- Computing CI width*
- Computing probability of CI width*

This entry describes the `ciwidth pairedmeans` command and the methodology for PrSS analysis for a CI for the difference between two means from paired samples. See [PSS-3] **Intro (ciwidth)** for a general introduction to PrSS analysis, and see [PSS-3] **ciwidth** for a general introduction to the `ciwidth` command. For PSS analysis for hypothesis tests, see [PSS-2] **power**.

Introduction

The analysis of paired means is commonly used in settings such as repeated-measures designs with before and after measurements on the same individual or cross-sectional studies of paired measurements from twins. For example, a company might initiate a voluntary exercise program and measure the average weight loss of participants from the first to sixth month. Or a school district might design an intensive remedial program for students with low math scores, and then analyze how much the students' math scores improve from the pretest to the posttest. For paired data, the inference is made on the mean difference accounting for the dependence between the two groups.

To compare two paired means, we assume that the two correlated samples are drawn from two normal populations with means μ_1 and μ_2 and standard deviations σ_1 and σ_2 . The construction of a CI for the paired-means difference and its PrSS analysis are analogous to those of a CI for one population mean from *Introduction* in [PSS-3] **ciwidth onemean**, with mean μ replaced by the mean difference $\mu_d = \mu_2 - \mu_1$ and standard deviation σ by the standard deviation of the paired differences σ_d .

The `ciwidth pairedmeans` command provides PrSS analysis for a CI for the difference between two correlated means. We discuss the command details in the next section.

Using ciwidth pairedmeans

`ciwidth pairedmeans` computes sample size, CI width, or probability of CI width for a paired-means-difference CI. By default, a two-sided CI is assumed, and the confidence level is set to 95%. You may change the confidence level by specifying the `level()` option. Alternatively, you can specify the significance level in the `alpha()` option. You can specify the `upper` and `lower` options to request upper and lower one-sided CIs.

To compute sample size, you must specify the CI width in the `width()` option and the probability of CI width in the `prowidth()` option. To compute CI width, you must specify the sample size in the `n()` option and the probability of CI width in the `prowidth()` option. You can also compute the probability of CI width given the sample size in `n()` and CI width in `width()`.

For CIs for means, the CI width does not depend on the mean point estimates, the sample means, so they are not needed in the computations.

For all computations, you must specify either the standard deviation of the differences in the `sddiff()` option or the correlation between the paired observations in the `corr()` option. If you specify the `corr()` option, then individual standard deviations of the pretreatment and posttreatment groups may also be specified in the respective `sd1()` and `sd2()` options. By default, their values are set to 1. When the two standard deviations are equal, you may specify the common standard deviation in the `sd()` option instead of specifying them individually. By default, all computations assume an unknown standard deviation of the differences. When the standard deviation of the differences is known, you can specify the `knownsd` option to request a normal-based CI instead of the default Student's *t*-based CI.

By default, the computed sample size is rounded up. You can specify the `nfractional` option to see the corresponding fractional sample size; see *Fractional sample sizes* in [PSS-4] **Unbalanced designs** for an example. The `nfractional` option is allowed only for sample-size determination.

Some of `ciwidth pairedmeans`'s computations require iteration. For example, when the standard deviation of the differences is unknown, the sample-size computation requires iteration. The default initial value of the estimated sample size is obtained by using a closed-form normal approximation. It may be changed by specifying the `init()` option. See [PSS-3] **ciwidth** for the descriptions of other options that control the iteration procedure.

All computations assume an infinite population. For a finite population, use the `fpc()` option to specify a sampling rate or a population size.

In the following sections, we describe the use of `ciwidth pairedmeans` accompanied by examples for computing sample size, CI width, and probability of CI width.

Computing sample size

To compute the sample size required for a paired-means-difference CI to have the width no larger than a target width, you must specify the target CI width in the `width()` option and the desired probability of achieving the target CI width in the `probwidth()` option.

► Example 1: Sample size for a paired-means-difference CI

Consider an example similar to [example 1](#) in [PSS-2] **power pairedmeans**. We study the low birthweight (LBW) infants as in [Howell \(2002, 186\)](#). The variable of interest is the decline of the Bayley mental development index (MDI) of 24-months-old infants compared with 6-months-old infants. We use an estimate of the standard deviation of the differences of 16.04.

We want to obtain the minimum sample size that is required to obtain a two-sided 95% CI with the width of at most 6, or the margin of error of 3, for the decline of the mean MDI score in a 24-month group. Because we use an estimate of the variance in our computations, we need to account for its sampling variability to ensure that the width of the actual CI in our new study does not exceed the target width. We use the probability of 0.98 that the width of a future CI does not exceed 6. We specify it in the `probwidth()` option, and the CI width of 6 in the `width()` option, and the estimate of the standard deviation of the differences 16.04 in the `sddiff()` option.

```

. ciwidth pairedmeans, width(6) probwidth(0.98) sddiff(16.04)
Performing iteration ...
Estimated sample size for a paired-means-difference CI
Student's t two-sided CI
Study parameters:
    level =    95.0000
    Pr_width =    0.9800
    width =     6.0000
    sd_d =    16.0400
Estimated sample size:
    N =        141

```

A sample of 141 infants is required for us to be 98% certain that the CI width will be no larger than 6 for the decline of the mean MDI score. The output reminds us that this was computed assuming a two-sided CI with the 95% confidence level.

As we mentioned in the [previous section](#), sample-size determination requires iteration in the case of an unknown standard deviation. By default, `ciwidth pairedmeans` suppresses the iteration log, but we can choose to display it by specifying the `log` option.

◀

► Example 2: Specifying individual standard deviations

For instances in which you do not have the standard deviation of the differences, you can also compute the required sample size by specifying the individual standard deviations as well as the correlation among the paired observations.

Howell (2002) reported the estimates of group-specific standard deviations: 13.85 in the 6-month group and 12.95 in the 24-month group. Using the values of individual standard deviations and the standard deviation of the differences from the [previous example](#), we obtain the correlation between the observations in the 6-month group and the 24-month group to be $(13.85^2 + 12.95^2 - 16.04^2) / (2 \times 13.85 \times 12.95) = 0.285$. To compute the sample size, we specify the group-specific standard deviations in `sd1()` and `sd2()` and the correlation in `corr()`.

```

. ciwidth pairedmeans, width(6) probwidth(0.98) corr(0.285) sd1(13.85)
> sd2(12.95)
Performing iteration ...
Estimated sample size for a paired-means-difference CI
Student's t two-sided CI
Study parameters:
    level =    95.0000          sd1 =    13.8500
    Pr_width =    0.9800          sd2 =    12.9500
    width =     6.0000          corr =     0.2850
    sd_d =    16.0403
Estimated sample size:
    N =        141

```

We obtain the same sample size as in [example 1](#).

Study parameters are divided into two columns. The parameters that are always displayed are listed in the first column, and the parameters that are displayed only if they are specified are listed in the second column. The correlation and standard deviations are reported in the second column.

◀

Computing CI width

To compute the CI width, you must specify the sample size in the `n()` option and the desired probability of achieving the target CI width in the `probwidth()` option.

► Example 3: Precision of a paired-means-difference CI

Continuing with [example 1](#), suppose that we have a sample of 120 subjects. To compute the CI width for this sample size, we specify 120 in the `n()` option and replace the `width()` option with it. We leave the other parameters unchanged:

```
. ciwidth pairedmeans, n(120) probwidth(0.98) sddiff(16.04)
Estimated width for a paired-means-difference CI
Student's t two-sided CI
Study parameters:
      level =   95.0000
         N =    120
Pr_width =   0.9800
      sd_d =  16.0400
Estimated width:
      width =   6.5705
```

Compared with the CI width of 6 in [example 1](#), the CI width of 6.6 in this example is larger, as would be expected with a smaller sample size of 120.

◀

► Example 4: One-sided CI

Rather than performing computations for the default two-sided CIs, we can request `upper` and `lower` one-sided CIs. We modify the command from [example 3](#) by instead requesting a lower one-sided 95% CI for the difference of the MDI scores:

```
. ciwidth pairedmeans, n(120) probwidth(0.98) sddiff(16.04) lower
Estimated width for a paired-means-difference CI
Student's t lower CI
Study parameters:
      level =   95.0000
         N =    120
Pr_width =   0.9800
      sd_d =  16.0400
Estimated width:
      width =   2.7504
```

Given the sample size of 120, the estimated width for a 95% lower one-sided CI for the differences of the MDI scores is about 2.75.

◀

► Example 5: Multiple values of study parameters

To investigate the effect of sample size on CI width, we can specify a list of sample sizes in the `n()` option:

```
. ciwidth pairedmeans, n(100(10)150) probwidth(0.98) sddiff(16.04)
```

Estimated width for a paired-means-difference CI

Student's t two-sided CI

level	N	Pr_width	width	sd_d
95	100	.98	7.294	16.04
95	110	.98	6.905	16.04
95	120	.98	6.57	16.04
95	130	.98	6.278	16.04
95	140	.98	6.021	16.04
95	150	.98	5.792	16.04

As expected, when the sample size increases, the CI width decreases.

For multiple values of parameters, the results are automatically displayed in a table. For more examples of tables, see [PSS-3] [ciwidth, table](#). If you wish to produce sample-size and other curves, see [PSS-3] [ciwidth, graph](#).

◀

Computing probability of CI width

To compute the probability that the width of a future CI will be no larger than the specified width, you must specify the sample size in the `n()` option and the target CI width in the `width()` option.

► Example 6: Probability of CI width for a paired-means-difference CI

Continuing with [example 1](#), let's determine how certain we can be that the CI width in a future study will be no larger than a prespecified value for a given sample size. We use the same CI width and standard deviation as in [example 1](#), and we use a sample size of 120 in `n()`.

```
. ciwidth pairedmeans, n(120) width(6) sddiff(16.04)
```

Estimated probability of width for a paired-means-difference CI

Student's t two-sided CI

Study parameters:

level = 95.0000

N = 120

width = 6.0000

sd_d = 16.0400

Estimated probability of width:

Pr_width = 0.7175

The estimated probability is 72%. Compared with [example 1](#), with a smaller sample size of 120, we are about 26% less certain that the CI width in a future study will be no larger than 6.

◀

Stored results

ciwidth pairedmeans stores the following in `r()`:

Scalars

<code>r(level)</code>	confidence level
<code>r(alpha)</code>	significance level
<code>r(N)</code>	sample size
<code>r(nfractional)</code>	1 if <code>nfractional</code> is specified, 0 otherwise
<code>r(onesided)</code>	1 for a one-sided CI, 0 otherwise
<code>r(Pr_width)</code>	probability of CI width
<code>r(Pr_width_a)</code>	actual probability of CI width (for sample-size determination when <code>proewidth()</code> specified)
<code>r(width)</code>	CI width
<code>r(width_a)</code>	actual CI width (for sample-size determination when <code>knownsd</code> specified)
<code>r(corr)</code>	correlation between paired observations
<code>r(sd_d)</code>	standard deviation of the differences
<code>r(sd1)</code>	standard deviation of the pretreatment group
<code>r(sd2)</code>	standard deviation of the posttreatment group
<code>r(sd)</code>	common standard deviation
<code>r(knownsd)</code>	1 if option <code>knownsd</code> is specified, 0 otherwise
<code>r(fpc)</code>	finite population correction (if specified)
<code>r(separator)</code>	number of lines between separator lines in the table
<code>r(divider)</code>	1 if <code>divider</code> is requested in the table, 0 otherwise
<code>r(init)</code>	initial value for sample size
<code>r(maxiter)</code>	maximum number of iterations
<code>r(iter)</code>	number of iterations performed
<code>r(tolerance)</code>	requested parameter tolerance
<code>r(deltax)</code>	final parameter tolerance achieved
<code>r(ftolerance)</code>	requested distance of the objective function from zero
<code>r(function)</code>	final distance of the objective function from zero
<code>r(converged)</code>	1 if iteration algorithm converged, 0 otherwise

Macros

<code>r(type)</code>	ci
<code>r(method)</code>	pairedmeans
<code>r(onesidedci)</code>	upper or lower (for a one-sided CI)
<code>r(columns)</code>	displayed table columns
<code>r(labels)</code>	table column labels
<code>r(widths)</code>	table column widths
<code>r(formats)</code>	table column formats

Matrices

<code>r(pss_table)</code>	table of results
---------------------------	------------------

Methods and formulas

Consider a sequence of n paired observations denoted by x_{ij} for $i = 1, \dots, n$ and groups $j = 1, 2$. Individual observations correspond to the pair (x_{i1}, x_{i2}) , and inference is made on the differences within the pairs. Let $\mu_d = \mu_2 - \mu_1$ denote the mean difference, where μ_j is the population mean of group j , and $D_i = x_{i2} - x_{i1}$ denote the difference between paired individual observations.

PrSS analysis for a paired-means-difference CI is analogous to a one-mean CI where the sample of differences D_i 's is treated as a single sample (Dixon and Massey 1983). The standard deviation of the differences, σ_d , is used in place of the one-sample standard deviation. For more information, see *Methods and formulas* in [PSS-3] **ciwidth onemean**.

References

- Dixon, W. J., and F. J. Massey, Jr. 1983. *Introduction to Statistical Analysis*. 4th ed. New York: McGraw-Hill.
- Howell, D. C. 2002. *Statistical Methods for Psychology*. 5th ed. Belmont, CA: Wadsworth.

Also see

- [PSS-3] **ciwidth** — Precision and sample-size analysis for CIs
- [PSS-3] **ciwidth, graph** — Graph results from the ciwidth command
- [PSS-3] **ciwidth, table** — Produce table of results from the ciwidth command
- [PSS-2] **power pairedmeans** — Power analysis for a two-sample paired-means test
- [PSS-5] **Glossary**
- [R] **ttest** — t tests (mean-comparison tests)
- [R] **ztest** — z tests (mean-comparison tests, known variance)