

power onemean — Power analysis for a one-sample mean test

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

Description

`power onemean` computes sample size, power, or target mean for a one-sample mean test. By default, it computes sample size for given power and the values of the mean parameters under the null and alternative hypotheses. Alternatively, it can compute power for given sample size and values of the null and alternative means or the target mean for given sample size, power, and the null mean. For power and sample-size analysis in a cluster randomized design, see [PSS-2] [power onemean, cluster](#). Also see [PSS-2] [power](#) for a general introduction to the `power` command using hypothesis tests.

For precision and sample-size analysis for a CI for a population mean, see [PSS-3] [ciwidth onemean](#).

Quick start

Sample size for two-sided test of $H_0: \mu = 10$ versus $H_a: \mu \neq 10$ with null mean $m_0 = 10$, alternative mean $m_a = 15$, and standard deviation of 12 using default power of 0.8 and significance level $\alpha = 0.05$

```
power onemean 10 15, sd(12)
```

As above, but for a one-sided test with power of 0.9

```
power onemean 10 15, sd(12) power(.9) onesided
```

Same as above, but specified as m_0 and difference $m_a - m_0 = 5$

```
power onemean 10, sd(12) power(.9) onesided diff(5)
```

Power for a sample size of 75

```
power onemean 10 15, sd(12) n(75)
```

Power for sample sizes of 50, 60, 70, and 80

```
power onemean 10 15, sd(12) n(50(10)80)
```

As above, but display results in a graph of power versus sample size

```
power onemean 10 15, sd(12) n(50(10)80) graph
```

Effect size and target mean for $m_0 = 10$ with standard deviation of 4, for a sample size of 40, power of 0.9, and $\alpha = 0.01$

```
power onemean 10, sd(4) n(40) power(.9) alpha(.01)
```

Menu

Statistics > Power, precision, and sample size

Syntax

Compute sample size

```
power onemean  $m_0$   $m_a$  [ , power(numlist) options ]
```

Compute power

```
power onemean  $m_0$   $m_a$  , n(numlist) [options]
```

Compute effect size and target mean

```
power onemean  $m_0$  , n(numlist) power(numlist) [options]
```

where m_0 is the null (hypothesized) mean or the value of the mean under the null hypothesis and m_a is the alternative (target) mean or the value of the mean under the alternative hypothesis. m_0 and m_a may each be specified either as one number or as a list of values in parentheses (see [U] [11.1.8 numlist](#)).

<i>options</i>	Description
Main	
* <u>a</u> lpha(<i>numlist</i>)	significance level; default is alpha(0.05)
* <u>p</u> ower(<i>numlist</i>)	power; default is power(0.8)
* <u>b</u> eta(<i>numlist</i>)	probability of type II error; default is beta(0.2)
* <u>n</u> (<i>numlist</i>)	sample size; required to compute power or effect size
<u>n</u> fractional	allow fractional sample size
* <u>d</u> iff(<i>numlist</i>)	difference between the alternative mean and the null mean, $m_a - m_0$; specify instead of the alternative mean m_a
* <u>s</u> d(<i>numlist</i>)	standard deviation; default is sd(1)
knownsd	request computation assuming a known standard deviation; 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>d</u> irection(<u>upper</u> <u>lower</u>)	direction of the effect for effect-size determination; default is direction(<u>upper</u>), which means that the postulated value of the parameter is larger than the hypothesized value
<u>o</u> nesided	one-sided test; default is two sided
<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>[no]</u> table [<i>tablespec</i>]	suppress table or display results as a table; see [PSS-2] power, 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-2] power, graph
Iteration	
<u>i</u> nit(#)	initial value for sample size or mean; default is to use normal approximation
<u>i</u> terate(#)	maximum number of iterations; default is iterate(500)
<u>t</u> olerance(#)	parameter tolerance; default is tolerance(1e-12)
<u>f</u> tolerance(#)	function tolerance; default is ftolerance(1e-12)
<u>[no]</u> log	suppress or display iteration log
<u>[no]</u> dots	suppress or display iterations as dots
<u>cluster</u>	perform computations for a CRD; see [PSS-2] power onemean, cluster
<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**.

cluster and notitle do 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
alpha	significance level	α
power	power	$1 - \beta$
beta	type II error probability	β
N	number of subjects	N
delta	effect size	δ
m0	null mean	μ_0
ma	alternative mean	μ_a
diff	difference between the alternative and null means	$\mu_a - \mu_0$
sd	standard deviation	σ
fpc	FPC as population size	N_{pop}
	FPC as sampling rate	γ
target	target parameter; synonym for ma	
_all	display all supported columns	

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

Columns diff and fpc are shown in the default table if specified.

Options

Main

`alpha()`, `power()`, `beta()`, `n()`, `nfractional`; see [PSS-2] **power**. The `nfractional` option is allowed only for sample-size determination.

`diff(numlist)` specifies the difference between the alternative mean and the null mean, $m_a - m_0$. You can specify either the alternative mean m_a as a command argument or the difference between the two means in `diff()`. If you specify `diff(#)`, the alternative mean is computed as $m_a = m_0 + \#$. This option is not allowed with the effect-size determination.

`sd(numlist)` specifies the sample standard deviation or the population standard deviation. The default is `sd(1)`. By default, `sd()` specifies the sample standard deviation. If `knownsd` is specified, `sd()` specifies the population standard deviation.

`knownsd` requests that the standard deviation be treated as known in the computation. By default, the standard deviation is treated as unknown, and the computation is based on a t test, which uses a Student's t distribution as a sampling distribution of the test statistic. If `knownsd` is specified, the computation is based on a z test, which uses a normal distribution as the sampling distribution of the test statistic.

`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.

`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 of the sample size for the sample-size determination or the initial value of the mean for the effect-size determination. The default is to use a closed-form normal approximation to compute an initial value of the sample size or mean.

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

The following options are available with `power onemean` but are not shown in the dialog box:

`cluster`; see [PSS-2] [power onemean, cluster](#).

`notitle`; see [PSS-2] [power](#).

Remarks and examples

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

Remarks are presented under the following headings:

- [Introduction](#)
- [Using power onemean](#)
- [Computing sample size](#)
- [Computing power](#)
- [Computing effect size and target mean](#)
- [Performing hypothesis tests on mean](#)
- [Video examples](#)

This entry describes the `power onemean` command and the methodology for power and sample-size analysis for a one-sample mean test. See [PSS-2] [Intro \(power\)](#) for a general introduction to power and sample-size analysis, and see [PSS-2] [power](#) for a general introduction to the `power` command using hypothesis tests. Also see [PSS-2] [power onemean, cluster](#) for power and sample-size analysis in a cluster randomized design.

Introduction

There are many examples of studies where a researcher would like to compare an observed mean with a hypothesized mean. A company that provides preparatory classes for a standardized exam would like to see if the mean score of students who take its classes is higher than the national average. A fitness center would like to know if its average clients' weight loss is greater than zero after six months. Or a government agency would like to know if a job training program results in higher wages than the national average.

This entry describes power and sample-size analysis for the inference about the population mean performed using hypothesis testing. Specifically, we consider the null hypothesis $H_0: \mu = \mu_0$ versus the two-sided alternative hypothesis $H_a: \mu \neq \mu_0$, the upper one-sided alternative $H_a: \mu > \mu_0$, or the lower one-sided alternative $H_a: \mu < \mu_0$.

The considered one-sample mean tests rely on the assumption that a random sample is normally distributed or that the sample size is large. Different test statistics can be based on whether the variance of the sampling process is known a priori. In the case of a known variance, the test statistic follows a standard normal distribution under the null hypothesis, and the corresponding test is known as a z test. In the case of an unknown variance, an estimate of the population variance is used to form a test statistic, which follows a Student's t distribution under the null hypothesis, and the corresponding test is known as a t test.

The random sample is typically drawn from an infinite population. When the sample is drawn from a population of a fixed size, sampling variability must be adjusted for a finite population size.

The `power onemean` command provides power and sample-size analysis for the comparison of a mean with a reference value using a t test or a z test.

Using `power onemean`

`power onemean` computes sample size, power, or target mean for a one-sample mean 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 are based on a t test, which assumes an unknown standard deviation, and use the default value of 1 as the estimate of the standard deviation. You may specify other values for the standard deviation in the `sd()` option. For a known standard deviation, you can specify the `knownsd` option to request a z test.

To compute sample size, you must specify the means under the null and alternative hypotheses, m_0 and m_a , respectively, and, optionally, the power of the test in the `power()` option. The default power is set to 0.8.

To compute power, you must specify the sample size in the `n()` option and the means under the null and alternative hypotheses, m_0 and m_a , respectively.

Instead of the alternative mean, m_a , you may specify the difference $m_a - m_0$ between the alternative mean and the null mean in the `diff()` option when computing sample size or power.

To compute effect size, the standardized difference between the alternative and null means, and the corresponding target mean, you must specify the sample size in the `n()` option, the power in the `power()` option, the null mean m_0 , and, optionally, the direction of the effect. The direction is upper by default, `direction(upper)`, which means that the target mean is assumed to be larger than the specified null mean value. This is also equivalent to the assumption of a positive effect size. You can change the direction to lower, which means that the target mean is assumed to be smaller than the specified null value, by specifying the `direction(lower)` option. This is equivalent to assuming a negative effect size.

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 `power onemean`'s computations require iteration. For example, when the standard deviation is unknown, computations use a noncentral Student's t distribution. Its degrees of freedom depends on the sample size, and the noncentrality parameter depends on the sample size and effect size. Therefore, the sample-size and effect-size determinations require iteration. The default initial values of the estimated parameters are obtained by using a closed-form normal approximation. They may be changed by specifying the `init()` option. See [PSS-2] **power** 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. When this option is specified, a finite population correction is applied to the population standard deviation. The correction factor depends on the sample size; therefore, computing sample size for a finite population requires iteration even for a known standard deviation. The initial value for the sample size is based on the corresponding sample size assuming an infinite population.

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

Computing sample size

To compute sample size, you must specify the means under the null and alternative hypotheses, m_0 and m_a , respectively, and, optionally, the power of the test in the `power()` option. A default power of 0.8 is assumed if `power()` is not specified.

► Example 1: Sample size for a one-sample mean test

Consider an example from [Tamhane and Dunlop \(2000, 209\)](#) that discusses the effectiveness of coaching programs in improving the verbal part of SAT scores. Previous studies found that students retaking the SAT exams without any coaching program improve their scores by 15 points on average with a standard deviation of about 40 points. A new coaching program claims to improve the SAT scores by 40 points above the average. The changes in scores are assumed to be approximately normally distributed. The parameter of interest in this example is the mean change in the test scores. To test the claim, investigators wish to conduct another study and compute the sample size that is required to detect a mean change in scores of 40 points with 80% power using a 5%-level two-sided test. We assume that the true population standard deviation is unknown and use its estimate from previous studies to compute the sample size:

```
. power onemean 15 40, sd(40)
Performing iteration ...
Estimated sample size for a one-sample mean test
t test
H0: m = m0 versus Ha: m != m0
Study parameters:
      alpha =    0.0500
      power =    0.8000
      delta =    0.6250
      m0 =    15.0000
      ma =    40.0000
      sd =    40.0000
Estimated sample size:
      N =          23
```

We find that a sample of 23 subjects is required to detect a shift of 40 points in average SAT scores given the standard deviation of 40 points with 80% power using a 5%-level two-sided test.

As we mentioned in [Using power onemean](#) and as is also indicated in the output, sample-size computation requires iteration when the standard deviation is unknown. The iteration log is suppressed by default, but you can display it by specifying the `log` option.

▷ Example 2: Specifying difference between means

Instead of the alternative mean change of 40 as in [example 1](#), we can specify the difference of 25 between the mean changes in scores under the alternative and null hypotheses in the `diff()` option and obtain the same results.

```
. power onemean 15, diff(25) sd(40)
Performing iteration ...
Estimated sample size for a one-sample mean test
t test
H0: m = m0 versus Ha: m != m0
Study parameters:
    alpha =    0.0500
    power =    0.8000
    delta =    0.6250
    m0 =    15.0000
    ma =    40.0000
    diff =    25.0000
    sd =    40.0000
Estimated sample size:
    N =          23
```

When we specify the `diff()` option, the difference between the alternative and null values is also reported in the output.



▷ Example 3: Known variance

If we know the population standard deviation, we can use the `knownsd` option to request a z test.

```
. power onemean 15 40, sd(40) knownsd
Performing iteration ...
Estimated sample size for a one-sample mean test
z test
H0: m = m0 versus Ha: m != m0
Study parameters:
    alpha =    0.0500
    power =    0.8000
    delta =    0.6250
    m0 =    15.0000
    ma =    40.0000
    sd =    40.0000
Estimated sample size:
    N =          21
```

The output now indicates that the computation is based on a z test instead of a t test. We find that a smaller sample of 21 subjects is required to detect the same effect size as in [example 1](#) when the standard deviation is known.



Computing power

To compute power, you must specify the sample size in the `n()` option and the means under the null and alternative hypotheses, m_0 and m_a , respectively.

► Example 4: Power of a one-sample mean test

Continuing with [example 1](#), we will suppose that we are designing a new study and anticipate to obtain a sample of 30 subjects. To compute the power corresponding to this sample size given the study parameters from [example 1](#), we specify the sample size of 30 in the `n()` option:

```
. power onemean 15 40, n(30) sd(40)
Estimated power for a one-sample mean test
t test
H0: m = m0 versus Ha: m != m0
Study parameters:
      alpha =    0.0500
         N =     30
      delta =    0.6250
         m0 =   15.0000
         ma =   40.0000
         sd =   40.0000
Estimated power:
      power =    0.9112
```

With a larger sample size, the power of the test increases to about 91.12%.

◀

► Example 5: Multiple values of study parameters

To investigate the effect of a finite population size on power, we can specify a list of population sizes in the `fpc()` option:

```
. power onemean 15 40, n(30) sd(40) fpc(100 500 1000)
Estimated power for a one-sample mean test
t test
H0: m = m0 versus Ha: m != m0
```

alpha	power	N	delta	m0	ma	sd	fpc
.05	.9769	30	.625	15	40	40	100
.05	.9267	30	.625	15	40	40	500
.05	.919	30	.625	15	40	40	1000

As expected, when the population size increases, the power tends to get closer to that obtained by assuming an infinite population size.

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](#).

◀

▷ Example 6: Reproducing published results from a text book

We can also reproduce the example from [Tamhane and Dunlop \(2000, 213–214\)](#). The authors consider a one-sided test with a 0.132 significance level and a known standard deviation and compute the power to be 95.3%. We can replicate their example by typing

```
. power onemean 15 40, n(20) sd(40) alpha(0.132) onesided knownsd
Estimated power for a one-sample mean test
z test
H0: m = m0 versus Ha: m > m0
Study parameters:
  alpha = 0.1320
  N = 20
  delta = 0.6250
  m0 = 15.0000
  ma = 40.0000
  sd = 40.0000
Estimated power:
  power = 0.9533
```

◀

Computing effect size and target mean

Effect size δ for a one-sample mean test is defined as the ratio of the difference between the alternative and null values of the mean to the standard deviation, $\delta = (\mu_a - \mu_0)/\sigma$.

Sometimes, we may be interested in determining the smallest effect and the corresponding alternative or target mean that yield a statistically significant result for prespecified sample size and power. In this case, power, sample size, and null mean must be specified. In addition, you must also decide on the direction of the effect: upper, which means $\mu_a > \mu_0$, or lower, which means $\mu_a < \mu_0$. The direction may be specified in the `direction()` option; `direction(upper)` is the default.

▷ Example 7: Minimum detectable value of the mean change in SAT scores

Continuing with [example 4](#), we may also be interested to find the smallest mean change in SAT scores that can be detected with a power of 80% given a sample of 30 subjects. To compute this, we specify only the null value of 15 as the command argument and also specify the sample size and power in the `n(30)` and `power(0.8)` options, respectively. We use the same value of 40 for the standard deviation as in [example 4](#).

```
. power onemean 15, n(30) power(0.8) sd(40)
Performing iteration ...
Estimated target mean for a one-sample mean test
t test
H0: m = m0 versus Ha: m != m0; ma > m0
Study parameters:
  alpha = 0.0500
  power = 0.8000
  N = 30
  m0 = 15.0000
  sd = 40.0000
Estimated effect size and target mean:
  delta = 0.5292
  ma = 36.1694
```

The estimated smallest mean change in SAT scores is 36.17, which corresponds to the effect size of 0.53. Compared with [example 1](#), for the same power of 80%, this example shows a smaller difference between the mean SAT scores of the two programs for a larger sample of 30 subjects.

In the above, we assumed the effect to be in the upper direction. By symmetry, the effect size in the lower direction will be -0.53 , which can also be obtained by specifying `direction(lower)` in the above example.

◀

Performing hypothesis tests on mean

In this section, we briefly demonstrate the use of the `ttest` command for testing hypotheses about means; see [\[R\] ttest](#) for details. Suppose we wish to test the hypothesis that the mean is different from a reference value on the collected data. We can use the `ttest` command to do this. Below we demonstrate the use of this command for the analysis of `sat.dta`.

▶ Example 8: Testing for mean

Suppose that we wish to test whether the mean verbal SAT score is equal to 600. We use the `ttest` command to do this as follows:

```
. use https://www.stata-press.com/data/r17/sat
(Fictional SAT data)
. ttest score == 600
```

One-sample t test

Variable	Obs	Mean	Std. err.	Std. dev.	[95% conf. interval]
score	75	504.8	15.24616	132.0356	474.4214 535.1786

```
mean = mean(score)                                t = -6.2442
HO: mean = 600                                    Degrees of freedom = 74
Ha: mean < 600                                    Ha: mean != 600      Ha: mean > 600
Pr(T < t) = 0.0000                                Pr(|T| > |t|) = 0.0000 Pr(T > t) = 1.0000
```

We find statistical evidence to reject the null hypothesis of $H_0: \mu_{\text{SAT}} = 600$ versus a two-sided alternative $H_a: \mu_{\text{SAT}} \neq 600$ at the 5% significance level; the p -value < 0.0000 .

We use the estimates based on this study to perform a sample-size analysis we would have conducted before the study.

```
. power onemean 600 505, sd(132)
Performing iteration ...
Estimated sample size for a one-sample mean test
t test
HO: m = m0 versus Ha: m != m0
Study parameters:
alpha = 0.0500
power = 0.8000
delta = -0.7197
m0 = 600.0000
ma = 505.0000
sd = 132.0000
Estimated sample size:
N = 18
```

We find that the sample size required to detect a mean score of 505 with 80% power using a 5%-level two-sided test is only 18. The current sample contains 75 subjects, which would allow us to detect a potentially smaller (in absolute value) difference between the alternative and null means.



Video examples

[Sample-size calculation for comparing a sample mean to a reference value](#)

[Power calculation for comparing a sample mean to a reference value](#)

[Minimum detectable effect size for comparing a sample mean to a reference value](#)

Stored results

`power onemean` stores the following in `r()`:

Scalars

<code>r(alpha)</code>	significance level
<code>r(power)</code>	power
<code>r(beta)</code>	probability of a type II error
<code>r(delta)</code>	effect size
<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 test, 0 otherwise
<code>r(m0)</code>	mean under the null hypothesis
<code>r(ma)</code>	mean under the alternative hypothesis
<code>r(diff)</code>	difference between the alternative and null means
<code>r(sd)</code>	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 or mean
<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>	test
<code>r(method)</code>	<code>onemean</code>
<code>r(direction)</code>	upper or lower
<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

Let x_1, \dots, x_n be a sequence of n independent and identically distributed random variables drawn from a normal population with mean μ and variance σ^2 . Let

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad \text{and} \quad s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

denote the sample mean and the sample variance, respectively. Let μ_0 and μ_a denote the null and alternative values of the mean parameter, respectively.

A one-sample mean test involves testing the null hypothesis $H_0: \mu = \mu_0$ versus the two-sided alternative hypothesis $H_a: \mu \neq \mu_0$, the upper one-sided alternative $H_a: \mu > \mu_0$, or the lower one-sided alternative $H_a: \mu < \mu_0$.

If the `nfractional` option is not specified, the computed sample size is rounded up.

The following formulas are based on [Chow et al. \(2018\)](#).

Methods and formulas are presented under the following headings:

Known standard deviation
Unknown standard deviation
Finite population size

Known standard deviation

In the case of a known standard deviation, the sampling distribution of the test statistic $z = \sqrt{n}(\bar{x} - \mu_0)/\sigma$ under the null hypothesis follows the standard normal distribution, and the corresponding test is known as a z 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 β th quantiles of the standard normal distribution.

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

$$\pi = \begin{cases} \Phi(\sqrt{n}\delta - z_{1-\alpha}) & \text{for an upper one-sided test} \\ \Phi(-\sqrt{n}\delta - z_{1-\alpha}) & \text{for a lower one-sided test} \\ \Phi(\sqrt{n}\delta - z_{1-\alpha/2}) + \Phi(-\sqrt{n}\delta - z_{1-\alpha/2}) & \text{for a two-sided test} \end{cases} \quad (1)$$

where $\Phi(\cdot)$ is the cdf of the standard normal distribution and $\delta = (\mu_a - \mu_0)/\sigma$ is the effect size.

The sample size n for a one-sided test is computed by inverting a one-sided power equation from (1):

$$n = \left(\frac{z_{1-\alpha} - z_\beta}{\delta} \right)^2 \quad (2)$$

Similarly, the absolute value of the effect size for a one-sided test is computed as follows:

$$|\delta| = \frac{(z_{1-\alpha} - z_\beta)}{\sqrt{n}} \quad (3)$$

Note that the magnitude of the effect size is the same regardless of the direction of the test.

The minimum detectable value of the mean for a one-sided test is computed as $\mu_a = \mu_0 + (z_{1-\alpha} - z_\beta)\sigma/\sqrt{n}$ when $\mu_a > \mu_0$ and as $\mu_a = \mu_0 - (z_{1-\alpha} - z_\beta)\sigma/\sqrt{n}$ when $\mu_a < \mu_0$.

Sample size and minimum detectable value of the mean for a two-sided test are computed iteratively using the two-sided power equation from (1). The initial values are obtained from (2) and (3), correspondingly, with $\alpha/2$ in place of α .

Unknown standard deviation

In the case of an unknown standard deviation, an unbiased estimator s is used in place of σ in the definition of a z test statistic. The sampling distribution of the test statistic $t = \sqrt{n}(\bar{x} - \mu_0)/s$ under the null hypothesis follows a Student's t distribution with $n - 1$ degrees of freedom, and the corresponding test is known as a t test.

Let $t_{n-1,\alpha}$ denote the α th quantile of a Student's t distribution with $n - 1$ degrees of freedom. Under the alternative hypothesis, the test statistic follows a noncentral Student's t distribution, and the power is computed using

$$\pi = \begin{cases} 1 - T_{n-1,\lambda}(t_{n-1,1-\alpha}) & \text{for an upper one-sided test} \\ T_{n-1,\lambda}(-t_{n-1,1-\alpha}) & \text{for a lower one-sided test} \\ 1 - T_{n-1,\lambda}(t_{n-1,1-\alpha/2}) + T_{n-1,\lambda}(-t_{n-1,1-\alpha/2}) & \text{for a two-sided test} \end{cases} \quad (4)$$

where $T_{n-1,\lambda}(\cdot)$ is the cumulative noncentral Student's t distribution with a noncentrality parameter $\lambda = \sqrt{n}\delta$.

Sample size and minimum detectable value of the mean are obtained by iteratively solving nonlinear equations in (4), for n and δ , respectively. The default initial values for the iterative procedure are calculated from (2) and (3), respectively, assuming a normal distribution.

Finite population size

The above formulas assume that the random sample is drawn from an infinite population. In cases when the size of the population is known, we need to make the following adjustment to the standard deviation,

$$\sigma_{\text{fpc}} = \sigma \sqrt{\left(1 - \frac{n}{N}\right)}$$

where σ_{fpc} is the population standard deviation adjusted for finite population size. The correction factor depends on the sample size; therefore, computing sample size in this case requires iteration. The initial value for the sample size is based on the corresponding normal approximation with infinite population size.

References

- Chow, S.-C., J. Shao, H. Wang, and Y. Lohknygina. 2018. *Sample Size Calculations in Clinical Research*. 3rd ed. Boca Raton, FL: CRC Press.
- Tamhane, A. C., and D. D. Dunlop. 2000. *Statistics and Data Analysis: From Elementary to Intermediate*. Upper Saddle River, NJ: Prentice Hall.

Also see

[PSS-2] **power onemean, cluster** — Power analysis for a one-sample mean test, CRD

[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-3] **ciwidth onemean** — Precision analysis for a one-mean CI

[PSS-5] **Glossary**

[R] **ttest** — t tests (mean-comparison tests)