Title

ztest — z tests (mean-comparison tests, known variance)

Description

`ztest` performs z tests on the equality of means, assuming known variances. The test can be performed for one sample against a hypothesized population value or for no difference in population means estimated from two samples. Two-sample tests can be conducted for paired and unpaired data. Clustered data are also supported.

`ztesti` is the immediate form of `ztest`; see **[U] 19 Immediate commands**.

For the comparison of means when variances are unknown, use `ttest`; see **[R] ttest**.

Quick start

One-sample test that the mean of `v1` is 3 at the 90% confidence level

```
ztest v1 == 3, level(90)
```

Same as above, and adjust for clustering with clusters defined by `cvar` and an intraclass correlation of 0.5

```
ztest v1 == 3, level(90) cluster(cvar) rho(0.5)
```

Unpaired z test that the mean of `v1` is equal between two groups defined by `catvar`

```
ztest v1, by(catvar)
```

Same as above, and adjust for clustering with clusters defined by `cvar` and an intraclass correlation of 0.5 in the two groups

```
ztest v1, by(catvar) cluster(cvar) rho(0.5)
```

Unpaired test of equality of the means of `v2` and `v3`

```
ztest v2 == v3, unpaired
```

Paired test of equality of the means of `v2` and `v3` with standard deviation of the differences between paired observations of 2.4

```
ztest v2 == v3, sddiff(2.4)
```

Same as above, specified using a common standard deviation of 2 and correlation between observations of 0.28

```
ztest v2 == v3, sd(2) corr(0.28)
```

Immediate form unpaired test of $\mu_1 = \mu_2$ if $\overline{x}_1 = 3.2$, $\text{sd}_1 = 0.1$, $\overline{x}_2 = 3.4$, and $\text{sd}_2 = 0.15$ with $n_1 = n_2 = 120$

```
ztesti 120 3.2 0.1 120 3.4 0.15
```

Menu

ztest

Statistics > Summaries, tables, and tests > Classical tests of hypotheses > z test (mean-comparison test, known variance)

ztesti

Statistics > Summaries, tables, and tests > Classical tests of hypotheses > z test calculator

Syntax

One-sample z test

`ztest` *varname* `==` # [*if*] [*in*] [, *onesampleopts*]

Two-sample z test using groups

`ztest` *varname* [*if*] [*in*] , `by(`*groupvar*`)` [*twosamplegropts*]

Two-sample z test using variables

`ztest` $varname_1$ `==` $varname_2$ [*if*] [*in*] , `unpaired` [*twosamplevaropts*]

Paired z test

`ztest` $varname_1$ `==` $varname_2$ [*if*] [*in*] , `sddiff(`#`)` [`level(`#`)`]

`ztest` $varname_1$ `==` $varname_2$ [*if*] [*in*] , `corr(`#`)` [*pairedopts*]

Immediate form of one-sample z test

`ztesti` $\#_{\rm obs}$ $\#_{\rm mean}$ $\#_{\rm sd}$ $\#_{\rm val}$ [, `level(`#`)`]

Immediate form of two-sample unpaired z test

`ztesti` $\#_{\rm obs1}$ $\#_{\rm mean1}$ $\#_{\rm sd1}$ $\#_{\rm obs2}$ $\#_{\rm mean2}$ $\#_{\rm sd2}$ [, `level(`#`)`]

onesampleopts	Description
Main	
`sd(`#`)`	one-population standard deviation; default is `sd(1)`
`level(`#`)`	confidence level; default is `level(95)`
`cluster(`*varname*`)`	variable defining the clusters
`rho(`#`)`	intraclass correlation

twosamplegropts	Description
Main	
* by(*groupvar*)	variable defining the groups
unpaired	unpaired test; implied when by() is specified
sd(#)	two-population common standard deviation; default is sd(1)
sd1(#)	standard deviation of the first population; requires sd2() and may not be combined with sd()
sd2(#)	standard deviation of the second population; requires sd1() and may not be combined with sd()
level(#)	confidence level; default is level(95)
cluster(*varname*)	variable defining the clusters
rho(#)	common intraclass correlation
rho1(#)	intraclass correlation for group 1
rho2(#)	intraclass correlation for group 2

*by(*groupvar*) is required.

twosamplevaropts	Description
Main	
* unpaired	unpaired test
sd(#)	two-population common standard deviation; default is sd(1)
sd1(#)	standard deviation of the first population; requires sd2() and may not be combined with sd()
sd2(#)	standard deviation of the second population; requires sd1() and may not be combined with sd()
level(#)	confidence level; default is level(95)

*unpaired is required.

pairedopts	Description
Main	
* corr(#)	correlation between paired observations
sd(#)	two-population common standard deviation; default is sd(1); may not be combined with sd1(), sd2(), or sddiff()
sd1(#)	standard deviation of the first population; requires corr() and sd2() and may not be combined with sd() or sddiff()
sd2(#)	standard deviation of the second population; requires corr() and sd1() and may not be combined with sd() or sddiff()
level(#)	confidence level; default is level(95)

*corr(#) is required.

by and collect are allowed with ztest and ztesti; see **[U] 11.1.10 Prefix commands**.

Options

Main

by(*groupvar*) specifies the *groupvar* that defines the two groups that ztest will use to test the hypothesis that their means are equal. Specifying by(*groupvar*) implies an unpaired (two-sample) z test. Do not confuse the by() option with the by prefix; you can specify both.

unpaired specifies that the data be treated as unpaired. The unpaired option is used when the two sets of values to be compared are in different variables.

sddiff(*#*) specifies the population standard deviation of the differences between paired observations for a paired z test. For this kind of test, either sddiff() or corr() must be specified.

corr(*#*) specifies the correlation between paired observations for a paired z test. This option along with sd1() and sd2() or with sd() is used to compute the standard deviation of the differences between paired observations unless that standard deviation is supplied directly in the sddiff() option. For a paired z test, either sddiff() or corr() must be specified.

sd(*#*) specifies the population standard deviation for a one-sample z test or the common population standard deviation for a two-sample z test. The default is sd(1). sd() may not be combined with sd1(), sd2(), or sddiff().

sd1(*#*) specifies the standard deviation of the first population or group. When sd1() is specified with by(*groupvar*), the first group is defined by the first category of the sorted *groupvar*. sd1() requires sd2() and may not be combined with sd() or sddiff().

sd2(*#*) specifies the standard deviation of the second population or group. When sd2() is specified with by(*groupvar*), the second group is defined by the second category of the sorted *groupvar*. sd2() requires sd1() and may not be combined with sd() or sddiff().

level(*#*) specifies the confidence level, as a percentage, for confidence intervals. The default is level(95) or as set by set level; see **[U] 20.8 Specifying the width of confidence intervals**.

cluster(*varname*) specifies the variable that identifies clusters. The cluster() option is required to adjust the computation for clustering.

rho(*#*) specifies the intraclass correlation for a one-sample test or the common intraclass correlation for a two-sample test. The rho() option is required to adjust the computation for clustering for a one-sample test.

rho1(*#*) specifies the intraclass correlation of the first group for a two-sample test using groups. The rho() option or both rho1() and rho2() options are required to adjust the computation for clustering.

rho2(*#*) specifies the intraclass correlation of the second group for a two-sample test using groups. The rho() option or both rho1() and rho2() options are required to adjust the computation for clustering.

When by() is used, sd1() and sd2() or sd() is used to specify the population standard deviations of the two groups defined by *groupvar* for an unpaired two-sample z test (using groups). By default, a common standard deviation of one, sd(1), is assumed.

When unpaired is used, sd1() and sd2() or sd() is used to specify the population standard deviations of $varname_1$ and $varname_2$ for an unpaired two-sample z test (using variables). By default, a common standard deviation of one, sd(1), is assumed.

Options corr(), sd1(), and sd2() or corr() and sd() are used for a paired z test to compute the standard deviation of the differences between paired observations. By default, a common standard deviation of one, `sd(1)`, is assumed for both populations. Alternatively, the standard deviation of the differences between paired observations may be supplied directly with the `sddiff()` option.

Remarks and examples

stata.com

Remarks are presented under the following headings:

One-sample z test
Two-sample z test
Paired z test
Adjust for clustering
Immediate form

For the purpose of illustration, we assume that variances are known in all the examples below.

One-sample z test

▷ Example 1

In the first form, `ztest` tests whether the mean of the sample is equal to a known constant under the assumption of known variance. Assume that we have a sample of 74 automobiles. We know each automobile's average mileage rating and wish to test whether the overall average for the sample is 20 miles per gallon. We also assume that the population standard deviation is 6.

```
. use https://www.stata-press.com/data/r18/auto
(1978 automobile data)

. ztest mpg==20, sd(6)

One-sample z test
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. err.   Std. dev.   [95% conf. interval]
---------+--------------------------------------------------------------------
     mpg |      74     21.2973    .6974858           6    19.93025    22.66434
------------------------------------------------------------------------------
    mean = mean(mpg)                                              z =   1.8600
H0: mean = 20

    Ha: mean < 20              Ha: mean != 20               Ha: mean > 20
 Pr(Z < z) = 0.9686       Pr(|Z| > |z|) = 0.0629         Pr(Z > z) = 0.0314
```

The p-value for the two-sided test is 0.0629, so we do not have statistical evidence to reject the null hypothesis that the mean equals 20 at a 5% significance level, but we would reject the null hypothesis at a 10% level.

◁

Two-sample z test

▷ Example 2: Two-sample z test using groups

We are testing the effectiveness of a new fuel additive. We run an experiment in which 12 cars are given the fuel treatment and 12 cars are not. The results of the experiment are as follows:

treated	mpg
0	20
0	23
0	21
0	25
0	18
0	17
0	18
0	24
0	20
0	24
0	23
0	19
1	24
1	25
1	21
1	22
1	23
1	18
1	17
1	28
1	24
1	27
1	21
1	23

The treated variable is coded as 1 if the car received the fuel treatment and 0 otherwise.

We can test the equality of means of the treated and untreated group by typing

```
. use https://www.stata-press.com/data/r18/fuel3
. ztest mpg, by(treated) sd(3)
Two-sample z test
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. err.   Std. dev.   [95% conf. interval]
---------+--------------------------------------------------------------------
       0 |      12          21    .8660254           3    19.30262    22.69738
       1 |      12       22.75    .8660254           3    21.05262    24.44738
---------+--------------------------------------------------------------------
    diff |               -1.75    1.224745               -4.150456    .6504558
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      z =  -1.4289
H0: diff = 0

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(Z < z) = 0.0765         Pr(|Z| > |z|) = 0.1530         Pr(Z > z) = 0.9235
```

We do not have evidence to reject the null hypothesis that the means of the two groups are equal at a 5% significance level.

In the above, we assumed that the two groups have the same standard deviation of 3. If the standard deviations for the two groups are different, we can specify group-specific standard deviations in options `sd1()` and `sd2()`:

```
. ztest mpg, by(treated) sd1(2.7) sd2(3.2)
Two-sample z test
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. err.   Std. dev.   [95% conf. interval]
---------+--------------------------------------------------------------------
       0 |      12          21    .7794229         2.7    19.47236    22.52764
       1 |      12       22.75    .9237604         3.2    20.93946    24.56054
---------+--------------------------------------------------------------------
    diff |               -1.75    1.208649                -4.118909    .6189093
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      z =  -1.4479
H0: diff = 0

    Ha: diff < 0               Ha: diff != 0                 Ha: diff > 0
 Pr(Z < z) = 0.0738         Pr(|Z| > |z|) = 0.1476          Pr(Z > z) = 0.9262
```

◁

❑ Technical note

In two-sample randomized designs, subjects will sometimes refuse the assigned treatment but still be measured for an outcome. In this case, take care to specify the group properly. You might be tempted to let *varname* contain missing where the subject refused and thus let `ztest` drop such observations from the analysis. Zelen (1979) argues that it would be better to specify that the subject belongs to the group in which he or she was randomized, even though such inclusion will dilute the measured effect.

❑

▷ Example 3: Two-sample z test using variables

There is a second, inferior way to organize the data in the preceding example. We ran a test on 24 cars, 12 without the additive and 12 with. We now create two new variables, `mpg1` and `mpg2`.

mpg1	mpg2
20	24
23	25
21	21
25	22
18	23
17	18
18	17
24	28
20	24
24	27
23	21
19	23

This method is inferior because it suggests a connection that is not there. There is no link between the car with 20 mpg and the car with 24 mpg in the first row of the data. Each column of data could be arranged in any order. Nevertheless, if our data are organized like this, `ztest` can accommodate us.

```
. use https://www.stata-press.com/data/r18/fuel
. ztest mpg1==mpg2, unpaired sd(3)
Two-sample z test
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. err.   Std. dev.   [95% conf. interval]
---------+--------------------------------------------------------------------
    mpg1 |      12          21    .8660254           3    19.30262    22.69738
    mpg2 |      12       22.75    .8660254           3    21.05262    24.44738
---------+--------------------------------------------------------------------
    diff |               -1.75    1.224745               -4.150456    .6504558
------------------------------------------------------------------------------
    diff = mean(mpg1) - mean(mpg2)                                z =  -1.4289
H0: diff = 0

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(Z < z) = 0.0765         Pr(|Z| > |z|) = 0.1530          Pr(Z > z) = 0.9235
```

◁

Paired z test

▷ Example 4

Suppose that the preceding data were actually collected by running a test on 12 cars. Each car was run once with the fuel additive and once without. Our data are stored in the same manner as in example 3, but this time, there is most certainly a connection between the mpg values that appear in the same row. These come from the same car. The variables `mpg1` and `mpg2` represent mileage without and with the treatment, respectively. Suppose that the two variables have a common standard deviation of 2 and the correlation between them is 0.4.

```
. use https://www.stata-press.com/data/r18/fuel
. ztest mpg1==mpg2, sd(2) corr(0.4)
Paired z test
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. err.   Std. dev.   [95% conf. interval]
---------+--------------------------------------------------------------------
    mpg1 |      12          21    .5773503           2    19.86841    22.13159
    mpg2 |      12       22.75    .5773503           2    21.61841    23.88159
---------+--------------------------------------------------------------------
    diff |      12       -1.75    .6324555     2.19089    -2.98959   -.5104099
------------------------------------------------------------------------------
     mean(diff) = mean(mpg1 - mpg2)                               z =  -2.7670
 H0: mean(diff) = 0

 Ha: mean(diff) < 0           Ha: mean(diff) != 0           Ha: mean(diff) > 0
 Pr(Z < z) = 0.0028         Pr(|Z| > |z|) = 0.0057          Pr(Z > z) = 0.9972
```

The p-value for the two-sided test is 0.0057, so we reject, for example, the null hypothesis that the two means are equal at a 5% significance level.

Equivalently, we could specify directly the standard deviation of the differences between paired observations with the `sddiff()` option:

```
. ztest mpg1==mpg2, sddiff(2.191)
Paired z test
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. err.   Std. dev.   [95% conf. interval]
---------+--------------------------------------------------------------------
    diff |      12       -1.75    .6324872       2.191   -2.989652   -.5103478
------------------------------------------------------------------------------
     mean(diff) = mean(mpg1 - mpg2)                               z =  -2.7669
 H0: mean(diff) = 0

 Ha: mean(diff) < 0           Ha: mean(diff) != 0           Ha: mean(diff) > 0
 Pr(Z < z) = 0.0028         Pr(|Z| > |z|) = 0.0057          Pr(Z > z) = 0.9972
```

◁

Adjust for clustering

When observations are not independent and can be grouped into clusters, we need to adjust for clustering in a z test. For example, in a cluster randomized design, groups of individuals are randomized instead of individuals. To adjust for clustering, we need to specify the cluster identifier variable in the `cluster()` option. In the case of a one-sample z test, we need to also specify the intraclass correlation in the `rho()` option. In the case of a two-sample z test, we need to also specify the common population intraclass correlation in the `rho()` option or group-specific population intraclass correlations in the `rho1()` and `rho2()` options.

▷ Example 5: One-sample z test, adjusting for clusters

Consider data on the SAT score of 75 students from 15 classes, with 5 students in each class. We want to test whether the mean verbal SAT score is different from 600. We assume a known standard deviation of 132 and a known intraclass correlation of 0.7. To perform the test, we specify the options `cluster(class)`, `rho(0.7)`, and `sd(132)`:

```
. use https://www.stata-press.com/data/r18/sat
(Fictional SAT data)
. ztest score == 600, cluster(class) rho(0.7) sd(132)
One-sample z test                               Number of clusters =        15
Cluster variable: class                         Cluster size       =         5
                                                Intraclass corr.   =    0.7000
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. err.   Std. dev.   [95% conf. interval]
---------+--------------------------------------------------------------------
   score |      75       504.8    29.71222         132    446.5651    563.0349
------------------------------------------------------------------------------
    mean = mean(score)                                            z =  -3.2041
H0: mean = 600

   Ha: mean < 600               Ha: mean != 600               Ha: mean > 600
 Pr(Z < z) = 0.0007         Pr(|Z| > |z|) = 0.0014          Pr(Z > z) = 0.9993
```

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 $= 0.0014 < 0.05$.

◁

Example 6: Two-sample z test using groups, adjusting for clusters

Consider a cluster randomized control trial that studies the effect of additional training of nurses and general practitioners in patient-centered care on the well-being and future disease risk of patients with type 2 diabetes (Kinmonth et al. [1998] and Campbell and Walters [2014]). Practices (`practice`) are randomly allocated to two groups—one trained to give patient-centered care (intervention group) and another trained to give routine care (comparison or control group). In our analysis, we transform the original `bmi` using the formula $\ln(\texttt{bmi} - 14.67355)$ to obtain a variable that is approximately normally distributed, `lbmi`. We want to test the equality of the means of `lbmi` for the two groups. We assume a known common standard deviation of 0.35 and a known common intraclass correlation of 0.028.

To perform the test, we need to specify the `rho(0.028)` and `sd(0.35)` options. We also need to specify the cluster identifier `practice` in the `cluster()` option and the group identifier `group` in the `by()` option.

```
. use https://www.stata-press.com/data/r18/dcfd_trial
(BMI data from Diabetes Care from Diagnosis trial (Kinmonth et al., 1998))

. ztest lbmi, by(group) cluster(practice) rho(0.028) sd(0.35)

Two-sample z test
Cluster variable: practice

Group: Control                              Group: Interv.
  Number of clusters =        20              Number of clusters =        18
  Avg. cluster size  =      5.10              Avg. cluster size  =      7.67
  CV cluster size    =    0.5330              CV cluster size    =    0.5126
  Intraclass corr.   =    0.0280              Intraclass corr.   =    0.0280

------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. err.   Std. dev.   [95% conf. interval]
---------+--------------------------------------------------------------------
 Control |     102     2.62954    .0372502         .35    2.556531    2.702549
  Interv. |    138    2.749023    .0332182         .35    2.683916     2.81413
---------+--------------------------------------------------------------------
    diff |           -.1194831    .0499102                -.2173054   -.0216608
------------------------------------------------------------------------------
    diff = mean(Control) - mean(Interv.)                          z =  -2.3940
H0: diff = 0

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(Z < z) = 0.0083         Pr(|Z| > |z|) = 0.0167          Pr(Z > z) = 0.9917
```

We find statistical evidence to reject the null hypothesis of H_0: $\mu_{\rm diff} = 0$ versus a two-sided alternative H_a: $\mu_{\rm diff} \neq 0$ at the 5% significance level; the p-value $= 0.0167 < 0.05$.

Immediate form

Example 7: One-sample z test

`ztesti` is like `ztest`, except that we specify summary statistics rather than variables as arguments. For instance, we are reading an article that reports the mean number of sunspots per month as 62.6 with a standard deviation of 15.8. We assume this standard deviation is the population standard deviation. There are 24 months of data. We wish to test whether the mean is 75:

```
. ztesti 24 62.6 15.8 75
One-sample z test
------------------------------------------------------------------------------
         |     Obs        Mean    Std. err.   Std. dev.   [95% conf. interval]
---------+--------------------------------------------------------------------
       x |      24        62.6    3.225161        15.8     56.2788     68.9212
------------------------------------------------------------------------------
    mean = mean(x)                                                z =  -3.8448
H0: mean = 75

    Ha: mean < 75               Ha: mean != 75               Ha: mean > 75
 Pr(Z < z) = 0.0001         Pr(|Z| > |z|) = 0.0001        Pr(Z > z) = 0.9999
```

Example 8: Two-sample z test

There is no immediate form of `ztest` with paired data because the test is also a function of the covariance, a number unlikely to be reported in any published source. For unpaired data, however, we might type

```
. ztesti 20 20 5 32 15 4
Two-sample z test
------------------------------------------------------------------------------
         |     Obs        Mean    Std. err.   Std. dev.   [95% conf. interval]
---------+--------------------------------------------------------------------
       x |      20          20    1.118034           5    17.80869    22.19131
       y |      32          15    .7071068           4     13.6141     16.3859
---------+--------------------------------------------------------------------
    diff |                   5    1.322876                2.407211    7.592789
------------------------------------------------------------------------------
    diff = mean(x) - mean(y)                                      z =   3.7796
H0: diff = 0

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(Z < z) = 0.9999         Pr(|Z| > |z|) = 0.0002        Pr(Z > z) = 0.0001
```

Stored results

One-sample `ztest` and `ztesti` store the following in `r()`:

Scalars

`r(N)`	sample size
`r(mu)`	sample mean
`r(sd)`	standard deviation
`r(se)`	standard error
`r(lb)`	lower confidence bound of one-sample mean
`r(ub)`	upper confidence bound of one-sample mean
`r(z)`	z statistic
`r(p_l)`	lower one-sided p-value
`r(p)`	two-sided p-value
`r(p_u)`	upper one-sided p-value
`r(level)`	confidence level

Cluster-adjusted one-sample `ztest` also stores the following in `r()`:

Scalars

`r(K)`	number of clusters K
`r(M)`	cluster size M
`r(rho)`	intraclass correlation
`r(CV_cluster)`	coefficient of variation for cluster sizes

Two-sample `ztest` and `ztesti` store the following in `r()`:

Scalars

`r(N1)`	sample size of population one
`r(N2)`	sample size of population two
`r(mu1)`	sample mean for population one
`r(mu2)`	sample mean for population two
`r(mu_diff)`	difference of means
`r(corr)`	correlation between paired observations; if the `corr()` option is specified
`r(sd)`	common standard deviation
`r(sd1)`	standard deviation for population one
`r(sd2)`	standard deviation for population two
`r(sd_diff)`	standard deviation of the differences between paired observations
`r(se1)`	standard error of population-one sample mean
`r(se2)`	standard error of population-two sample mean
`r(se_diff)`	standard error of the difference of means
`r(lb1)`	lower confidence bound of population-one sample mean
`r(ub1)`	upper confidence bound of population-one sample mean
`r(lb2)`	lower confidence bound of population-two sample mean
`r(ub2)`	upper confidence bound of population-two sample mean
`r(lb_diff)`	lower confidence bound of the difference of means
`r(ub_diff)`	upper confidence bound of the difference of means
`r(z)`	z statistic
`r(p_l)`	lower one-sided p-value
`r(p)`	two-sided p-value
`r(p_u)`	upper one-sided p-value
`r(level)`	confidence level

Cluster-adjusted two-sample `ztest` using the `by()` option also stores the following in `r()`:

Scalars

`r(K1)`	population-one number of clusters K_1
`r(K2)`	population-two number of clusters K_2
`r(M1)`	population-one cluster size M_1
`r(M2)`	population-two cluster size M_2
`r(rho)`	common intraclass correlation
`r(rho1)`	population-one intraclass correlation
`r(rho2)`	population-two intraclass correlation
`r(CV_cluster1)`	population-one coefficient of variation for cluster sizes
`r(CV_cluster2)`	population-two coefficient of variation for cluster sizes

Methods and formulas

Methods and formulas are presented under the following headings:

One-sample z test
Two-sample unpaired z test
Paired z test

For all the tests below, the test statistic z is distributed as standard normal, and the p-value is computed as

$$p = \begin{cases} 1 - \Phi(z) & \text{for an upper one-sided test} \\ \Phi(z) & \text{for a lower one-sided test} \\ 2\{1 - \Phi(|z|)\} & \text{for a two-sided test} \end{cases}$$

where $\Phi(\cdot)$ is the cdf of a standard normal distribution and $|z|$ is an absolute value of z.

Also see, for instance, Hoel (1984, 140–161), Dixon and Massey (1983, 100–130), and Tamhane and Dunlop (2000, 237–290) for more information about z tests.

One-sample z test

Suppose that we observe a random sample $x_1, x_2, \dots, x_n$ of size n, which follows a normal distribution with mean μ and standard deviation σ. We are interested in 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: $\mu < \mu_0$. Assuming a known standard deviation σ, we use the following test statistic,

$$z = \frac{(\overline{x} - \mu_0)}{s}$$

where $\overline{x} = (\sum_{i=1}^{n} x_i)/n$ is the sample mean and $s = \sigma/\sqrt{n}$ is the standard error of $\overline{x}$.

The $100(1-\alpha)\%$ confidence interval for $\overline{x}$ is given by

$$\overline{x} \pm z_{1-\alpha/2}s$$

where $z_{1-\alpha/2}$ is the $(1-\alpha/2)$th quantile of the standard normal distribution.

With clustered data, suppose that there are K clusters. The ith cluster of size M_i contains the observations $x_{i1}, x_{i2}, \dots, x_{iM_i}$, such that $n = \sum_{i=1}^{K} M_i$ and $\overline{x} = \frac{1}{n}\sum_{i=1}^{K}\sum_{j=1}^{M_i} x_{ij}$. Let ρ be the intraclass correlation. Following Ahn, Heo, and Zhang (2015), we assume that the cluster sizes M_i are independent and identically distributed. Let C_{adj} be the adjustment to the standard error for clustered data,

$$C_{\text{adj}} = \sqrt{\sum_{i=1}^{K} M_i\{1 + \rho(M_i - 1)\}/n}$$

such that $s_{\text{cl}} = C_{\text{adj}}s$.

C_{adj} can be equivalently written as

$$C_{\text{adj}} = \sqrt{1 + \rho(\overline{M} - 1) + \rho\overline{M}\text{CV}_{\text{cl}}^2}$$

where $\overline{M} = \sum_{i=1}^{K} M_i/K$ is the average cluster size and CV_{cl} is the coefficient of variation for cluster sizes:

$$\text{CV}_{\text{cl}} = \frac{\sqrt{\sum_{i=1}^{K}(M_i - \overline{M})^2/K}}{\overline{M}}$$

To adjust the test statistic z and the confidence interval for clustering, replace s with $s_{\rm cl}$ in the corresponding formulas. In the presence of clustering, the test statistic z is asymptotically normally distributed conditional on the empirical distribution of M_i's.

Two-sample unpaired z test

Suppose that we observe a random sample $x_{11}, x_{12}, \ldots, x_{1n_1}$ of size n_1, which follows a normal distribution with mean μ_1 and standard deviation σ_1, and another random sample $x_{21}, x_{22}, \ldots, x_{2n_2}$ of size n_2, which follows a normal distribution with mean μ_2 and standard deviation σ_2. We are interested in testing the null hypothesis $H_0\colon \mu_2 = \mu_1$ versus the two-sided alternative hypothesis $H_a\colon \mu_2 \neq \mu_1$, the upper one-sided alternative $H_a\colon \mu_2 > \mu_1$, or the lower one-sided alternative $H_a\colon \mu_2 < \mu_1$. Assuming known standard deviations σ_1 and σ_2, we use the following test statistic,

$$z = \frac{\overline{x}_2 - \overline{x}_1}{\sqrt{s_1^2 + s_2^2}}$$

where $\overline{x}_1 = (\sum_{i=1}^{n_1} x_{1i})/n_1$ and $\overline{x}_2 = (\sum_{i=1}^{n_2} x_{2i})/n_2$ are the two sample means and $s_1 = \sigma_1/\sqrt{n_1}$ and $s_2 = \sigma_2/\sqrt{n_2}$ are the corresponding two standard errors.

The $100(1-\alpha)\%$ confidence intervals for $\overline{x}_1$ and $\overline{x}_2$ are given by

$$\overline{x}_1 \pm z_{1-\alpha/2} s_1$$
$$\overline{x}_2 \pm z_{1-\alpha/2} s_2$$

where $z_{1-\alpha/2}$ is the $(1-\alpha/2)$th quantile of the standard normal distribution.

The $100(1-\alpha)\%$ confidence interval for $\overline{x}_1 - \overline{x}_2$ is given by

$$\overline{x}_1 - \overline{x}_2 \pm z_{1-\alpha/2}\sqrt{s_1^2 + s_2^2}$$

With clustered data, similar to the discussion for the one-sample test, suppose that population one has K_1 clusters and population two has K_2 clusters. Let ρ_1 and ρ_2 be the intraclass correlations, $\overline{M}_1$ and $\overline{M}_2$ be the average cluster sizes, $\overline{x}_1 = (1/n_1)\sum_{i=1}^{K_1}\sum_{j=1}^{M_{1i}} x_{1ij}$ and $\overline{x}_2 = (1/n_2)\sum_{i=1}^{K_2}\sum_{j=1}^{M_{2i}} x_{2ij}$ be the sample means, and $\mathrm{CV}_{\rm cl,1}$ and $\mathrm{CV}_{\rm cl,2}$ be the coefficients of variation for cluster sizes for population one and population two. Let $s_{1,\rm cl} = C_{\rm adj,1} s_1$ and $s_{2,\rm cl} = C_{\rm adj,2} s_2$ be the standard errors of the population-specific sample means adjusted for clustered data, where the population-specific adjustment factors are defined as described for the one-sample test. To adjust the two-sample test statistic and the confidence intervals for clustering, replace s_1 with $s_{1,\rm cl}$ and s_2 with $s_{2,\rm cl}$ in the corresponding formulas.

Paired z test

Some experiments have paired observations (also known as matched observations, correlated pairs, or permanent components). Consider a sequence of n paired observations denoted by x_{ij} for subjects $i = 1, 2, \ldots, n$ and groups $j = 1, 2$. An individual observation corresponds 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 let $D_i = x_{i2} - x_{i1}$ denote the difference between individual observations. D_i follows a normal distribution with mean $\mu_2 - \mu_1$ and standard deviation σ_d, where $\sigma_d = \sqrt{{\sigma_1}^2 + {\sigma_2}^2 - 2\rho_{\rm pair}\sigma_1\sigma_2}$, σ_j is the population standard deviation of group j and $\rho_{\rm pair}$ is the correlation between paired observations.

We are interested in testing the null hypothesis H_0: $\mu_2 = \mu_1$ versus the two-sided alternative hypothesis H_a: $\mu_2 \neq \mu_1$, the upper one-sided alternative H_a: $\mu_2 > \mu_1$, or the lower one-sided alternative H_a: $\mu_2 < \mu_1$. Assuming the standard deviation of the differences σ_d is known, we use the following test statistic,

$$z = \frac{\overline{d}}{s_d}$$

where $\overline{d} = (\sum_{i=1}^{n} D_i)/n$ is the sample mean of the differences between paired observations and $s_d = \sigma_d/\sqrt{n}$ is the standard error of $\overline{d}$.

The $100(1-\alpha)\%$ confidence interval for $\overline{d}$ is given by

$$\overline{d} \pm z_{1-\alpha/2} s_d$$

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.

Dixon, W. J., and F. J. Massey, Jr. 1983. *Introduction to Statistical Analysis*. 4th ed. New York: McGraw–Hill.

Hoel, P. G. 1984. *Introduction to Mathematical Statistics*. 5th ed. New York: Wiley.

Kinmonth, A. L., A. Woodcock, S. Griffin, N. Spiegal, and M. J. Campbell. 1998. Randomised controlled trial of patient centred care of diabetes in general practice: Impact on current wellbeing and future disease risk. *BMJ* 317: 1202–1208. https://doi.org/10.1136/bmj.317.7167.1202.

Tamhane, A. C., and D. D. Dunlop. 2000. *Statistics and Data Analysis: From Elementary to Intermediate*. Upper Saddle River, NJ: Prentice Hall.

Zelen, M. 1979. A new design for randomized clinical trials. *New England Journal of Medicine* 300: 1242–1245. https://doi.org/10.1056/NEJM197905313002203.

Also see

[R] **ci** — Confidence intervals for means, proportions, and variances

[R] **esize** — Effect size based on mean comparison

[R] **mean** — Estimate means

[R] **oneway** — One-way analysis of variance

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

[MV] **hotelling** — Hotelling's T^2 generalized means test

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

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

[PSS-2] **power pairedmeans** — Power analysis for a two-sample paired-means test

[PSS-2] **power twomeans** — Power analysis for a two-sample means test

[PSS-2] **power twomeans, cluster** — Power analysis for a two-sample means test, CRD

For suggested citations, see the FAQ on citing Stata documentation.