ttest — t tests (mean-comparison tests)

Description

`ttest` performs t tests on the equality of means. The test can be performed for one sample against a hypothesized population mean. Two-sample tests can be conducted for paired and unpaired data. The assumption of equal variances can be optionally relaxed in the unpaired two-sample case.

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

Quick start

Test that the mean of `v1` is equal between two groups defined by `catvar`

```
ttest v1, by(catvar)
```

Same as above, but assume unequal variances

```
ttest v1, by(catvar) unequal
```

Paired t test of `v2` and `v3`

```
ttest v2 == v3
```

Same as above, but with unpaired data and conduct test separately for each level of `catvar`

```
by catvar: ttest v2 == v3, unpaired
```

Test that the mean of `v4` is 3 at the 90% confidence level

```
ttest v4 == 3, level(90)
```

Test $\mu_1 = \mu_2$ if $\overline{x}_1 = 3.2$, $sd_1 = 0.1$, $\overline{x}_2 = 3.4$, and $sd_2 = 0.15$ with $n_1 = n_2 = 12$

```
ttesti 12 3.2 .1 12 3.4 .15
```

Menu

ttest

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

ttesti

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

Syntax

One-sample t test

ttest *varname* == # [*if*] [*in*] [, level(#)]

Two-sample t test using groups

ttest *varname* [*if*] [*in*] , by(*groupvar*) [*options*$_1$]

Two-sample t test using variables

ttest *varname*$_1$ == *varname*$_2$ [*if*] [*in*] , unpaired [unequal welch level(#)]

Paired t test

ttest *varname*$_1$ == *varname*$_2$ [*if*] [*in*] [, level(#)]

Immediate form of one-sample t test

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

Immediate form of two-sample t test

ttesti $\#_{obs1}$ $\#_{mean1}$ $\#_{sd1}$ $\#_{obs2}$ $\#_{mean2}$ $\#_{sd2}$ [, *options*$_2$]

options$_1$	Description
Main	
* by(*groupvar*)	variable defining the groups
reverse	reverse group order for mean difference computation
unequal	unpaired data have unequal variances
welch	use Welch's approximation
level(#)	set confidence level; default is level(95)

*by(*groupvar*) is required.

options$_2$	Description
Main	
unequal	unpaired data have unequal variances
welch	use Welch's approximation
level(#)	set confidence level; default is level(95)

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

Options

Main

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

`reverse` reverses the order of the mean difference between groups defined in `by()`. By default, the mean of the group corresponding to the largest value in the variable in `by()` is subtracted from the mean of the group with the smallest value in `by()`. `reverse` reverses this behavior and the order in which variables appear on the table.

`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.

`unequal` specifies that the unpaired data not be assumed to have equal variances.

`welch` specifies that the approximate degrees of freedom for the test be obtained from Welch's formula (1947) rather than from Satterthwaite's approximation formula (1946), which is the default when `unequal` is specified. Specifying `welch` implies `unequal`.

`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**.

Remarks and examples

Remarks are presented under the following headings:

One-sample t test

▷ Example 1

In the first form, `ttest` tests whether the mean of the sample is equal to a known constant under the assumption of unknown 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.

```
. use https://www.stata-press.com/data/r19/auto
(1978 automobile data)

. ttest mpg==20

One-sample t test
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. err.   Std. dev.   [95% conf. interval]
---------+--------------------------------------------------------------------
     mpg |      74     21.2973    .6725511    5.785503     19.9569    22.63769
------------------------------------------------------------------------------
    mean = mean(mpg)                                              t =   1.9289
H0: mean = 20                                    Degrees of freedom =       73

    Ha: mean < 20               Ha: mean != 20                 Ha: mean > 20
 Pr(T < t) = 0.9712         Pr(|T| > |t|) = 0.0576          Pr(T > t) = 0.0288
```

The test indicates that the underlying mean is not 20 with a significance level of 5.8%.

◁

Two-sample t test

▷ Example 2: Two-sample t 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/r19/fuel3
. ttest mpg, by(treated)
Two-sample t test with equal variances
```

Group	Obs	Mean	Std. err.	Std. dev.	[95% conf.	interval]
0	12	21	.7881701	2.730301	19.26525	22.73475
1	12	22.75	.9384465	3.250874	20.68449	24.81551
Combined	24	21.875	.6264476	3.068954	20.57909	23.17091
diff		-1.75	1.225518		-4.291568	.7915684

```
    diff = mean(0) - mean(1)                                      t =  -1.4280
H0: diff = 0                                     Degrees of freedom =       22

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0837         Pr(|T| > |t|) = 0.1673          Pr(T > t) = 0.9163
```

We do not find a statistically significant difference in the means.

If we were not willing to assume that the variances were equal and wanted to use Welch's formula, we could type

```
. ttest mpg, by(treated) welch
Two-sample t test with unequal variances
```

Group	Obs	Mean	Std. err.	Std. dev.	[95% conf.	interval]
0	12	21	.7881701	2.730301	19.26525	22.73475
1	12	22.75	.9384465	3.250874	20.68449	24.81551
Combined	24	21.875	.6264476	3.068954	20.57909	23.17091
diff		-1.75	1.225518		-4.28369	.7836902

```
    diff = mean(0) - mean(1)                                      t =  -1.4280
H0: diff = 0                             Welch's degrees of freedom =  23.2465

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0833         Pr(|T| > |t|) = 0.1666          Pr(T > t) = 0.9167
```

◁

❑ 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 `ttest` 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 t 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, `ttest` can accommodate us.

```
. use https://www.stata-press.com/data/r19/fuel

. ttest mpg1==mpg2, unpaired

Two-sample t test with equal variances
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. err.   Std. dev.   [95% conf. interval]
---------+--------------------------------------------------------------------
    mpg1 |      12          21    .7881701    2.730301    19.26525    22.73475
    mpg2 |      12       22.75    .9384465    3.250874    20.68449    24.81551
---------+--------------------------------------------------------------------
Combined |      24      21.875    .6264476    3.068954    20.57909    23.17091
---------+--------------------------------------------------------------------
    diff |               -1.75    1.225518               -4.291568    .7915684
------------------------------------------------------------------------------
    diff = mean(mpg1) - mean(mpg2)                                t =  -1.4280
H0: diff = 0                                     Degrees of freedom =       22

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0837         Pr(|T| > |t|) = 0.1673          Pr(T > t) = 0.9163
```

◁

Paired t 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.

```
. use https://www.stata-press.com/data/r19/fuel
. ttest mpg1==mpg2
Paired t test
```

Variable	Obs	Mean	Std. err.	Std. dev.	[95% conf.	interval]
mpg1	12	21	.7881701	2.730301	19.26525	22.73475
mpg2	12	22.75	.9384465	3.250874	20.68449	24.81551
diff	12	-1.75	.7797144	2.70101	-3.46614	-.0338602

```
     mean(diff) = mean(mpg1 - mpg2)                          t =  -2.2444
 H0: mean(diff) = 0                         Degrees of freedom =       11

 Ha: mean(diff) < 0           Ha: mean(diff) != 0           Ha: mean(diff) > 0
 Pr(T < t) = 0.0232         Pr(|T| > |t|) = 0.0463          Pr(T > t) = 0.9768
```

We find that the means are statistically different from each other at any level greater than 4.6%.

◁

Two-sample t test compared with one-way ANOVA

▷ Example 5

In example 2, we saw that ttest can be used to test the equality of a pair of means; see [R] **oneway** for an extension that allows testing the equality of more than two means.

Suppose that we have data on the 50 states. The dataset contains the median age of the population (medage) and the region of the country (region) for each state. Region 1 refers to the Northeast, region 2 to the North Central, region 3 to the South, and region 4 to the West. Using oneway, we can test the equality of all four means.

```
. use https://www.stata-press.com/data/r19/census
(1980 Census data by state)
. oneway medage region
```

Analysis of variance

Source	SS	df	MS	F	Prob > F
Between groups	46.3961903	3	15.4653968	7.56	0.0003
Within groups	94.1237947	46	2.04616945		
Total	140.519985	49	2.8677548		

```
Bartlett's equal-variances test: chi2(3) =  10.5757    Prob>chi2 = 0.014
```

We find that the means are different, but we are interested only in testing whether the means for the Northeast (`region==1`) and West (`region==4`) are different. We could use `oneway`:

```
. oneway medage region if region==1 | region==4
                        Analysis of variance
    Source              SS         df      MS            F     Prob > F
------------------------------------------------------------------------
Between groups      46.241247      1    46.241247     20.02     0.0002
 Within groups      46.1969169     20   2.30984584
------------------------------------------------------------------------
    Total           92.4381638     21   4.40181733
Bartlett's equal-variances test: chi2(1) =   2.4679    Prob>chi2 = 0.116
```

We could also use `ttest`:

```
. ttest medage if region==1 | region==4, by(region)
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. err.   Std. dev.   [95% conf. interval]
---------+--------------------------------------------------------------------
      NE |       9    31.23333    .3411581    1.023474    30.44662    32.02005
    West |      13    28.28462    .4923577    1.775221    27.21186    29.35737
---------+--------------------------------------------------------------------
Combined |      22    29.49091    .4473059    2.098051    28.56069    30.42113
---------+--------------------------------------------------------------------
    diff |            2.948718    .6590372                 1.57399    4.323445
------------------------------------------------------------------------------
    diff = mean(NE) - mean(West)                                  t =   4.4743
H0: diff = 0                                     Degrees of freedom =       20

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

The significance levels of both tests are the same.

◁

Immediate form

▷ Example 6

`ttesti` is like `ttest`, 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. There are 24 months of data. We wish to test whether the mean is 75:

```
. ttesti 24 62.6 15.8 75
One-sample t test
------------------------------------------------------------------------------
         |     Obs        Mean    Std. err.   Std. dev.   [95% conf. interval]
---------+--------------------------------------------------------------------
       x |      24        62.6    3.225161        15.8    55.92825    69.27175
------------------------------------------------------------------------------
    mean = mean(x)                                                t =  -3.8448
H0: mean = 75                                    Degrees of freedom =       23

    Ha: mean < 75               Ha: mean != 75                 Ha: mean > 75
 Pr(T < t) = 0.0004         Pr(|T| > |t|) = 0.0008          Pr(T > t) = 0.9996
```

◁

▷ Example 7

There is no immediate form of ttest 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

```
. ttesti 20 20 5  32 15 4
Two-sample t test with equal variances

------------------------------------------------------------------------------
         |     Obs        Mean    Std. err.   Std. dev.   [95% conf. interval]
---------+--------------------------------------------------------------------
       x |      20          20    1.118034           5    17.65993    22.34007
       y |      32          15    .7071068           4    13.55785    16.44215
---------+--------------------------------------------------------------------
Combined |      52    16.92308    .6943785    5.007235    15.52905     18.3171
---------+--------------------------------------------------------------------
    diff |                   5    1.256135                2.476979    7.523021
------------------------------------------------------------------------------
    diff = mean(x) - mean(y)                                      t =   3.9805
H0: diff = 0                                     Degrees of freedom =       50

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

If we had typed ttesti 20 20 5 32 15 4, unequal, the test would have assumed unequal variances.

◁

Video examples

One-sample t test in Stata

t test for two independent samples in Stata

t test for two paired samples in Stata

One-sample t-test calculator

Two-sample t-test calculator

Stored results

One-sample ttest and ttesti store the following in r():

Scalars

r(N_1)	sample size
r(mu_1)	sample mean
r(sd_1)	standard deviation
r(se)	standard error
r(lb_1)	lower confidence bound for the mean
r(ub_1)	upper confidence bound for the mean
r(t)	t statistic
r(df_t)	degrees of freedom
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

Two-sample ttest and ttesti store the following in r():

Scalars

r(N_1)	sample size n_1
r(N_2)	sample size n_2
r(N_combined)	combined sample size
r(mu_1)	estimated mean for sample 1
r(mu_2)	estimated mean for sample 2
r(mu_combined)	combined sample mean
r(mu_diff)	difference of means
r(sd_1)	standard deviation for sample 1
r(sd_2)	standard deviation for sample 2
r(sd)	combined standard deviation
r(se)	standard error
r(se_1)	standard error of the mean for sample 1
r(se_2)	standard error of the mean for sample 2
r(se_combined)	standard error of combined sample mean
r(lb_1)	lower confidence bound for the mean for sample 1
r(ub_1)	upper confidence bound for the mean for sample 1
r(lb_2)	lower confidence bound for the mean for sample 2
r(ub_2)	upper confidence bound for the mean for sample 2
r(lb_combined)	lower confidence bound for combined sample mean
r(ub_combined)	upper confidence bound for combined sample mean
r(lb_diff)	lower confidence bound for the difference of means
r(ub_diff)	upper confidence bound for the difference of means
r(t)	t statistic
r(df_t)	degrees of freedom
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

Paired-sample ttest stores the following in r():

Scalars

r(N_1)	sample size n_1
r(N_2)	sample size n_2
r(mu_1)	estimated mean for first variable
r(mu_2)	estimated mean for second variable
r(mu_diff)	difference of means
r(sd_1)	standard deviation for first variable
r(sd_2)	standard deviation for second variable
r(sd_diff)	standard deviation of the paired differences
r(se)	standard error
r(se_1)	standard error of the mean for the first variable
r(se_2)	standard error of the mean for the second variable
r(lb_1)	lower confidence bound for the mean of the first variable
r(ub_1)	upper confidence bound for the mean of the first variable
r(lb_2)	lower confidence bound for the mean of the second variable
r(ub_2)	upper confidence bound for the mean of the second variable
r(lb_diff)	lower confidence bound for the difference of means
r(ub_diff)	upper confidence bound for the difference of means
r(t)	t statistic
r(df_t)	degrees of freedom
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

Methods and formulas

See, for instance, Hoel (1984, 140–161) or Dixon and Massey (1983, 121–130) for an introduction and explanation of the calculation of these tests. Acock (2023, 165–179) and Hamilton (2013, 145–150) describe t tests using applications in Stata.

The test for $\mu = \mu_0$ for unknown σ is given by

$$t = \frac{(\overline{x} - \mu_0)\sqrt{n}}{s}$$

The statistic is distributed as Student's t with $n - 1$ degrees of freedom (Gosset [Student, pseud.] 1908).

The test for $\mu_x = \mu_y$ when σ_x and σ_y are unknown but $\sigma_x = \sigma_y$ is given by

$$t = \frac{\overline{x} - \overline{y}}{\left\{\frac{(n_x-1)s_x^2 + (n_y-1)s_y^2}{n_x+n_y-2}\right\}^{1/2} \left(\frac{1}{n_x} + \frac{1}{n_y}\right)^{1/2}}$$

The result is distributed as Student's t with $n_x + n_y - 2$ degrees of freedom.

You could perform `ttest` (without the `unequal` option) in a regression setting given that regression assumes a homoskedastic error model. To compare with the `ttest` command, denote the underlying observations on x and y by x_j, $j = 1, \ldots, n_x$, and y_j, $j = 1, \ldots, n_y$. In a regression framework, typing `ttest` without the `unequal` option is equivalent to

1. creating a new variable z_j that represents the stacked observations on x and y (so that $z_j = x_j$ for $j = 1, \ldots, n_x$ and $z_{n_x+j} = y_j$ for $j = 1, \ldots, n_y$)
2. and then estimating the equation $z_j = \beta_0 + \beta_1 d_j + \epsilon_j$, where $d_j = 0$ for $j = 1, \ldots, n_x$ and $d_j = 1$ for $j = n_x + 1, \ldots, n_x + n_y$ (that is, $d_j = 0$ when the z observations represent x, and $d_j = 1$ when the z observations represent y).

The estimated value of β_1, b_1, will equal $\overline{y} - \overline{x}$, and the reported t statistic will be the same t statistic as given by the formula above.

The test for $\mu_x = \mu_y$ when σ_x and σ_y are unknown and $\sigma_x \neq \sigma_y$ is given by

$$t = \frac{\overline{x} - \overline{y}}{\left(s_x^2/n_x + s_y^2/n_y\right)^{1/2}}$$

The result is distributed as Student's t with ν degrees of freedom, where ν is given by (with Satterthwaite's [1946] formula)

$$\frac{\left(s_x^2/n_x + s_y^2/n_y\right)^2}{\frac{\left(s_x^2/n_x\right)^2}{n_x-1} + \frac{\left(s_y^2/n_y\right)^2}{n_y-1}}$$

With Welch's formula (1947), the number of degrees of freedom is given by

$$-2+\frac{\left(s_x^2/n_x+s_y^2/n_y\right)^2}{\frac{\left(s_x^2/n_x\right)^2}{n_x+1}+\frac{\left(s_y^2/n_y\right)^2}{n_y+1}}$$

The test for $\mu_x = \mu_y$ for matched observations (also known as paired observations, correlated pairs, or permanent components) is given by

$$t=\frac{\overline{d}\sqrt{n}}{s_d}$$

where $\overline{d}$ represents the mean of $x_i - y_i$ and s_d represents the standard deviation. The test statistic t is distributed as Student's t with $n-1$ degrees of freedom.

You can also use `ttest` without the `unpaired` option in a regression setting because a paired comparison includes the assumption of constant variance. The `ttest` with an unequal variance assumption does not lend itself to an easy representation in regression settings and is not discussed here. $(x_j - y_j) = \beta_0 + \epsilon_j$.

William Sealy Gosset (1876–1937) was born in Canterbury, England. He studied chemistry and mathematics at Oxford and worked as a chemist with the brewers Guinness in Dublin. Gosset became interested in statistical problems, which he discussed with Karl Pearson and later with Fisher and Neyman. He published several important papers under the pseudonym "Student", and he lent that name to the t test he invented.

Stella Cunliffe (1917–2012) was an advocate for increased understanding of the role of human nature in experiments and methodological rigor in social statistics. She was born in Battersea, England. She was the first person from her local public girls' school to attend college, obtaining a bachelor of science from the London School of Economics. Her first job was with the Danish Bacon Company during World War II, where she was in charge of bacon rations for London. After the war, she moved to Germany and again helped to ration food, this time for refugees.

She then spent a long career in quality control at the Guinness Brewing Company. Cunliffe observed that the weights of rejected casks skewed lighter. Noting that workers had to roll casks that were too light or too heavy uphill to be remade, she had the scales moved to the top of the hill. With workers able to roll rejected casks downhill, the weight of these casks began to follow a normal distribution.

After 25 years at Guinness, Cunliffe joined the British Home Office, where she would go on to become the first woman to serve as director of statistics. During her tenure at the Home Office, she emphasized applying principles of experimental design she had learned at Guinness to the study of such topics as birthrates, recidivism, and criminology. In 1975, she became the first woman to serve as president of the Royal Statistical Society.

References

Acock, A. C. 2023. *A Gentle Introduction to Stata*. Rev. 6th ed. College Station, TX: Stata Press.

Boland, P. J. 2000. "William Sealy Gosset—alias 'Student' 1876–1937". In *Creators of Mathematics: The Irish Connection*, edited by K. Houston, 105–112. Dublin: University College Dublin Press.

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

Earnest, A. 2017. *Essentials of a Successful Biostatistical Collaboration*. Boca Raton, FL: CRC Press.

Gosset, W. S. 1943. *"Student's" Collected Papers*. London: Biometrika Office, University College.

Gosset [Student, pseud.], W. S. 1908. The probable error of a mean. *Biometrika* 6: 1–25. https://doi.org/10.2307/2331554.

Hamilton, L. C. 2013. *Statistics with Stata: Updated for Version 12*. 8th ed. Boston: Brooks/Cole.

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

Huber, C. 2013. Measures of effect size in Stata 13. *The Stata Blog: Not Elsewhere Classified*. https://blog.stata.com/2013/09/05/measures-of-effect-size-in-stata-13/.

Kaplan, D. M. 2019. distcomp: Comparing distributions. *Stata Journal* 19: 832–848.

Mehmetoglu, M., and T. G. Jakobsen. 2022. *Applied Statistics Using Stata: A Guide for the Social Sciences*. 2nd ed. Thousand Oaks, CA: Sage.

Pearson, E. S., R. L. Plackett, and G. A. Barnard. 1990. *'Student': A Statistical Biography of William Sealy Gosset*. Oxford: Oxford University Press.

Preece, D. A. 1982. t is for trouble (and textbooks): A critique of some examples of the paired-samples t-test. *Statistician* 31: 169–195. https://doi.org/10.2307/2987888.

Satterthwaite, F. E. 1946. An approximate distribution of estimates of variance components. *Biometrics Bulletin* 2: 110–114. https://doi.org/10.2307/3002019.

Senn, S. J., and W. Richardson. 1994. The first t-test. *Statistics in Medicine* 13: 785–803. https://doi.org/10.1002/sim.4780130802.

Welch, B. L. 1947. The generalization of 'student's' problem when several different population variances are involved. *Biometrika* 34: 28–35. https://doi.org/10.2307/2332510.

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] **bitest** — Binomial probability test

[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] **prtest** — Tests of proportions

[R] **sdtest** — Variance-comparison tests

[R] **ztest** — z tests (mean-comparison tests, known variance)

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

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