Title

stata.com

pkequiv — Perform bioequivalence tests

Description

`pkequiv` performs bioequivalence testing for two treatments. By default, `pkequiv` calculates a standard confidence interval symmetric about the difference between the two treatment means. `pkequiv` also calculates confidence intervals symmetric about zero and intervals based on Fieller's theorem. Also, `pkequiv` can perform interval hypothesis tests for bioequivalence.

`pkequiv` is one of the pk commands. Please read [R] **pk** before reading this entry.

Quick start

Classic CI for difference in pharmacokinetic outcome `y1` between treatments `v1` given over period `v2` in sequence `v3` with subjects identified by `idvar`

```
pkequiv y1 v1 v2 v3 idvar
```

As above, but calculate an exact CI by Fieller's theorem

```
pkequiv y1 v1 v2 v3 idvar, fieller
```

Schuirmann's two one-sided tests for bioequivalence

```
pkequiv y1 v1 v2 v3 idvar, tost
```

Specify the two treatments, 2 and 3, that are to be tested for equivalence

```
pkequiv y1 v1 v2 v3 idvar, compare(2 3)
```

Menu

Statistics > Epidemiology and related > Other > Bioequivalence tests

Syntax

```
pkequiv outcome treatment period sequence id [if] [in] [, options]
```

options	Description
Options	
`compare(`*string*`)`	compare the two specified values of the treatment variable
`limit(#)`	equivalence limit (between 0.10 and 0.99); default is 0.2
`level(#)`	set confidence level; default is `level(90)`
`fieller`	calculate confidence interval by Fieller's theorem
`symmetric`	calculate symmetric equivalence interval
`anderson`	Anderson and Hauck hypothesis test for bioequivalence
`tost`	two one-sided hypothesis tests for bioequivalence
`noboot`	do not estimate probability that CI lies within confidence limits

Options

Options

`compare(`*string*`)` specifies the two treatments to be tested for equivalence. Sometimes there may be more than two treatments, but the equivalence can be determined only between any two treatments.

`limit(#)` specifies the equivalence limit. The default is 0.2. The equivalence limit can be changed only symmetrically; that is, it is not possible to have a 0.15 lower limit and a 0.2 upper limit in the same test.

`level(#)` specifies the confidence level, as a percentage, for confidence intervals. The default is `level(90)`. This setting is not controlled by the `set level` command.

`fieller` specifies that an equivalence interval based on Fieller's theorem be calculated.

`symmetric` specifies that a symmetric equivalence interval be calculated.

`anderson` specifies that the Anderson and Hauck (1983) hypothesis test for bioequivalence be computed. This option is ignored when calculating equivalence intervals based on Fieller's theorem or when calculating a confidence interval that is symmetric about zero.

`tost` specifies that the two one-sided hypothesis tests for bioequivalence be computed. This option is ignored when calculating equivalence intervals based on Fieller's theorem or when calculating a confidence interval that is symmetric about zero.

`noboot` prevents the estimation of the probability that the confidence interval lies within the confidence limits. If this option is not specified, this probability is estimated by resampling the data.

Remarks and examples

stata.com

`pkequiv` is designed to conduct tests for bioequivalence based on data from a crossover experiment. `pkequiv` requires that the user specify the *outcome*, *treatment*, *period*, *sequence*, and *id* variables. The data must be in the same format as that produced by `pkshape`; see [R] **pkshape**.

Example 1

We have the following data on which we want to conduct a bioequivalence test between `treat` $= A$ and `treat` $= B$.

```
. use http://www.stata-press.com/data/r14/pkdata3
. list, sep(4)
```

	id	sequence	outcome	treat	carry	period
1.	1	1	150.9643	A	0	1
2.	2	1	146.7606	A	0	1
3.	3	1	160.6548	A	0	1
4.	4	1	157.8622	A	0	1
5.	5	1	133.6957	A	0	1
6.	7	1	160.639	A	0	1
7.	8	1	131.2604	A	0	1
8.	9	1	168.5186	A	0	1
9.	10	2	137.0627	B	0	1
10.	12	2	153.4038	B	0	1
11.	13	2	163.4593	B	0	1
12.	14	2	146.0462	B	0	1
13.	15	2	158.1457	B	0	1
14.	18	2	147.1977	B	0	1
15.	19	2	164.9988	B	0	1
16.	20	2	145.3823	B	0	1
17.	1	1	218.5551	B	A	2
18.	2	1	133.3201	B	A	2
19.	3	1	126.0635	B	A	2
20.	4	1	96.17461	B	A	2
21.	5	1	188.9038	B	A	2
22.	7	1	223.6922	B	A	2
23.	8	1	104.0139	B	A	2
24.	9	1	237.8962	B	A	2
25.	10	2	139.7382	A	B	2
26.	12	2	202.3942	A	B	2
27.	13	2	136.7848	A	B	2
28.	14	2	104.5191	A	B	2
29.	15	2	165.8654	A	B	2
30.	18	2	139.235	A	B	2
31.	19	2	166.2391	A	B	2
32.	20	2	158.5146	A	B	2

```
. set seed 1
. pkequiv outcome treat period seq id
      Classic confidence interval for bioequivalence
------------------------------------------------------------------------------
                     [equivalence limits]        [    test limits   ]
------------------------------------------------------------------------------
      difference:   -30.296        30.296        -11.332        26.416
           ratio:       80%          120%         92.519%     117.439%
------------------------------------------------------------------------------
Probability test limits are within equivalence limits =    0.6380
Note: Reference treatment = 1
```

The default output for pkequiv shows a confidence interval for the difference of the means (test limits), the ratio of the means, and the federal equivalence limits. The classic confidence interval can be constructed around the difference between the average measure of effect for the two drugs or around the ratio of the average measure of effect for the two drugs. pkequiv reports both the difference measure and the ratio measure. For these data, U.S. federal government regulations state that the confidence interval for the difference must be entirely contained within the range $[-30.296, 30.296]$ and between 80% and 120% for the ratio. Here the test limits are within the equivalence limits. Although the test limits are inside the equivalence limits, there is only a 64% assurance that the observed confidence interval will be within the equivalence limits in the long run. This is an interesting case because, although this sample shows bioequivalence, the evaluation of the long-run performance indicates possible problems. These fictitious data were generated with high intersubject variability, which causes poor long-run performance.

If we conduct a bioequivalence test with the data published in Chow and Liu (2009, 71), which we introduced in [R] **pk** and fully described in [R] **pkshape**, we observe that the probability that the test limits are within the equivalence limits is high.

```
. use http://www.stata-press.com/data/r14/chowliu2
. set seed 1
. pkequiv outcome treat period seq id
      Classic confidence interval for bioequivalence
------------------------------------------------------------------------------
                     [equivalence limits]        [    test limits   ]
------------------------------------------------------------------------------
     difference:   -16.512        16.512         -8.698         4.123
          ratio:       80%          120%         89.464%      104.994%
------------------------------------------------------------------------------
Probability test limits are within equivalence limits =    0.9960
Note: Reference treatment = 1
```

For these data, the test limits are well within the equivalence limits, and the probability that the test limits are within the equivalence limits is 99.8%. ◁

▷ Example 2

We compute a confidence interval that is symmetric about zero:

```
. pkequiv outcome treat period seq id, symmetric
      Westlake's symmetric confidence interval for bioequivalence
------------------------------------------------------------------------------
                     [Equivalence limits]       [    Test mean   ]
------------------------------------------------------------------------------
Test formulation:     75.145    89.974               80.272
------------------------------------------------------------------------------
Note: Reference treatment = 1
```

The reported equivalence limit is constructed symmetrically about the reference mean, which is equivalent to constructing a confidence interval symmetric about zero for the difference in the two drugs. In the output above, we see that the test formulation mean of 80.272 is within the equivalence limits, indicating that the test drug is bioequivalent to the reference drug.

pkequiv displays interval hypothesis tests of bioequivalence if you specify the tost or the anderson option, or both. For example,

```
. pkequiv outcome treat period seq id, tost anderson
      Classic confidence interval for bioequivalence
------------------------------------------------------------------------------
                    [equivalence limits]        [    test limits   ]
------------------------------------------------------------------------------
     difference:   -16.512       16.512         -8.698        4.123
          ratio:       80%         120%         89.464%     104.994%
------------------------------------------------------------------------------
Probability test limits are within equivalence limits =    0.9960
Schuirmann's two one-sided tests
------------------------------------------------------------------------------
upper test statistic =    -5.036                    p-value =     0.000
lower test statistic =     3.810                    p-value =     0.001
Anderson and Hauck's test
------------------------------------------------------------------------------
noncentrality parameter =     4.423
         test statistic =    -0.613      empirical p-value =    0.0005
Note: Reference treatment = 1
```

Both of Schuirmann's one-sided tests are highly significant, suggesting that the two drugs are bioequivalent. A similar conclusion is drawn from the Anderson and Hauck test of bioequivalence. ◁

Stored results

pkequiv stores the following in r():

Scalars

r(stddev)	pooled-sample standard deviation of period differences from both sequences
r(uci)	upper confidence interval for a classic interval
r(lci)	lower confidence interval for a classic interval
r(delta)	delta value used in calculating a symmetric confidence interval
r(u3)	upper confidence interval for Fieller's confidence interval
r(l3)	lower confidence interval for Fieller's confidence interval

Methods and formulas

The lower confidence interval for the difference in the two treatments for the classic shortest confidence interval is

$$L_1 = \left(\overline{Y}_T - \overline{Y}_R\right) - t_{(\alpha,n_1+n_2-2)}\widehat{\sigma}_d\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

The upper limit is

$$U_1 = \left(\overline{Y}_T - \overline{Y}_R\right) + t_{(\alpha,n_1+n_2-2)}\widehat{\sigma}_d\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

The limits for the ratio measure are

$$L_2 = \left(\frac{L_1}{\overline{Y}_R} + 1\right) 100\%$$

and

$$U_2 = \left(\frac{U_1}{\overline{Y}_R} + 1\right) 100\%$$

where $\overline{Y}_T$ is the mean of the test formulation of the drug, $\overline{Y}_R$ is the mean of the reference formulation of the drug, and $t_{(\alpha,n_1+n_2-2)}$ is the t distribution with $n_1 + n_2 - 2$ degrees of freedom. $\widehat{\sigma}_d$ is the pooled sample variance of the period differences from both sequences, defined as

$$\widehat{\sigma}_d = \frac{1}{n_1 + n_2 - 2} \sum_{k=1}^{2} \sum_{i=1}^{n_k} \left(d_{ik} - \overline{d}_{.k}\right)^2$$

The finite sample performance of the classical confidence interval is assessed via bootstrap simulation of the confidence interval. One thousand bootstrap samples are drawn using the patient IDs as clusters. For each sample, the classical confidence interval is constructed and compared with the equivalence limits.

The upper and lower limits for the symmetric confidence interval are $\overline{Y}_R + \Delta$ and $\overline{Y}_R - \Delta$, where

$$\Delta = k_1 \widehat{\sigma}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} - \left(\overline{Y}_T - \overline{Y}_R\right)$$

and (simultaneously)

$$\Delta = -k_2 \widehat{\sigma}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} + 2\left(\overline{Y}_T - \overline{Y}_R\right)$$

and k_1 and k_2 are computed iteratively to satisfy the above equalities and the condition

$$\int_{k_1}^{k_2} f(t)dt = 1 - 2\alpha$$

where $f(t)$ is the probability density function of the t distribution with $n_1 + n_2 - 2$ degrees of freedom.

See Chow and Liu (2009, 88–92) for details about calculating the confidence interval based on Fieller's theorem.

The two test statistics for the two one-sided tests of equivalence are

$$T_L = \frac{\left(\overline{Y}_T - \overline{Y}_R\right) - \theta_L}{\widehat{\sigma}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

and

$$T_U = \frac{\left(\overline{Y}_T - \overline{Y}_R\right) - \theta_U}{\widehat{\sigma}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

where $-\theta_L = \theta_U$ and are the regulated confidence limits.

The logic of the Anderson and Hauck test is tricky; see Chow and Liu (2009) for a complete explanation. However, the test statistic is

$$T_{AH} = \frac{\left(\overline{Y}_T - \overline{Y}_R\right) - \left(\frac{\theta_L + \theta_U}{2}\right)}{\widehat{\sigma}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

and the noncentrality parameter is estimated by

$$\widehat{\delta} = \frac{\theta_U - \theta_L}{2\widehat{\sigma}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

The empirical p-value is calculated as

$$p = F_t\left(|T_{AH}| - \widehat{\delta}\right) - F_t\left(-|T_{AH}| - \widehat{\delta}\right)$$

where F_t is the cumulative distribution function of the t distribution with $n_1 + n_2 - 2$ degrees of freedom.

References

Anderson, S., and W. W. Hauck. 1983. A new procedure for testing equivalence in comparative bioavailability and other clinical trials. *Communications in Statistics—Theory and Methods* 12: 2663–2692.

Chow, S.-C., and J.-P. Liu. 2009. *Design and Analysis of Bioavailability and Bioequivalence Studies*. 3rd ed. Boca Raton, FL: Chapman & Hall/CRC.

Fieller, E. C. 1954. Some problems in interval estimation. *Journal of the Royal Statistical Society, Series B* 16: 175–185.

Kutner, M. H., C. J. Nachtsheim, J. Neter, and W. Li. 2005. *Applied Linear Statistical Models*. 5th ed. New York: McGraw–Hill/Irwin.

Locke, C. S. 1984. An exact confidence interval from untransformed data for the ratio of two formulation means. *Journal of Pharmacokinetics and Biopharmaceutics* 12: 649–655.

Schuirmann, D. J. 1989. Confidence intervals for the ratio of two means from a cross-over study. In *Proceedings of the Biopharmaceutical Section*, 121–126. Washington, DC: American Statistical Association.

Westlake, W. J. 1976. Symmetrical confidence intervals for bioequivalence trials. *Biometrics* 32: 741–744.

Also see

[R] **pk** — Pharmacokinetic (biopharmaceutical) data