---

pkequiv — Perform bioequivalence tests

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

Description

pkequiv performs bioequivalence testing for two treatments. By default, pkequiv calculates a standard confidence interval (CI) symmetric about the difference between the two treatment means. pkequiv also calculates CIs symmetric about 0 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

Same 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	
<u>compare</u>(*string*)	compare the two specified values of the treatment variable
<u>limit</u>(#)	equivalence limit (between 0.01 and 0.99); default is `limit(0.2)`
<u>level</u>(#)	set confidence level; default is `level(90)`
<u>fie</u>ller	calculate CI by Fieller's theorem
<u>symmetric</u>	calculate symmetric CI
<u>anderson</u>	Anderson and Hauck hypothesis test for bioequivalence
<u>tost</u>	two one-sided hypothesis tests for bioequivalence
<u>noboot</u>	do not estimate probability that CI lies within confidence limits

collect is allowed; see [U] **11.1.10 Prefix commands**.

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 limit(0.2). The equivalence limit can be changed only symmetrically; for example, 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 CIs. The default is level(90). This setting is not controlled by the set level command.

fieller specifies that a CI based on Fieller's theorem be calculated.

symmetric specifies that a symmetric CI be calculated.

anderson specifies that the Anderson and Hauck (1983) hypothesis test for bioequivalence be computed. This option is ignored when calculating a CI based on Fieller's theorem or when calculating a CI that is symmetric about 0.

tost specifies that the two one-sided hypothesis tests for bioequivalence be computed. This option is ignored when calculating a CI based on Fieller's theorem or when calculating a CI that is symmetric about 0.

noboot prevents the estimation of the probability that the CI lies within the equivalence 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 use pkdata.dta from example 2 of [R] **pk**. We use pkcollapse and pkshape on the data as discussed in that example. After sorting the data with sort, our data appear as follows:

```
. use https://www.stata-press.com/data/r18/pkdata, clear
(Fictional drug concentration data)
. pkcollapse time conc1 conc2, id(id) keep(seq) stat(auc)
...............................
. pkshape id seq auc*, order(RT TR)
. sort period id
. list, sep(4)
```

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

We use pkequiv to conduct a bioequivalence test between reference treatment "R" and test treatment "T".

```
. set seed 123
. 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 =     0.6590
        Note: Reference treatment = 1.
```

The default output for pkequiv shows a CI for the difference between the means (test limits), the ratio of the means, and the user-specified equivalence limits. The classic CI 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. Following Chow and Liu (2009), we can apply the ± 20 rule to the difference and ratio measures to determine equivalence limits. For these data, the CI for the difference must be entirely contained within the range $\left[-30.296, 30.296\right]$ and for the ratio between 80% and 120%. Here the test limits are within the equivalence limits. Although the test limits are inside the equivalence limits, there is only a 66% assurance that the observed CI 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 intrasubject variability, which causes poor long-run performance.

We now use the data published in Chow and Liu (2009, 71), which we describe in example 1 of [R] **pkshape**. As shown in example 1, we also use pkshape to reshape the data. We use the same pkequiv command used above to conduct a bioequivalence test on the data.

```
. use https://www.stata-press.com/data/r18/chowliu, clear
. pkshape id seq period1 period2, order(RT TR)
. set seed 123
. 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 =     0.9940
        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 99.4%.

◁

▷ Example 2

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

```
. set seed 123
. 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 =   0.9940

Schuirmann's two one-sided tests

upper test statistic =	-5.036	p-value =	0.000
lower test statistic =	3.810	p-value =	0.000

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 limit of a classic CI
r(lci)	lower limit of a classic CI
r(delta)	delta value used in calculating a symmetric CI
r(u3)	upper limit of Fieller's CI
r(l3)	lower limit of Fieller's CI

Methods and formulas

The lower limit for the difference in the two treatments for the classic shortest CI 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 marginal mean of the test formulation of the drug, $\overline{Y}_R$ is the marginal 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 CI is assessed via bootstrap simulation of the CI. One thousand bootstrap samples are drawn using the patient IDs as clusters. For each sample, the classical CI is constructed and compared with the equivalence limits.

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

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

and (simultaneously)

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

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 CI 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 $-\theta_L = \theta_U$, both of these being 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. https://doi.org/10.1080/03610928308828634.

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

Fieller, E. C. 1954. Some problems in interval estimation. *Journal of the Royal Statistical Society, Series B* 16: 175–185. https://doi.org/10.1111/j.2517-6161.1954.tb00159.x.

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. https://doi.org/10.1007/BF01059558.

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. https://doi.org/10.2307/2529259.

Also see

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

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