Package ‘adk’

February 6, 2008

Type Package

Title Anderson-Darling K-Sample Test and Combinations of Such Tests

Version 1.0.1

Date 2008-01-20

Author Fritz Scholz

Maintainer <fscholz@u.washington.edu>

Description The Anderson-Darling K-sample test can be used to test whether several independent random samples of various sizes come from the same but unspecified continuous distribution. It is a rank test and consistent against all alternatives. A low to moderate number of tied observations can be tolerated. The combination of such tests can be used to test whether M groups of samples (with K allowed to vary from group to group) come from respective common distributions, which may vary from group to group. This is useful in testing for treatment effects in randomized (incomplete) block designs or in examining whether several laboratories perform equally well when asked to measure a sufficient number of test speciments from different batches or materials.

R topics documented:

adk.combined.test 1
adk-package 3
adk.test 5
print.adk 7

`adk.combined.test` *Combined Anderson-Darling K-Sample Tests*

Description

This function combines several independent Anderson-Darling k-sample tests into one overall test of the hypothesis that the independent samples within each group come from a common unspecified distribution, while the common distributions may vary from group to group. The k for each group of k independent samples may also change from group to group. All samples from all groups are independent and are assumed to come from continuous distributions. Also provided is a version that adjusts for a moderate number of ties (due to rounding).

Usage

```
adk.combined.test(...)
```

Arguments

`...` Either several lists, say L.1, ... , L.M, where list L.i contains K.i sample vectors of respective sizes n.i[1], ..., n.i[K.i], where n.i[j] > 4 is recommended

or a list of such lists.

Details

If AD.i is the Anderson-Darling criterion for the i-th group of K.i samples, its standardized test statistic is T.i = (AD.i - mu.i)/sig.i, with mu.i and sig.i representing mean and standard deviation of AD.i. This test statistic is used to test the hypothesis that the samples in the i-th group all come from the same but unspecified continuous distribution function F(x).

The combined Anderson-Darling criterion is AD.c=AD.1 + $\ldots$ + AD.M and T.combined = (AD.c - mu.c)/sig.c is the standardized test statistic, where mu.c and sig.c represent the mean and standard deviation of AD.c. This test statistic is used to simultaneously test whether the samples in each group come from the same continuous distribution function F(x), However, the unspecified common distribution function F(x) may change from group to group.

Value

A list with components

`M` number of groups of samples being compared

`n.samples` list of vectors, each vector giving the sample sizes for each group of samples being compared

`nt` vector of total sample sizes involved in each of the k.data comparisons

`n.ties` vector giving the number of ties in each the k.data comparison group

`adk.i`	(2*k.data) * 3 matrix containing the T.kN, P-value, and extrapolation for the M individual Anderson-Darling tests, not adjusted for ties and adjusted for ties
`mu`	vector of means of the k.data AD statistics
`sig`	vector of standard deviations of the k.data AD statistics
`adk.c`	2*3 matrix containing T.combined, P-value, and extrapolation for the combined test not adjusted for ties and adjusted for ties
`mu.c`	mean of the combined AD statistic
`sig.c`	standard deviation of the combined AD statistic
`warning`	logical indicator, warning = TRUE when at least one of the sample sizes is < 5.

Note

This test is useful in analyzing treatment effects in randomized (incomplete) block experiments and in examining performance equivalence of several laboratories when presented with different test materials for comparison.

Author(s)

Fritz Scholz

References

Scholz F.W. and Stephens M.A. (1987), K-sample Anderson-Darling Tests, *Journal of the American Statistical Association,* **Vol 82, No. 399**, 918–924.

See Also

See also `adk.test`

Examples

```
## Create two lists of sample vectors.
x1 <- list( c(1, 3, 2, 5, 7), c(2, 8, 1, 6, 9, 4), c(12, 5, 7, 9, 11) )
x2 <- list( c(51, 43, 31, 53, 21, 75), c(23, 45, 61, 17, 60) )

## Run adk.combined.test.
adk.combined.out <- adk.combined.test(x1,x2)  # or out <- adk.combined.test(list(x1,x2))

## Examine list objects in adk.combined.out.
names(adk.combined.out)

## Extract matrix components adk.i and adk.c
adk.combined.out$adk.i
adk.combined.out$adk.c

## Fully formatted output
adk.combined.out
```

`adk-package` *The Package adk Contains a K-Sample Anderson-Darling Test and its Combinations*

Description

The K-sample Anderson-Darling test `adk.test` is used to test the hypothesis that K samples of various sizes come from a common continuous distribution that is otherwise unspecified. It is a rank test and it is consistent against all alternatives. The combined version of the test `adk.combined.test` is used to test several such hypotheses at the same time and the common distribution may vary from hypothesis to hypothesis.

Details

Package:	adk
Type:	Package
Version:	1.0
Date:	2008-01-21
License:	GPL version 2 or newer

Author(s)

Fritz Scholz

Maintainer: <fscholz@u.washington.edu>

References

Scholz F.W. and Stephens M.A. (1987), K-sample Anderson-Darling Tests, *Journal of the American Statistical Association,* **Vol 82, No. 399**, 918–924.

Examples

```
## Example using adk.test
x <- list(c(1,3,2,5,7),c(2,8,1,6,9,4),c(12,5,7,9,11))
out <- adk.test(x) # or out <- adk.test(c(1,3,2,5,7),c(2,8,1,6,9,4),c(12,5,7,9,11))

## Example using adk.combined.test
x1 <- list( c(1, 3, 2, 5, 7), c(2, 8, 1, 6, 9, 4), c(12, 5, 7, 9, 11) )
x2 <- list( c(51, 43, 31, 53, 21, 75), c(23, 45, 61, 17, 60) )
adk.combined.out <- adk.combined.test(x1,x2)  # or out <- adk.combined.test(list(x1,x2))
```

`adk.test` *Anderson-Darling K-Sample Test*

Description

The Anderson-Darling `k`-sample test may be used to test the hypothesis that `k` samples of various sizes (> 4) come from one common continuous distribution. It is a rank test and it is consistent against all alternatives, a property not shared by the Kruskal-Wallis `k`-sample rank test. Also provided is a version that adjusts for a moderate number of ties (due to rounding).

Usage

```
adk.test(...)
```

Arguments

`...`	Either several sample vectors of respective sizes n.1, ..., n.k, with n.i > 4 recommended, or a list of such sample vectors

Details

See the given reference for details on the Anderson-Darling `k`-sample criterion AD and its modification in case of ties. The standardized value of AD, i.e., T = (AD - mu)/`sig`, is used as test statistic. Here mu = `k-1` and `sig` are the mean and standard deviation of AD. The `P-value` = P(T > `t.obs`) corresponding to an observed `t.obs` of T is computed by quadratic interpolation w.r.t. 1/sqrt(`mu`) and by quadratic interpolation w.r.t. log(p/(1-p)), where p is the tail probability corresponding to the quantiles given in Table 1 of the cited reference. Both interpolations are reasonably accurate. For p beyond the range [.01,.25] of Table 1 linear exptrapolation is used w.r.t. the log(p/(1-p)) fit. Such extrapolation affects the accuracy of the P-value caluculation to some extent but this should not strongly affect any decisions regarding the tested hypothesis.

Value

A list of class adk with components

`k`	number of samples being compared
`ns`	vector of the k sample sizes c(n.1, ...,n.k)
`n`	total sample size = n.1 + ... + n.k
`n.ties`	number of ties in the combined set of all n observations
`sig`	standard deviation of the AD statistic
`adk`	2 x 3 matrix containing `t.obs`, `P-value`, `extrapolation`, not adjusting for ties and adjusting for ties. `extrapolation = 1` when the `P-value` was extrapolated.

`warning` logical variable, `warning = TRUE` if n.i < 5 for at least one of the samples, otherwise `warning = FALSE` .

Author(s)

Fritz Scholz

References

Scholz F.W. and Stephens M.A. (1987), K-sample Anderson-Darling Tests, *Journal of the American Statistical Association,* **Vol 82, No. 399**, 918–924.

See Also

`kruskal.test` as a nonparametric alternative to `adk.test` and `adk.combined.test` for combining several such tests for different and independent groups of samples

Examples

```
## Create input list of 3 sample vectors.
x <- list(c(1,3,2,5,7),c(2,8,1,6,9,4),c(12,5,7,9,11))
out <- adk.test(x) # or out <- adk.test(c(1,3,2,5,7),c(2,8,1,6,9,4),c(12,5,7,9,11))
## Examine the component names of out
names(out)

## Examine the matrix adk of out.
out$adk

## Fully print formatted object out of class adk.
out
```

---

print.adk *Print Function for Objects of Class adk*

Description

This print function is invoked for printing formatted output of class adk as it is output from adk.test or adk.combined.test.

Usage

```
print.adk(x,...)
```

Arguments

`x`	an object of class adk, as ouput by adk.test or adk.combined.test
`...`	further arguments passed to or from other methods

Details

The formatted output is different for objects generated by adk.test and for object generated by adk.combined.test.

Author(s)

Fritz Scholz

See Also

`adk.test` and `adk.combined.test`

Examples

```
## Create input list of 3 sample vectors.
x <- list(c(1,3,2,5,7),c(2,8,1,6,9,4),c(12,5,7,9,11))
adk.test(x)
```

Index

∗Topic **classes**
    print.adk, 7
∗Topic **htest**
    adk-package, 3
    adk.combined.test, 1
    adk.test, 5
∗Topic **nonparametric**
    adk-package, 3
    adk.combined.test, 1
    adk.test, 5
∗Topic **package**
    adk-package, 3
∗Topic **print**
    print.adk, 7

adk-package, **3**
adk.combined.test, **1**, *6*, *7*
adk.test, *3*, **5**, *7*

kruskal.test, *6*

print.adk, **7**