

# Package ‘runDRT’

June 1, 2024

**Title** Run Doubly Ranked Tests

**Version** 0.1.0

**Description** Doubly ranked tests are nonparametric tests for grouped functional and multivariate data. The testing procedure first ranks a matrix (or three dimensional array) of data by column (if a matrix) or by cell (across the third dimension if an array). By default, it calculates a sufficient statistic for the subject's order within the sample using the observed ranks, taken over the columns or cells. Depending on the number of groups,  $G$ , the summarized ranks are then analyzed using either a Wilcoxon Rank Sum test ( $G = 2$ ) or a Kruskal-Wallis ( $G$  greater than 2).

**Encoding** UTF-8

**RoxygenNote** 7.3.1

**License** MIT + file LICENSE

**Suggests** fda ( $\geq 6.1.4$ ), FDboost ( $\geq 1.1.2$ ), refund ( $\geq 0.1.32$ ),  
testthat ( $\geq 3.0.0$ )

**Config/testthat/edition** 3

**NeedsCompilation** no

**Author** Mark J. Meyer [aut, cre, cph] (<<https://orcid.org/0000-0003-3942-9675>>)

**Maintainer** Mark J. Meyer <mjm556@georgetown.edu>

**Repository** CRAN

**Date/Publication** 2024-06-01 10:30:03 UTC

## R topics documented:

run\_drt . . . . . 2

**Index** . . . . . 5

run\_drt

*Run Doubly Ranked Tests***Description**

Performs two (or more) sample doubly ranked tests on pre-processed functional data, formatted as either a matrix (for functions) or an array (for surfaces).

**Usage**

```
run_drt(X, G, method = c("suff.rank", "avg.rank"), data.names = NULL)

## Default S3 method:
run_drt(X, G, method = c("suff.rank", "avg.rank"), data.names = NULL)

## S3 method for class 'formula'
run_drt(formula, ...)
```

**Arguments**

X	an n by T matrix or an S by T by n array containing the functions (or surfaces) to analyze.
G	a vector of length n containing the grouping variable.
method	statistic for summarizing the ranks: 'suff.rank' for sufficient statistic (the default) or 'avg.rank' for arithmetic average.
data.names	a vector of length two containing names that describe X and G.
formula	a formula of the form $X \sim G$ .
...	additional arguments to pass to <code>run_drt.default()</code> , e.g. method.

**Details**

Doubly ranked tests are non-parametric tests that first rank functions (or surfaces) by time (or location). Next, the procedure summarizes the observed ranks using a statistic. The summarized ranks are then analyzed using either a Wilcoxon rank sum test or a Kruskal-Wallis test. To perform a doubly ranked test, realizations of functions must be stored in an n by T matrix where n is the total number of observed functions and T is the number of realizations per function (commonly time points or locations). Surface data in an S by T by n array can be analyzed as well, although currently this feature has under gone only limited testing.

By default, `run_drt()` implements a sufficient statistic when summarizing the ranks of each observed function across T, i.e. the argument method defaults to `method = suff.rank`. This statistic has the form

$$t(z) = \frac{1}{T} \sum_{t=1}^T \log \left[ \left( \frac{z_t}{n} - \frac{1}{2n} \right) / \left( 1 - \frac{z_t}{n} + \frac{1}{2n} \right) \right],$$

where  $z_t$  is the observed rank at time  $t$ . See Meyer (2024) for additional details. The average rank may also be used by setting `method = 'avg.rank'`, although this summary has not undergone testing in the doubly ranked context.

Regardless of the statistic used, the summarized ranks are the analyzed using either `wilcox.test()` or `kruskal.test()`, depending on the number of groups in  $G$ .

For functional data, Meyer (2024) suggests using `refund::fpca.face()` for pre-processing the data, but  $X$  can be pre-processed using any functional data approach or it can just be the raw data. `run_drt()` itself performs no pre-processing and takes  $X$  as inputted.

## Value

A list with class "htest" containing the following components:

<code>statistic</code>	the value of the test statistic with a name describing it.
<code>parameter</code>	the parameter(s) for the exact distribution of the test statistic.
<code>p.value</code>	the p-value for the test.
<code>null.value</code>	the location parameter.
<code>alternative</code>	a character string describing the alternative hypothesis.
<code>data.name</code>	a character string giving the names of the data.
<code>test_details</code>	the output from the internally run Wilcoxon rank sum or Kruskal-Wallis test.
<code>method</code>	character string giving the type of doubly ranked test performed.
<code>ranks</code>	a list containing the ranks by column (if $X$ is a matrix) and the summarized ranks.
<code>data</code>	a list containing $X$ and $G$ .

## References

Meyer, MJ (2024). Doubly ranked tests for grouped functional data. *Available on arXiv* at <https://arxiv.org/abs/2306.14761>.

## Examples

```
#### Two Sample Problem: Resin Viscosity ####
library(FDboost)
data("viscosity")

Xv <- matrix(viscosity$visAll, nrow = nrow(viscosity$visAll), ncol = ncol(viscosity$visAll))
fXv <- refund::fpca.face(Xv)
Yvis <- fXv$Yhat
TR <- viscosity$T_A
```

```
run_drt(Yvis ~ TR)

#### Four Sample Problem: Canadian Weather ####
R    <- fda::CanadianWeather$region
XT   <- t(fda::CanadianWeather$dailyAv[,,'Temperature.C'])
fXT  <- refund::fpca.face(XT)
YT   <- fXT$Yhat

run_drt(YT ~ R)
```

# Index

`kruskal.test()`, 3

`refund::fpca.face()`, 3

`run_drt`, 2

`run_drt.default()`, 2

`wilcox.test()`, 3