

dPvalAggregate

January 19, 2018

dPvalAggregate	<i>Function to aggregate p values</i>
----------------	---------------------------------------

Description

dPvalAggregate is supposed to aggregate a input matrix p-values into a vector of aggregated p-values. The aggregate operation is applied to each row of input matrix, each resulting in an aggregated p-value. The method implemented can be based on the order statistics of p-values or according to Fisher's method or Z-transform method.

Usage

```
dPvalAggregate(pmatrix, method = c("orderStatistic", "fishers",  
  "Ztransform",  
  "logistic"), order = ncol(pmatrix), weight = rep(1, ncol(pmatrix)))
```

Arguments

pmatrix	a data frame or matrix of p-values
method	the method used. It can be either "orderStatistic" for the method based on the order statistics of p-values, or "fishers" for Fisher's method (summation of logs), or "Ztransform" for Z-transform test (summation of z values, Stouffer's method) and the weighted Z-test, or "logistic" for summation of logits
order	an integer specifying the order used for the aggregation according to the order statistics of p-values
weight	a vector specifying the weights used for the aggregation according to Z-transform method

Value

- ap: a vector with the length nrow(pmatrix), containing aggregated p-values

Note

For each row of input matrix with the c columns, there are c p-values that are uniformly independently distributed over $[0,1]$ under the null hypothesis (uniform distribution). According to the order statistics, they follow the Beta distribution with the parameters $a = \text{order}$ and $b = c - \text{order} + 1$. According to the Fisher's method, after transformation by $-2 * \sum^c \log(pvalue)$, they follow Chi-Squared distribution. According to the Z-transform method, first converts the one-tailed P-values into standard normal deviates Z , then combines Z via $\frac{\sum^c (w * Z)}{\sqrt{\sum^c (w^2)}}$, where w is the weight (usually square root of the sample size if the weighted Z-test; 1 if Z-transform test), and finally the combined Z follows the standard normal distribution to test the cumulative/aggregated evidence on the common null hypothesis. The logistic method is defined as $\sum^c \log(\frac{pvalue}{1-pvalue}) * 1/C$, where $C = \text{sqr}t((k\pi^2(5k+2))/(3(5k+4)))$, following Student's t distribution. Generally speaking, Fisher's method places greater emphasis on small p-values, while the Z-transform method on equal footings, the logistic method provides a compromise between these two. In other words, the Z-transform method does well in problems where evidence against the combined null is spread more than a small fraction of the individual tests, or when the total evidence is weak; Fisher's method does best in problems where the evidence is concentrated in a relatively small fraction of the individual tests or when the evidence is at least moderately strong.

See Also

[dPvalAggregate](#)

Examples

```
# 1) generate an iid uniformly-distributed random matrix of 1000x3
pmatrix <- cbind(runif(1000), runif(1000), runif(1000))

# 2) aggregate according to the order statistics
ap <- dPvalAggregate(pmatrix, method="orderStatistic")

# 3) aggregate according to the Fisher's method
ap <- dPvalAggregate(pmatrix, method="fishers")

# 4) aggregate according to the Z-transform method
ap <- dPvalAggregate(pmatrix, method="Ztransform")

# 5) aggregate according to the logistic method
ap <- dPvalAggregate(pmatrix, method="logistic")
```