gTests

Yang Meng

11/9/2019

```
library('gTests')
```

Load the data

```
data(example_discrete)
```

Two sample data

Data structure

```
# count in two samples
# 720 unique values
head(counts1)
```

```
##      sample1 sample2
## [1,]       1       0
## [2,]       1       5
## [3,]       1       4
## [4,]       0       1
## [5,]       0       4
## [6,]       2       4
```

Distance

```
# distance matrix of 720 unique values
dim(ds1)
```

```
## [1] 720 720
```

Test result

```
# 5-MST
E = getGraph(counts1, ds1, 5)
g.tests_discrete(E, counts1, test.type = 'w')
```

```
## $weighted
## $weighted$test.statistic_a
## [1] 1.474036
##
## $weighted$pval.approx_a
```

```
## [1] 0.07023595
##
## $weighted$test.statistic_u
## [1] 2.588258
##
## $weighted$pval.approx_u
## [1] 0.004823137
```

Comparison

```
pval = matrix(nrow = 30, ncol = 3)
for (k in 1:30){
  E1 = getGraph(counts1, ds1, k)
  pval[k, 1] = g.tests_discrete(E1, counts1, test.type = 'w')$weighted$pval.approx_u
  E2 = getGraph(counts2, ds2, k)
  pval[k, 2] = g.tests_discrete(E2, counts2, test.type = 'w')$weighted$pval.approx_u
  E3 = getGraph(counts3, ds3, k)
  pval[k, 3] = g.tests_discrete(E3, counts3, test.type = 'w')$weighted$pval.approx_u
}
```

```
plot(pval[,1], ylab = 'p-value', xlab = 'k (K-MST)')
points(pval[,2], pch=2)
points(pval[,3], pch=3)
abline(h = .05)
legend("topright", legend = c("data 1", "data 2", "data 3"),pch = 1:3)
```