GTestClustersModules

BV

18.06.2019

```
# Version 0.1.
# Bernhard Voelkl
# Last modified: 18.06.2019
#
#-------------------------------------------------------------------------
---------------------------
# The purpose of this script is to create a contingency table for DTW
clusters and AI modules
# for each pen and calculate the overall G-statistic (likelihood ratio X2
value) to see whether
# DTW clusters and AI modules are independent (uncorrelated)
# The critical statistic is the X2 value (plus df and p-value) from the
Likelihood Ratio Test
#-------------------------------------------------------------------------
---------------------------

# INITILALIZATION

library(MASS)
setwd("M:/Chicken Research/ChickenSocialNetwork")

# Data Import

nwm<-read.csv("networkmeasures_DTWcluster.csv")
set.seed(1492)

# Main


# This is for pen 11:

contingencytable<-c()
for(i in 1:4)
{
  for(m in 1:5)
  {
    contingencytable<-c(contingencytable,(sum(nwm$Community==11+0.1*m &
nwm$mycl==i)))
  }
}
data <- as.table(matrix(contingencytable, nrow = 4, byrow = TRUE, dimnames =
```

```
list(DTW_Clusters = c('dtw1', 'dtw2', 'dtw3','dtw4'), AI_Modules = c('m1',
'm2', 'm3','m4','m5'))))
loglm <- loglm( ~ DTW_Clusters + AI_Modules, data = data)
loglm

## Call:
## loglm(formula = ~DTW_Clusters + AI_Modules, data = data)
##
## Statistics:
##                       X^2 df     P(> X^2)
## Likelihood Ratio 89.18889 12 7.083223e-14
## Pearson          85.84591 12 3.128608e-13

fisher.test(data, workspace = 5000000)

##
##  Fisher's Exact Test for Count Data
##
## data:  data
## p-value = 7.333e-16
## alternative hypothesis: two.sided

# This is for pen 12:

contingencytable<-c()
for(i in 1:4)
{
  for(m in 1:4)
  {
    contingencytable<-c(contingencytable,(sum(nwm$Community==12+0.1*m &
nwm$mycl==i)))
  }
}
data <- as.table(matrix(contingencytable, nrow = 4, byrow = TRUE, dimnames =
list(DTW_Clusters = c('dtw1', 'dtw2', 'dtw3','dtw4'), AI_Modules = c('m1',
'm2', 'm3','m4'))))
loglm <- loglm( ~ DTW_Clusters + AI_Modules, data = data)
loglm

## Call:
## loglm(formula = ~DTW_Clusters + AI_Modules, data = data)
##
## Statistics:
##                       X^2 df     P(> X^2)
## Likelihood Ratio 61.70193  9 6.292301e-10
## Pearson          70.92053  9 1.004608e-11

fisher.test(data, workspace = 5000000)

##
##  Fisher's Exact Test for Count Data
```

```
##
## data:  data
## p-value = 1.014e-11
## alternative hypothesis: two.sided

# This is for pen 13:

contingencytable<-c()
for(i in 1:4)
{
  for(m in 1:5)
  {
    contingencytable<-c(contingencytable,(sum(nwm$Community==13+0.1*m &
nwm$mycl==i)))
  }
}
data <- as.table(matrix(contingencytable, nrow = 4, byrow = TRUE, dimnames =
list(DTW_Clusters = c('dtw1', 'dtw2', 'dtw3','dtw4'), AI_Modules = c('m1',
'm2', 'm3','m4','m5'))))
loglm <- loglm( ~ DTW_Clusters + AI_Modules, data = data)
loglm

## Call:
## loglm(formula = ~DTW_Clusters + AI_Modules, data = data)
##
## Statistics:
##                       X^2 df     P(> X^2)
## Likelihood Ratio 88.33345 12 1.036948e-13
## Pearson               NaN 12          NaN

fisher.test(data, workspace = 5000000)

##
##  Fisher's Exact Test for Count Data
##
## data:  data
## p-value < 2.2e-16
## alternative hypothesis: two.sided

# This is for pen 14:

contingencytable<-c()
for(i in 1:4)
{
  for(m in 1:6)
  {
    contingencytable<-c(contingencytable,(sum(nwm$Community==14+0.1*m &
nwm$mycl==i)))
  }
}
data <- as.table(matrix(contingencytable, nrow = 4, byrow = TRUE, dimnames =
```

```r
list(DTW_Clusters = c('dtw1', 'dtw2', 'dtw3','dtw4'), AI_Modules = c('m1',
'm2', 'm3','m4','m5','m6'))))
loglm <- loglm( ~ DTW_Clusters + AI_Modules, data = data)
loglm
```

```
## Call:
## loglm(formula = ~DTW_Clusters + AI_Modules, data = data)
##
## Statistics:
##                       X^2 df     P(> X^2)
## Likelihood Ratio 60.51299 15 2.058136e-07
## Pearson               NaN 15          NaN
```

```r
fisher.test(data, workspace = 5000000)
```

```
##
##  Fisher's Exact Test for Count Data
##
## data:  data
## p-value = 9e-12
## alternative hypothesis: two.sided
```