Tutorial: functions for biocultural complexity

M. Humberto Reyes-Valdés

5/30/2018

This tutorial shows the use of the functions coded in the script *biocultural.R*, to analyzie biocultural complexity from a table of cultures and species usage. In this case, I use the file *toyset.csv*, which is an artificial simple data set. The platform and language is R.

1. Open an R console

2. Read the data set

```
dat.sim<-read.csv("toyset.csv",head=T)
```

3. Visualize data features

```
dat.sim
```

```
##   culture rep sp1 sp2 sp3 sp4 sp5
## 1      c1   1   0  25  39  27   9
## 2      c2   1 110  10   0   0   0
## 3      c3   1   0  26  24  28  10
## 4      c4   1   0  11 119   0   0
```

```
#rep is the representation frequency of cultural groups.
#In this case the assumption is that all groups are equally represented,
#and one can write 1 for each entry.
str(dat.sim)
```

```
## 'data.frame':    4 obs. of  7 variables:
##  $ culture: Factor w/ 4 levels "c1","c2","c3",..: 1 2 3 4
##  $ rep    : int  1 1 1 1
##  $ sp1    : int  0 110 0 0
##  $ sp2    : int  25 10 26 11
##  $ sp3    : int  39 0 24 119
##  $ sp4    : int  27 0 28 0
##  $ sp5    : int  9 0 10 0
```

4. Source script

```
source("biocultural.R")
```

5. Create an object with biocultural parameters

```
ob.sim<-bc.info(dat.sim)
names(ob.sim)
```

```
##  [1] "cu"               "species"          "specificities"
##  [4] "r.specificities"  "s.diversity"      "specializations"
##  [7] "r.specializations" "mutualInfo"       "BC"
## [10] "r.BC"
```

6. Table of species and their specificities and relative specificities

```
sp.tab<-data.frame("Species"=ob.sim$species,"Specificity"=ob.sim$specificities,
                   "Relative.specificity"=ob.sim$r.specificities)
sp.tab
```

```
##   Species Specificity Relative.specificity
## 1    sp1   2.0000000            1.0000000
## 2    sp2   0.2165054            0.1082527
## 3    sp3   0.6081559            0.3040780
## 4    sp4   1.0048459            0.5024230
## 5    sp5   1.0097404            0.5048702
```

```r
library(knitr)
kable(sp.tab) #Just to get fancy
```

Species	Specificity	Relative.specificity
sp1	2.0000000	1.0000000
sp2	0.2165054	0.1082527
sp3	0.6081559	0.3040780
sp4	1.0048459	0.5024230
sp5	1.0097404	0.5048702

7. Cultures and their specializations, relative specializations and diversities

```r
cu.tab<-data.frame("Cultures"=ob.sim$cu,"Specialization"=ob.sim$specializations,
                   "Relative.specialization"=ob.sim$r.specializations,
                   "S.Diversity"=ob.sim$s.diversity)
kable(cu.tab)
```

Cultures	Specialization	Relative.specialization	S.Diversity
c1	0.6534922	0.3267461	1.8524724
c2	1.8513755	0.9256877	0.4138169
c3	0.6642951	0.3321476	1.9131163
c4	0.5750163	0.2875081	0.4182366

8. Extract mutual information, biocultural complexity and relative biocultural complexity.

```r
#Mutual information
ob.sim$mutualInfo
```

```
## [1] 0.9360448
```

```r
#Biocultural complexity
ob.sim$BC
```

```
## [1] 1.913276
```

```r
#Relative biocultural complexity
ob.sim$r.BC
```

```
## [1] 0.4783189
```

9. Perform a chi square test for association between cultures and species. A montecarlo simulation is used to get the p value>

```r
chisq.test(dat.sim[-c(1,2)],simulate.p.value=T)
```

```
##
##  Pearson's Chi-squared test with simulated p-value (based on 2000
##  replicates)
##
## data:  dat.sim[-c(1, 2)]
```

```
## X-squared = 547.36, df = NA, p-value = 0.0004998
```