Spatial Inequalities with R

Stamatis Kalogirou

2015-07-05

Introduction

This is a quick presentation for calculating the spatial decomposition of the Gini coefficient (Rey and Smith, 2013) in R with the package `lctools`. For more details about lctools see http://cran.r-project.org/web/packages/lctools/index.html. Please make sure that the R packages lctools and rgdal have been installed.

The dataset GR.Municipalities that comes with the package lctoos is a SpatialPolygonsDataFrame and refers to Municipalities of Greece. The descriptive data for each municipality include demographic and economic variables the source of which are the Hellenic Population Census of 2001 and the General Secretariat for Information Systems of the Hellenic Republic, respectively.

```
library(lctools)
data(GR.Municipalities)
names(GR.Municipalities@data)
```

```
##  [1] "OBJECTID"   "X"          "Y"          "Name"       "CodeELSTAT"
##  [6] "PopM01"     "PopF01"     "PopTot01"   "UnemrM01"   "UnemrF01"
## [11] "UnemrT01"   "PrSect01"   "Foreig01"   "Income01"
```

Analysis

Out of the above 14 variables, *X,Y* are the geographical coordinates of the municipalities' centroids that will help calculating the distance matrix and then the weights matrix based on a fixed number of nearest neighbours; and *Income01* is the variable to be analysed and refers to the mean annual recorded income in 2001 (in Euros)

In order to calculate the spatial Gini coefficient for the above variable, it is necessary to prepare the input for the function spGini: to set the number of nearest neighbours, the variable for which the coefficient will be calculated and the x,y coordinates of the data. If present, one should check for and remove missing values as this is not yet controlled for within the function spGini. The code below is an example where the spatial Gini coefficient is calculated for the variable *Income01* with 12 neighbours and binary weights (this is 1 for neighbours and 0 for non-neighbours).

```
myDF<-cbind(GR.Municipalities@data$Income01,GR.Municipalities@data$X, GR.Municipalities@data$Y)
myDF[!complete.cases(myDF),]
```

```
##      [,1] [,2] [,3]
```

```
myDF.new<-na.omit(myDF)
bw<-12
wt<-'Binary'
spGini(myDF.new[,2:3],bw,myDF.new[,1],wt)
```

```
##        Gini      gwGini      nsGini gwGini.frac nsGini.frac
## 0.145501498 0.003420536 0.142080961 0.023508599 0.976491401
```

The overall Gini coefficient suggests that there is inequality among municipalities in terms of the mean annual recorded income. The first term of the spatially disaggregated Gini suggests that neighbour municipalities have rather similar level of income, whereas the second term suggests that most of the inequality is due to non-neighbours inequality. To test if this finding is significant, one can run a Monte Carlo simulation:

```
spGini.Sim20<-mc.spGini(Nsim=19,bw,myDF.new[,1],myDF.new[,2],myDF.new[,3],wt)
spGini.Sim20$pseudo.p
```

```
## [1] 0.05
```

Since the pseudo p-value is 0.05 for 19 permutations, it can be argued that the terms of the spatial Gini coefficient are significant at the 95% level. This level of significance can be increased if more permutations are tested (eg 99, 999 etc.).

References

Rey, S.J., Smith, R.J. , 2013, A spatial decomposition of the Gini coefficient, Letters in Spatial and Resource Sciences, 6 (2), pp. 55-70.