

# Package ‘SpatialPack’

February 19, 2015

**Version** 0.2-3

**Date** 2014-08-07

**Title** Package for analysis of spatial data

**Author** Felipe Osorio <felipe.osorios@usm.cl> and Ronny Vallejos  
<ronny.vallejos@usm.cl>, with contributions of Francisco Cuevas  
and Diego Mancilla.

**Maintainer** Felipe Osorio <felipe.osorios@usm.cl>

**Suggests** geoR

**Description** This package provides tools to assess the association between two spatial processes. Currently, three methodologies are implemented: An adapted t-test to perform hypothesis testing about the independence between the processes, a suitable nonparametric correlation coefficient, and the codispersion coefficient. SpatialPack gives methods to complement methodologies that are available in geoR for one spatial process.

**License** GPL (>= 2)

**URL** <http://spatialpack.mat.utfsm.cl>

**LazyLoad** yes

**LazyData** yes

**NeedsCompilation** yes

**Repository** CRAN

**Date/Publication** 2014-08-08 08:04:01

## R topics documented:

codisp . . . . .	2
cor.spatial . . . . .	3
modified.ttest . . . . .	4
murray . . . . .	6

<b>Index</b>	<b>7</b>
--------------	----------

---

codisp *Codispersion Coefficient*

---

### Description

Computes the codispersion coefficient between two spatial variables for a given number of classes for the lag distance.

### Usage

```
codisp(x, y, coords, nclass = 13)
```

### Arguments

x	an n-dimensional vector of data values.
y	an n-dimensional vector of data values.
coords	an n x 2 matrix containing coordinates of the n data locations in each row.
nclass	a single number giving the number of cells for the codispersion coefficient. The default is 13. If this argument is NULL Sturges' formula is used.

### Details

The procedure computes the codispersion coefficient for two spatial sequences defined on general (non-rectangular) grids. First, a given number of bins are constructed for the lag distance. Then the codispersion is computed for each bin.

### Value

A list with class "codisp" containing the following components:

coef	a vector of size nclass containing the values of the codispersion coefficient.
upper.bounds	upper bounds of the intervals constructed to compute the codispersion coefficient.
card	number of elements in each interval generated to compute the codispersion coefficient.

The function plot can be used to obtain a graph of the codispersion coefficient versus the lag distance.

### References

- Matheron, G. (1965), *Les Variables Regionalisees et leur Estimation*. Masson, Paris.
- Rukhin, A., Vallejos, R. (2008), Codispersion coefficient for spatial and temporal series. *Statistics and Probability Letters* **78**, 1290–1300.
- Vallejos, R. (2008). Assessing the association between two spatial or temporal sequences. *Journal of Applied Statistics* **35**, 1323–1343.

## Examples

```
## This example uses the camg data set, available in the geoR Package
## (Ribeiro and Diggle, 2001).

## required to loading the camg data set from geoR package
library(geoR)
data(camg)

## defining the variables elevation and ca020 from the camg data set
x <- camg$elevation
y <- camg$ca020

## extracting the coordinates from the camg data set
coords <- camg[,1:2]

## computing the codispersion coefficient
z <- codisp(x, y, coords)
z

## plotting the codispersion coefficient vs. the lag distance
plot(z)
```

---

cor.spatial

*Tjostheim's Coefficient*

---

## Description

Computes Tjostheim's coefficient for two spatial sequences observed over the same locations on the plane.

## Usage

```
cor.spatial(x, y, coords)
```

## Arguments

x	an n-dimensional vector of data values.
y	an n-dimensional vector of data values.
coords	an n x 2 matrix containing coordinates of the n data locations in each row.

## Details

The implemented technique is a nonparametric coefficient that summarizes the association between two spatial variables. This coefficient was first introduced by Tjostheim (1978) and later generalized by Hubert and Golledge (1992). The computation of the coefficient is based on the construction of ranks associated to suitable modifications of the coordinates. Tjostheim's coefficient is a variant of the correlation coefficient (`cor`) to be used in a spatial statistics context.

**Value**

Tjostheim's coefficient. The variance is returned as the attribute "variance".

**References**

Tjostheim, D., (1978), A measure of association for spatial variables. *Biometrika* **65**, 109–114.

Hubert, L., Golledge, R.G., (1982), Measuring association between spatially defined variables: Tjostheim's coefficient index and some extensions. *Geographical Analysis* **14**, 273–278.

**Examples**

```
## This example uses the camg data set, available in the geoR Package
## (Ribeiro and Diggle, 2001).

## required to loading the camg data set from geoR package
library(geoR)
data(camg)

## defining the variables elevation and ca020 from the camg data set
x <- camg$elevation
y <- camg$ca020

## extracting the coordinates from the camg data set
coords <- camg[,1:2]

## computing Tjostheim's coefficient
cor.spatial(x, y, coords)
```

---

modified.ttest

*Modified t test*

---

**Description**

Performs a modified version of the  $t$  test to assess the correlation between two spatial processes.

**Usage**

```
modified.ttest(x, y, coords, nclass = 13)
```

**Arguments**

x	an n-dimensional vector of data values.
y	an n-dimensional vector of data values.
coords	an n x 2 matrix containing coordinates of the n data locations in each row.
nclass	a single number giving the number of cells for Moran's index. The default is 13. If this argument is NULL Sturges' formula is used.

## Details

The methodology implemented is a modified  $t$  test of spatial association based on the work of Clifford and Richardson (1989). The test is based on corrections of the sample correlation coefficient between the two spatially correlated sequences and required the estimation of an effective sample size. This factor takes into account the spatial association of both processes.

## Value

A list with class "mod.ttest" containing the following components:

corr	the sample correlation coefficient.
Fstat	the value of the F-statistic.
dof	the estimated degrees of freedom for the F-statistic.
p.value	the p-value for the test.
upper.bounds	upper bounds of the intervals constructed to compute Moran's I.
card	number of elements in each interval generated to compute Moran's I.
imoran	a matrix containing Moran's index for each interval associated to both variables.

The generic functions `print` and `summary` are used to obtain and print additional details about the modified  $t$  test.

## References

- Clifford, P., Richardson, S., Hemon, D. (1989), Assessing the significance of the correlation between two spatial processes. *Biometrics* **45**, 123–134.
- Dutilleul, P. (1993), Modifying the  $t$  test for assessing the correlation between two spatial processes. *Biometrics* **49**, 305–314.

## Examples

```
## This example uses the camg data set, available in the geoR Package
## (Ribeiro and Diggle, 2001).

## required to loading the camg data set from geoR package
library(geoR)
data(camg)

## defining the variables elevation and ca020 from the camg data set
x <- camg$elevation
y <- camg$ca020

## extracting the coordinates from the camg data set
coords <- camg[,1:2]

## computing the modified t-test of spatial association
z <- modified.ttest(x, y, coords)
z

## display the upper bounds, cardinality and the computed Moran's index
summary(z)
```

---

murray

*The Murray smelter site dataset*

---

### Description

The dataset consists of soil samples collected in and around the vacant, industrially contaminated, Murray smelter site (Utah, USA). This area was polluted by airborne emissions and the disposal of waste slag from the smelting process. A total of 253 locations were included in the study, and soil samples were taken from each location. Each georeferenced sample point is a pool composite of four closely adjacent soil samples in which the concentration of the heavy metals arsenic (As) and lead (Pb) was determined.

### Usage

```
data(murray)
```

### Format

A data frame with 253 observations on the following 5 variables.

**As** arsenic concentrations measurements.

**Pb** lead concentrations measurements.

**xpos** x coordinates.

**ypos** y coordinates.

**quad** a factor where numbers indicate different sub-regions within the area.

### Source

Griffith, D., Paelinck, J.H.P. (2011). *Non-Standard Spatial Statistics*. Springer, New York.

# Index

\*Topic **datasets**

murray, 6

\*Topic **htest**

modified.ttest, 4

\*Topic **multivariate**

codisp, 2

cor.spatial, 3

codisp, 2

cor, 3

cor.spatial, 3

modified.ttest, 4

murray, 6