Package 'sandwich'

August 19, 2021

Version 0.1.0

Date 2021-06-01

Title sandwich: Spatial Interpolation Based on Spatial Stratified Heterogeneity

Author Yue Lin [aut, cre],
Chengdong Xu [aut, ctb],
Jinfeng Wang [aut, ctb]

Maintainer Yue Lin <lin.3326@osu.edu>

Depends R (>= 3.5.0), sf, ggplot2, gridExtra, geodetector, dplyr, caret, lwgeom

Suggests rmarkdown,
knitr,
MASS

Description This package enables the implementation of Sandwich model-based mapping in R.
It provides functions to output the summarized statistics, the interpolated surface,
and the standard error and confidence interval maps.

License GPL (>= 2)

URL http://www.sssampling.cn/, https:
//github.com/linyuehzzz/sandwich_spatial_interpolator

BugReports https://github.com/linyuehzzz/sandwich_spatial_interpolator/issues

LazyData false

RoxygenNote 7.1.1

VignetteBuilder knitr

Encoding UTF-8

R topics documented:

Index   **10**

---

bc   *Breast cancer incidence in mainland China*

Description

The bc dataset consists of three sf objects:

- bc.sampling: Breast cancer incidence at 271 sampling units in mainland China, which serves as the sampling layer.

- bc.ssh: A classification map with 20 strata that serves as a candidate SSH layer.

- bc.reporting: The administrative division of 32 provices in mainland China that served as the reporting layer.

Examples

```
library(sf)
library(ggplot2)
data("bc.sampling")
summary(bc.sampling$Incidence)
ggplot(data=bc.sampling) + geom_sf(aes(color=Incidence))
```

---

hs   *Heshun human population dataset*

Description

The hs dataset consists of four sf objects:

- hs.sampling: Human population of 326 sampling villages within the Heshun region, China, which serves as the sampling layer.

- hs.ssh: A zonation map with 5 strata that serves as a candidate SSH layer.

- hs.ssh2: A zonation map with 6 strata that serves as another candidate SSH layer.

- hs.reporting: The administrative division of 10 towns in Heshun county that served as the reporting layer.

References

Li, L., Wang, J., & Wu, J. (2012). A spatial model to predict the incidence of neural tube defects. *BMC Public Health*, 12(1), 951. doi:10.1186/1471-2458-12-951

Examples

```
library(sf)
library(ggplot2)
data("hs.sampling")
summary(hs.sampling$Population)
ggplot(data=hs.sampling) + geom_sf(aes(color=Population))
```

jl	*Jilin rodent density dataset*

Description

The `jl` dataset consists of four `sf` objects:

- `jl.sampling`: Rodent density at 627 sampling quadrats in Jilin, China, which serves as the sampling layer.
- `jl.ssh`: A classification map with 8 strata that serves as a candidate SSH layer.
- `jl.ssh2`: A zonation map with 11 strata that serves as another candidate SSH layer.
- `jl.reporting`: The administrative division of 10 counties in Jilin that served as the reporting layer.

References

Liu, T., Wang, J., Xu, C., Ma, J., Zhang, H., & Xu, C. (2018). Sandwich mapping of rodent density in Jilin Province, China. *Journal of Geographical Sciences*, 28(4), 445-458.. doi:10.1007/s11442-018-1483-z

Examples

```
library(sf)
library(ggplot2)
data("jl.sampling")
summary(jl.sampling$Shushu)
ggplot(data=jl.sampling) + geom_sf(aes(color=Shushu))
```

load.data.shp	*Load shapefiles into* sf*s*

Description

`load.data.shp` converts shapefiles into a list of `sf` objects.

Usage

```
load.data.shp(path, sampling.file, ssh.file, reporting.file)
```

Arguments

`path`	Directory of the shapefile.
`sampling.file`	Name of the sampling layer, minus extension (.shp). The sampling layer should be a point feature that contains an attribute that is to be interpolated.
`ssh.file`	Name of the SSH layer, minus extension (.shp). The SSH layer should be a polygon feature.
`reporting.file`	Name of the reporting layer, minus extension (.shp). The reporting llayer should be a polygon feature.

`load.data.txt`   *Load text files into data frames*

Description

`load.data.txt` converts text files into a list of data frames.

Usage

```
load.data.txt(path, sampling_ssh.file, reporting_ssh.file)
```

Arguments

`sampling_ssh.file`	
	Name of the file linking sampling and SSH layers. Each row denotes a sampling unit. Two columns need to be included in this file: value of the sampling attribute and the stratum each sampling unit falling into.
`reporting_ssh.file`	
	Name of the file linking reporting and SSH layers. Each row denotes a reporting unit. The weight W_{rh} should be specified in each column.
`path`	Directory of the shapefile.

`plot.ci`   *Plot confidence interval of the Sandwich estimates*

Description

`plot.ci` plots the lower and upper bounds of the confidence interval for the mean value of each reporting unit.

Usage

```
plot.ci(object)
```

Arguments

`object`	An `sf` object generated from the `sandwich.ci` function.

---

plot.mean	*Plot Sandwich estimates*

Description

`plot.mean` plots the mean value of an attribute for each reporting unit estimated by the Sandwich model.

Usage

```
plot.mean(object)
```

Arguments

object   An `sf` object generated from the `sandwich.model` function.

References

Wang, J. F., Haining, R., Liu, T. J., Li, L. F., & Jiang, C. S. (2013). Sandwich estimation for multi-unit reporting on a stratified heterogeneous surface. *Environment and Planning A*, 45(10), 2515-2534. doi:10.1068/a44710

---

plot.se	*Plot standard error of the Sandwich estimates*

Description

`plot.se` plots the standard error of estimated mean value for each reporting unit.

Usage

```
plot.se(object)
```

Arguments

object   An `sf` object generated from the `sandwich.model` function.

References

Wang, J. F., Haining, R., Liu, T. J., Li, L. F., & Jiang, C. S. (2013). Sandwich estimation for multi-unit reporting on a stratified heterogeneous surface. *Environment and Planning A*, 45(10), 2515-2534. doi:10.1068/a44710

---

sandwich   sandwich: *Spatial Interpolation Based on Spatial Stratified Hetero-*
geneity

Description

The package sandwich provides tools to interpolate a spatially stratified heterogeneous population with high accuracy, even when its spatial autocorrelation is weak. These tools allow the calculation of critical summarized statistics such as Geodetector q. They also enable the visualization of the interpolated surface as well as the standard error and confidence interval maps.

Author(s)

Yue Lin, Chengdong Xu, and Jinfeng Wang

References

Wang, J. F., Haining, R., Liu, T. J., Li, L. F., & Jiang, C. S. (2013). Sandwich estimation for multi-unit reporting on a stratified heterogeneous surface. *Environment and Planning A*, 45(10), 2515-2534. doi:10.1068/a44710

---

sandwich.ci   *Calculate confidence interval of the Sandwich estimates*

Description

sandwich.ci calculates the confidence interval of estimated mean value for each reporting unit.

Usage

```
sandwich.ci(object, level=.95)
```

Arguments

object	An sf object generated from the sandwich.model function.
level	The confidence level required. By default, level = .95.

---

sandwich.cv	*Perform* k-*fold cross validation*

Description

sandwich.cv perfoms *k*-fold cross validation to evaluate the overall model accuracy and outputs the average root mean square error (RMSE).

Usage

```
sandwich.cv(object, sampling.attr, k=10, type="shp", ssh.id.col=NULL, ssh.weights=NULL)
```

Arguments

object
: When type="shp", object is a list of 3 sf objects generated by load.data.shp, including a POINT sf object used as the sampling layer, a POLYGON sf object used as the SSH layer, and a POLYGON sf object used as the SSH layer. When type="txt", object is a list of 2 data frames generated by load.data.txt, including a file linking sampling and SSH layers linking reporting and SSH layers and a file.

sampling.attr
: A string denoting the name of the attribute in the sampling layer to be interpolated.

k
: The number of folds (k > 1). By default, k = 10.

type
: A string denoting the type of input data. type="shp" denotes shapefiles, and type="txt" denotes text files. By default, type="shp".

ssh.id.col
: A string denoting the column that specifies which stratum each sampling unit falls into in the file linking sampling and SSH layers.

reporting.id.col
: A string denoting the column that specifies which reporting unit each sampling unit falls into in the file linking sampling and SSH layers.

ssh.weights
: A list that specifies the strata in the SSH layer and and their corresponding columns of weights in the file linking reporting and SSH layers.

---

sandwich.model	*Perform Sandwich model-based mapping*

Description

sandwich.model estimate the mean value of an attribute and its standard error for each reporting unit using the Sandwich method documented by Wang et al. (2013).

Usage

```
sandwich.model(object, sampling.attr, type="shp", ssh.id.col=NULL, ssh.weights=NULL)
```

Arguments

object	When `type="shp"`, `object` is a list of 3 `sf` objects generated by `load.data.shp`, including a POINT `sf` object used as the sampling layer, a POLYGON `sf` object used as the SSH layer, and a POLYGON `sf` object used as the SSH layer. When `type="txt"`, `object` is a list of 2 data frames generated by `load.data.txt`, including a file linking sampling and SSH layers linking reporting and SSH layers and a file.
sampling.attr	A `string` denoting the name of the attribute in the sampling layer to be interpolated.
type	A `string` denoting the type of input data. `type="shp"` denotes shapefiles, and `type="txt"` denotes text files. By default, `type="shp"`.
ssh.id.col	A `string` denoting the column that specifies which stratum each sampling unit falls into in the file linking sampling and SSH layers.
ssh.weights	A `list` that specifies the strata in the SSH layer and and their corresponding columns of weights in the file linking reporting and SSH layers.

References

Wang, J. F., Haining, R., Liu, T. J., Li, L. F., & Jiang, C. S. (2013). Sandwich estimation for multi-unit reporting on a stratified heterogeneous surface. *Environment and Planning A*, 45(10), 2515-2534. doi:10.1068/a44710

---

ssh.data.shp	*Prepare data for Geodetector*

Description

`ssh.data.shp` joins a candidate SSH layer to an existing layer.

Usage

```
ssh.data.shp(object, ssh.lyr, ssh.id, ssh.name=NULL)
```

Arguments

object	A POINT `sf` object to be joint. Its geometry should be consistent with the sampling layer, but it may have additional attributes from the SSH layer(s).
ssh.lyr	A POLYGON `sf` object used as the candidate SSH layer. Its strata ID will be linked to `object`.
ssh.id	A `string` denoting the attribute name of strata ID in the SSH layer.
ssh.name	A `string` denoting the new attribute name assigned to the strata ID in the output object. If `NULL`, the attribute name will be the same as that in the old SSH layer.

See Also

geodetector R package

ssh.test	*Calculate Geodetector factor detector* q-*statistic and interaction detector*

Description

`ssh.test` calculates the factor detector *q*-statistic and the interaction detector in the geographical detector model described by Wang et al. (2010). The *q*-statistic measures the SSH of the sampling attribute in terms of a given stratification, which can be used for the selection of an SSH layer for Sandwich model-based mapping. The interactive effects indicate whether a combination of two stratifications enhances the SSH of the sampling attribute.

Usage

```
ssh.test(object, y, x, test="factor")
```

Arguments

object	A POINT `sf` object linking the attributes of the sampling layer and the strata ID in the SSH layer(s) generated from `ssh.data`.
y	A `string` denoting the name of the explained variable (sampling attribute) in `object`.
x	A `string` denoting the name(s) of the explanatory variable(s) (stratification(s)) in `object`.
test	A `string` denoting the type of test. `test="factor"` denotes the factor detector, and `test="interaction"` denotes the interaction detector. By default, `test="factor"`.

References

Wang, J. F., Li, X. H., Christakos, G., Liao, Y. L., Zhang, T., Gu, X., & Zheng, X. Y. (2010). Geographical detectors-based health risk assessment and its application in the neural tube defects study of the Heshun Region, China. *International Journal of Geographical Information Science*, 24(1), 107-127. doi:10.1080/13658810802443457

See Also

geodetector R package

Index