

# Package ‘smacpod’

September 23, 2015

**Type** Package

**Title** Statistical Methods for the Analysis of Case-Control Point Data

**Version** 1.4.1

**Date** 2015-09-22

**Author** Joshua French

**Maintainer** Joshua French <joshua.french@ucdenver.edu>

**Description** Various statistical methods for analyzing case-control point data.  
The methods available closely follow those in chapter 6 of Applied Spatial  
Statistics for Public Health Data by Waller and Gotway (2004).

**License** GPL (>= 2)

**LazyLoad** yes

**Depends** R (>= 3.1.1)

**Imports** spatstat, SpatialTools, plotrix, parallel

**Suggests** testthat

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2015-09-23 03:07:51

## R topics documented:

circles.intersect . . . . .	2
grave . . . . .	3
kd . . . . .	3
kdest . . . . .	5
kdplus.test . . . . .	6
logrr . . . . .	7
logrr.test . . . . .	9
nn . . . . .	10
noc . . . . .	11
plot.kdenv . . . . .	12
plot.logrrenv . . . . .	12

plot.scan . . . . .	13
qnn.test . . . . .	14
spdensity . . . . .	15
spscan.test . . . . .	16

<b>Index</b>	<b>19</b>
--------------	-----------

---

circles.intersect	<i>Determine whether circles intersection</i>
-------------------	---

---

## Description

circles.intersect determines whether circles intersect with each other.

## Usage

```
circles.intersect(coords, r)
```

## Arguments

coords	A matrix of coordinates contained the centroid of each circle.
r	A vector containing the radii of the circles. The have length equal to the number of rows of coords.

## Details

The algorithm is based on the premise that two circles intersect if, and only if, the distance between their centroids is between the sum and the difference of their radii. I have squared the respective parts of the inequality in the implemented algorithm.

## Value

Returns a matrix of logical values indicating whether the circles intersect.

## Author(s)

Joshua French

## Examples

```
# first two intersect, then next two, last doesn't intersect any other
co = cbind(c(1, 2, 5, 6, 9), c(1, 2, 5, 6, 9))
r = c(1.25, 1.25, 1.25, 1.25, 1.25)
circles.intersect(co, r)
# nested circles
co = matrix(rep(0, 4), nrow = 2)
r = c(1, 1.5)
circles.intersect(co, r)
```

---

grave

*Medieval Grave Site Data*

---

### **Description**

This data set contains 143 observations of medieval grave site data stored as a ppp object from the spatstat package. The data are marked as being "affected" by a tooth deformity or "unaffected" by a tooth deformity.

### **Usage**

grave

### **Format**

ppp (planar point process)

### **Author(s)**

Joshua French

### **Source**

Waller, L.A. and Gotway, C.A. (2005). Applied Spatial Statistics for Public Health Data. Hoboken, NJ: Wiley.

---

kd

*Difference of estimated K functions*

---

### **Description**

kd determines the difference in estimated K functions for a set of cases and controls.

### **Usage**

```
kd(x, case = 2, r = NULL, breaks = NULL, correction = c("border",  
  "isotropic", "Ripley", "translate"), nlarge = 3000, domain = NULL,  
  var.approx = FALSE, ratio = FALSE)
```

**Arguments**

x	A ppp object from the spatstat package with marks for the case and control groups.
case	The position of the name of the "case" group in levels(x\$marks). The default is 2.
r	Optional. Vector of values for the argument r at which $K(r)$ should be evaluated. Users are advised not to specify this argument; there is a sensible default.
breaks	This argument is for internal use only.
correction	Optional. A character vector containing any selection of the options "none", "border", "bord.modif", "isotropic", "Ripley", "translate", "translation", "none", "good" or "best". It specifies the edge correction(s) to be applied.
nlarge	Optional. Efficiency threshold. If the number of points exceeds nlarge, then only the border correction will be computed (by default), using a fast algorithm.
domain	Optional. Calculations will be restricted to this subset of the window. See Details.
var.approx	Logical. If TRUE, the approximate variance of $Kest(r)$ under CSR will also be computed.
ratio	Logical. If TRUE, the numerator and denominator of each edge-corrected estimate will also be saved, for use in analysing replicated point patterns.

**Value**

Returns an fv object. See documentation for `spatstat::Kest`.

**Author(s)**

Joshua French

**References**

Waller, L.A. and Gotway, C.A. (2005). Applied Spatial Statistics for Public Health Data. Hoboken, NJ: Wiley. Kulldorff, M. (1997) A spatial scan statistic. Communications in Statistics – Theory and Methods 26, 1481-1496.

**Examples**

```
data(grave)
# kd = kd(grave)
# plot(kd)
```

---

kdest	<i>Difference of estimated K functions</i>
-------	--

---

### Description

kdest determines the difference in estimated K functions for a set of cases and controls.

### Usage

```
kdest(x, case = 2, nsim = 0, level = 0.95, r = NULL, breaks = NULL,
      correction = c("border", "isotropic", "Ripley", "translate"),
      nlarge = 3000, domain = NULL, var.approx = FALSE, ratio = FALSE)
```

### Arguments

x	A ppp object from the spatstat package with marks for the case and control groups.
case	The position of the name of the "case" group in levels(x\$marks). The default is 2.
nsim	An non-negative integer. Default is 0. The difference in estimated K functions will be calculated for nsim data sets generated under the random labeling hypothesis.
level	Confidence level of confidence envelopes. Ignored if nsim is 0.
r	Optional. Vector of values for the argument r at which K(r) should be evaluated. Users are advised not to specify this argument; there is a sensible default.
breaks	This argument is for internal use only.
correction	Optional. A character vector containing any selection of the options "none", "border", "bord.modif", "isotropic", "Ripley", "translate", "translation", "none", "good" or "best". It specifies the edge correction(s) to be applied.
nlarge	Optional. Efficiency threshold. If the number of points exceeds nlarge, then only the border correction will be computed (by default), using a fast algorithm.
domain	Optional. Calculations will be restricted to this subset of the window. See Details.
var.approx	Logical. If TRUE, the approximate variance of Kest(r) under CSR will also be computed.
ratio	Logical. If TRUE, the numerator and denominator of each edge-corrected estimate will also be saved, for use in analysing replicated point patterns.

### Details

This function relies internally on the Kest and eval.fv functions from the spatstat package. So the arguments are essentially the same as the Kest function. See the documentation of the Kdest for more details about the various arguments.

**Value**

Returns an fv object. See documentation for `spatstat::Kest`.

**Author(s)**

Joshua French

**References**

Waller, L.A. and Gotway, C.A. (2005). *Applied Spatial Statistics for Public Health Data*. Hoboken, NJ: Wiley. Kulldorff, M. (1997) A spatial scan statistic. *Communications in Statistics – Theory and Methods* 26, 1481-1496.

**See Also**

[Kest](#)

**Examples**

```
data(grave)
kd1 = kdest(grave)
plot(kd1, iso ~ r, ylab = "difference", legend = FALSE, main = "")
kd2 = kdest(grave, nsim = 9, level = 0.8)
plot(kd2)
```

---

kdplus.test

*Global test of clustering using difference in K functions*

---

**Description**

`kdplus.test` performs a global test of clustering for comparing cases and controls using the method of Diggle and Chetwynd (1991). It relies on the difference in estimated K functions.

**Usage**

```
kdplus.test(x)
```

**Arguments**

`x` A `kdenv` object from the `kdest` function.

**Value**

A list providing the observed test statistic (`kdplus`) and the estimate p-value `pvalue`.

**Author(s)**

Joshua French

## References

Waller, L.A. and Gotway, C.A. (2005). Applied Spatial Statistics for Public Health Data. Hoboken, NJ: Wiley. Diggle, Peter J., and Amanda G. Chetwynd. "Second-order analysis of spatial clustering for inhomogeneous populations." Biometrics (1991): 1155-1163.

## See Also

[kdest](#)

## Examples

```
data(grave)
kdsim = kdest(grave, nsim = 9)
kdplus.test(kdsim)
```

---

logrr

*Log ratio of spatial densities*

---

## Description

logrr computes envelopes for the log ratio of spatial density functions. The numerator in this ratio is related to the "cases" and the denominator to the "controls".

## Usage

```
logrr(x, sigma = NULL, sigmacon = NULL, case = 2, nsim = 0,
      level = 0.9, alternative = "two.sided", ..., bwargs = list(),
      weights = NULL, edge = TRUE, varcov = NULL, at = "pixels",
      leaveoneout = TRUE, adjust = 1, diggle = FALSE, nreport = 50)
```

## Arguments

x	Point pattern (object of class "ppp").
sigma	Standard deviation of isotropic Gaussian smoothing kernel for cases. Either a numerical value, or a function that computes an appropriate value of sigma.
sigmacon	Standard deviation of isotropic Gaussian smoothing kernel for controls. Default is the same as sigma.
case	The position of the name of the "case" group in levels(x\$marks). The default is 2.
nsim	The number of simulated data sets from which to construct the envelopes under the random labeling hypothesis. Default is 0.
level	The level for Monte Carlo test using tolerance intervals.
alternative	The direction of the test. Default is "two.sided". The values "less" and "greater" are also valid.

...	Additional arguments passed to <code>spatstat::pixellate.ppp</code> or <code>spatstat::as.mask</code> to determine the pixel resolution.
<code>bwargs</code>	A list of arguments for the bandwidth function supplied to <code>sigma</code> , if applicable.
<code>weights</code>	Optional weights to be attached to the points. A numeric vector, numeric matrix, or an expression.
<code>edge</code>	Logical flag: if TRUE, apply edge correction.
<code>varcov</code>	Variance-covariance matrix of anisotropic Gaussian kernel. Incompatible with <code>sigma</code> .
<code>at</code>	String specifying whether to compute the intensity values at a grid of pixel locations ( <code>at="pixels"</code> ) or only at the points of <code>x</code> ( <code>at="points"</code> ).
<code>leaveoneout</code>	Logical value indicating whether to compute a leave-one-out estimator. Applicable only when <code>at="points"</code> .
<code>adjust</code>	Optional. Adjustment factor for the smoothing parameter.
<code>diggle</code>	Logical. If TRUE, use Diggle's edge correction, which is more accurate but slower to compute than the correction described under Details.
<code>nreport</code>	How frequently to report progress on the simulation. Default is 50.

### Details

The `two.sided` alternative test assesses whether the observed ratio of log densities deviates more what is expected under the random labelling hypothesis. When the test is significant, this suggests that the cases and controls are clustered, relative to the either. The `greater` alternative test assesses whether the cases are more clustered than the controls. The `less` alternative test assesses whether the controls are more clustered than the cases. If the estimated density of the case or control group becomes too small, this function may produce warnings due to numerical underflow. Increasing the bandwidth (`sigma`) may help.

### Value

The function produces an object of type `logrrenv`. Its components are similar to those returned by the `density.ppp` function from the `spatstat` package, with the intensity values replaced by the log ratio of spatial densities of `f` and `g`. Includes an array `simr` of dimension `c(nx, ny, nsim + 1)`, where `nx` and `ny` are the number of `x` and `y` grid points used to estimate the spatial density. `simr[, , 1]` is the log ratio of spatial densities for the observed data, and the remaining `nsim` elements in the third dimension of the array are the log ratios of spatial densities from a new `ppp` simulated under the random labeling hypothesis.

### Author(s)

Joshua French (and a small chunk by the authors of the )

### References

Waller, L.A. and Gotway, C.A. (2005). Applied Spatial Statistics for Public Health Data. Hoboken, NJ: Wiley. Kelsall, Julia E., and Peter J. Diggle. "Kernel estimation of relative risk." *Bernoulli* (1995): 3-16. Kelsall, Julia E., and Peter J. Diggle. "Non-parametric estimation of spatial variation in relative risk." *Statistics in Medicine* 14.21-22 (1995): 2335-2342.



**Examples**

```
data(grave)
r = logrr(grave)
plot(r)
r2 = logrr(grave, sigma = spatstat::bw.scott)
plot(r2)
rsim = logrr(grave, nsim = 9)
plot(rsim)
```

---

`logrr.test`*Global test of clustering using log ratio of spatial densities*

---

**Description**

`logrr.test` performs a global test of clustering for comparing cases and controls using the log ratio of spatial densities the method of Kelsall and Diggle (1995).

**Usage**

```
logrr.test(x)
```

**Arguments**

`x` An `renv` object from the `logrr.env` function.

**Value**

A list providing the observed test statistic (`islogrr`) and the estimated p-value `pvalue`.

**Author(s)**

Joshua French

**References**

Waller, L.A. and Gotway, C.A. (2005). Applied Spatial Statistics for Public Health Data. Hoboken, NJ: Wiley. Kelsall, Julia E., and Peter J. Diggle. "Non-parametric estimation of spatial variation in relative risk." *Statistics in Medicine* 14.21-22 (1995): 2335-2342.

**Examples**

```
data(grave)
renv = logrr(grave, nsim = 9)
logrr.test(renv)
```

---

nn *Determine nearest neighbors*

---

### Description

nn determines the nearest neighbors for a set of observations based on the distance matrix.

### Usage

```
nn(d, k, method = "c", self = FALSE)
```

### Arguments

d	A square distance matrix for the set of coordinates of interest.
k	The number of numbers to return (if method = "c") or the distance for which observations are considered neighbors (if method = "d").
method	The method of determining the neighbors. The default is "c", specifying that the k nearest neighbors (the count of neighbors) for each observation should be returned. The alternative is "d", meaning that neighbors are determined by their distance from an observation. In that case, two observations are neighbors if their separation distance is less or equal to k.
self	A logical indicating whether an observation is a neighbor with itself. The default is FALSE.

### Details

This function can determine nearest neighbors in two ways: 1. by total count or 2. by distance. If method = "c", then k determines the total number of neighbors to return for each observation. If method = "d", then k determines the distance for which an observation is considered a neighbor.

### Value

Returns the indexes of the nearest neighbors as a matrix if method = "c" and a list otherwise. For each row or element of the list, the indexes are ordered from nearest to farthest.

### Author(s)

Joshua French

### Examples

```
data(grave)
# make distance matrix
d = as.matrix(dist(cbind(grave$x, grave$y)))
# 3 nearest neighbors
nnc = nn(d, k = 3, method = "c")
# nearest neighbors within k units of each observation
nnd = nn(d, k = 200, method = "d")
```

---

`noc`*Determine non-overlapping clusters*

---

**Description**

Determine the indexes of the non-overlapping clusters

**Usage**

```
noc(x)
```

**Arguments**

`x` A list containing the indexes of the potential clusters.

**Details**

The functions takes a list of potential cluster indexes. Each element of the list contains a potential cluster. The potential clusters are defined by the location indexes of the regions comprising the clusters. Starting with the first potential cluster, the function excludes every potential cluster that intersects the first (look at the location indexes comprising each cluster). Moving onto the next non-overlapping cluster, the process is repeated. The function returns the indexes (in the list of clusters) that do not overlap.

**Value**

A vector with the list indexes of the non-overlapping clusters.

**Author(s)**

Joshua French

**Examples**

```
x = list(1:2, 1:3, 4:5, 4:6, 7:8)
noc(x)
```

---

plot.kdenv	<i>Plots object from kdest of class kdenv.</i>
------------	--

---

**Description**

Plots object from kdest of class kdenv.

**Usage**

```
## S3 method for class 'kdenv'
plot(x, ..., shadecol1 = "grey", shadecol2 = "lightblue",
      main = "", legend = FALSE)
```

**Arguments**

x	An object of class kdenv to be plotted.
...	Additional graphical parameters passed to plot.fv function.
shadecol1	Shade color for max/min envelopes.
shadecol2	Shade color for confidence envelopes.
main	A main title for the plot. Default is blank.
legend	Logical for whether a legend should automatically be displayed. Default if FALSE.

**See Also**

[plot.fv](#)

**Examples**

```
data(grave)
kd1 = kdest(grave, nsim = 19, level = 0.9)
plot(kd1)
```

---

plot.logrrenv	<i>Plots object from logrr of class logrrenv.</i>
---------------	---

---

**Description**

Plots object from logrr of class logrrenv.

**Usage**

```
## S3 method for class 'logrrenv'
plot(x, ..., conlist = list(), main = "")
```

**Arguments**

x	An object of class logrrenv to be plotted.
...	Additional graphical parameters passed to the spatstat::image.im function. See details.
conlist	Additional argument passed to the spatstat::contour.im function.
main	A main title for the plot. Default is blank.

**Details**

An important aspect of this plot is the color argument (col) used for displaying the regions outside the tolerance envelopes. If NULL (the implicit default), then the default color palette used by spatstat::image.im will be used. Simpler schemes, e.g., c("blue", "white", "orange") can suffice. See the examples.

**See Also**

[plot.im](#)  
[contour.im](#)

**Examples**

```
data(grave)
rsim = logrr(grave, nsim = 9)
plot(rsim)
# no border or ribbon (legend). Simple color scheme.
plot(rsim, col = c("blue", "white", "orange"), ribbon = FALSE, box = FALSE)
# alternate color scheme
plot(rsim, col = topo.colors(12))
```

---

plot.scan

*Plots object from spscan.test of class scan.*

---

**Description**

Plots object from spscan.test of class scan.

**Usage**

```
## S3 method for class 'scan'
plot(x, ..., nv = 100, border = NULL, ccol = NULL,
      clty = NULL, clwd = NULL)
```

**Arguments**

x	An object of class scan to be plotted.
...	Additional graphical parameters passed to the spatstat::plot.ppp function.
nv	The number of vertices to draw the circle. Default is 100.
border	The border color of the circle. Default is NULL, meaning black.
ccol	Fill color of the circles. Default is NULL, indicating empty.
clty	Line type of circles. Default is NULL, indicating lty = 1.
clwd	Line width of circles. Default is NULL, indicating lwd = 2 for the most likely cluster and lwd = 1 for the rest.

**See Also**

[plot.ppp](#)

**Examples**

```
data(grave)
out = spscan.test(grave, case = 2, alpha = 0.1)
plot(out, chars = c(1, 20), main = "most likely cluster")
```

---

qnn.test

*q Nearest Neighbors Test*

---

**Description**

qnn.test calculates statistics related to the q nearest neighbors method of comparing case and control point patterns under the random labeling hypothesis.

**Usage**

```
qnn.test(x, q = 5, nsim = 499, case = 2)
```

**Arguments**

x	A ppp object from the spatstat package with marks for the case and control groups.
q	A vector of positive integers indicating the values of q for which to do the q nearest neighbors test.
nsim	The number of simulations from which to compute p-value.
case	The position of the name of the "case" group in levels(x\$marks). The default is 2.

**Value**

Returns a list with the following components:

qsum	A dataframe with the number of neighbors (q), test statistic (Tq), and p-value for each test.
consum	A dataframe with the contrasts (contrast), test statistic (Tcon), and p-value (pvaluecon) for the test of contrasts.

**Author(s)**

Joshua French

**References**

Waller, L.A., and Gotway, C.A. (2005). Applied Spatial Statistics for Public Health Data. Hoboken, NJ: Wiley. Cuzick, J., and Edwards, R. (1990). Spatial clustering for inhomogeneous populations. Journal of the Royal Statistical Society. Series B (Methodological), 73-104. Alt, K.W., and Vach, W. (1991). The reconstruction of "genetic kinship" in prehistoric burial complexes-problems and statistics. Classification, Data Analysis, and Knowledge Organization, 299-310.

**Examples**

```
data(grave)
qnn.test(grave, q = c(3, 5, 7, 9, 11, 13, 15))
```

---

spdensity	<i>Kernel smoothed spatial density of point pattern</i>
-----------	---

---

**Description**

spdensity computes a kernel smoothed spatial density function from a point pattern. This is essentially a slight modification of the density.ppp function from the spatstat package, which computes the spatial intensity of a point pattern.

**Usage**

```
spdensity(x, sigma = NULL, ..., weights = NULL, edge = TRUE,
          varcov = NULL, at = "pixels", leaveoneout = TRUE, adjust = 1,
          diggle = FALSE)
```

**Arguments**

x	Point pattern (object of class "ppp").
sigma	Standard deviation of isotropic Gaussian smoothing kernel. Either a numerical value, or a function that computes an appropriate value of sigma.
...	Additional arguments passed to pixellate.ppp and as.mask to determine the pixel resolution, or passed to sigma if it is a function.

weights	Optional weights to be attached to the points. A numeric vector, numeric matrix, or an expression.
edge	Logical flag: if TRUE, apply edge correction.
varcov	Variance-covariance matrix of anisotropic Gaussian kernel. Incompatible with sigma.
at	String specifying whether to compute the intensity values at a grid of pixel locations (at="pixels") or only at the points of x (at="points").
leaveoneout	Logical value indicating whether to compute a leave-one-out estimator. Applicable only when at="points".
adjust	Optional. Adjustment factor for the smoothing parameter.
diggle	Logical. If TRUE, use Diggle's edge correction, which is more accurate but slower to compute than the correction described under Details.

### Value

This function produces an object of class `im` from the `spatstat` package, in nearly the exact same way as `spatstat::density.ppp`. The difference is that the values are scaled so that a true spatial density function is produced (i.e., the function integrates to 1).

### Author(s)

Joshua French

### References

Waller, L.A. and Gotway, C.A. (2005). *Applied Spatial Statistics for Public Health Data*. Hoboken, NJ: Wiley.

### See Also

[density.ppp](#)

### Examples

```
data(grave)
contour(spdensity(grave))
```

---

spscan.test

*Spatial Scan Test*

---

### Description

`spscan.test` performs the spatial scan test of Kulldorf (1997).



**Usage**

```
spscan.test(x, case = 2, nsim = 499, alpha = 0.1, nreport = nsim + 1,
            maxd = NULL, parallel = TRUE)
```

**Arguments**

x	A ppp object from the spatstat package with marks for the case and control groups.
case	The position of the name of the "case" group in levels(x\$marks). The default is 2.
nsim	The number of simulations from which to compute p-value.
alpha	The significance level to determine whether a cluster is significant.
nreport	The frequency with which to report simulation progress. The default is nsim+ 1, meaning no progress will be displayed.
maxd	The radius of the largest possible cluster to consider.
parallel	A logical indicating whether the test should be parallelized using the parallel:mclapply function. Default is TRUE.

**Details**

The test is performed using the random labeling hypothesis. The windows are circular and extend from the observed data locations. The clusters returned are non-overlapping, ordered from most significant to least significant. The first cluster is the most likely to be a cluster. If no significant clusters are found, then the most likely cluster is returned (along with a warning).

**Value**

Returns a list of length two of class scan. The first element (clusters) is a list containing the significant, non-overlapping clusters, and has the following components:

coords	The centroid of the significant clusters.
r	The radius of the window of the clusters.
pop	The total population in the cluster window.
cases	The observed number of cases in the cluster window.
expected	The expected number of cases in the cluster window.
smr	Standardized mortality ratio (observed/expected) in the cluster window.
rr	Relative risk in the cluster window.
propcases	Proportion of cases in the cluster window.
loglikrat	The loglikelihood ratio for the cluster window (i.e., the log of the test statistic).
pvalue	The pvalue of the test statistic associated with the cluster window.

The second element of the list is the centroid coordinates. This is needed for plotting purposes.

**Author(s)**

Joshua French

**References**

Waller, L.A. and Gotway, C.A. (2005). Applied Spatial Statistics for Public Health Data. Hoboken, NJ: Wiley. Kulldorff M., Nagarwalla N. (1995) Spatial disease clusters: Detection and Inference. *Statistics in Medicine* 14, 799-810. Kulldorff, M. (1997) A spatial scan statistic. *Communications in Statistics – Theory and Methods* 26, 1481-1496.

**Examples**

```
data(grave)
out = spscan.test(grave, parallel = FALSE)
plot(out, chars = c(1, 20), main = "most likely cluster")
# get warning if no significant cluster
out2 = spscan.test(grave, alpha = 0.01)
```

# Index

`circles.intersect`, 2

`contour.im`, 13

`density.ppp`, 16

`grave`, 3

`kd`, 3

`kdest`, 5, 7

`kdplus.test`, 6

`Kest`, 6

`logrr`, 7

`logrr.test`, 9

`nn`, 10

`noc`, 11

`plot.fv`, 12

`plot.im`, 13

`plot.kdenv`, 12

`plot.logrrenv`, 12

`plot.ppp`, 14

`plot.scan`, 13

`qnn.test`, 14

`spdensity`, 15

`spscan.test`, 16