# Package 'Ake'

May 6, 2026

**Type** Package

**Title** Associated Kernel Estimations

**Version** 1.0.2

**Date** 2025-07-02

**Author** W. E. Wansouwé [aut, cre],
S. M. Somé [aut],
C. C. Kokonendji [aut]

**Maintainer** W. E. Wansouwé <ericwansouwe@gmail.com>

**Description** Continuous and discrete (count or categorical) estimation of density, probability mass function (p.m.f.) and regression functions are performed using associated kernels. The cross-validation technique and the local Bayesian procedure are also implemented for bandwidth selection.

**License** GPL (>= 2)

**LazyLoad** yes

**URL** https://www.r-project.org

**NeedsCompilation** no

**Encoding** UTF-8

**Repository** CRAN

**Date/Publication** 2025-07-02 23:10:09 UTC

## Contents

---

`Ake-package` *Associated kernel estimations*

---

## Description

Continuous and discrete estimation of density `dke.fun`, probability mass function (p.m.f.) `kpmfe.fun` and regression `reg.fun` functions are performed using continuous and discrete associated kernels. The cross-validation technique `hcvc.fun`, `hcvreg.fun` and the Bayesian procedure `hbay.fun` are also implemented for bandwidth selection.

## Details

**The estimated density or p.m.f:** The associated kernel estimator $\widehat{f}_n$ of $f$ is defined as

$$\widehat{f}_n(x) = \frac{1}{n}\sum_{i=1}^{n} K_{x,h}(X_i),$$

where $K_{x,h}$ is one of the kernels `kef` defined below. In practice, we first calculate the global normalizing constant

$$C_n = \int_{x\in T} \widehat{f}_n(x)\nu(dx),$$

where $T$ is the support of the density or p.m.f. function and $\nu$ is the Lebesgue or count measure on $T$. For both continuous and discrete associated kernels, this normalizing constant is not generally equal to 1 and it will be computed. The represented density or p.m.f. estimate is then $\tilde{f}_n = \widehat{f}_n / C_n$.

For **discrete data**, the integrated squared error (ISE) defined by

$$ISE_0 = \sum_{x\in N} \{\tilde{f}_n(x) - f_0(x)\}^2$$

is the criteria used to measure the smoothness of the associated kernel estimator $\tilde{f}_n$ with the empirical p.m.f. $f_0$; see Kokonendji and Senga Kiessé (2011).

**The estimated regressor:** Both in continuous and discrete cases, considering the relation between a response variable $y$ and an explanatory variable $x$ given by

$$y = m(x) + \epsilon,$$

where $m$ is an unknown regession function on $T$ and $\epsilon$ the disturbance term with null mean and finite variance. Let $(x_1, y_1), \ldots, (x_n, y_n)$ be a sequence of independent and identically distributed (iid) random vectors on $T \times R$ with $m(x) = E(y|x)$. The well-known Nadaraya-Watson estimator using associated kernels is $\widehat{m}_n$ defined as

$$\widehat{m}_n(x) = \sum_{i=1}^{n} \omega_x(X_i) Y_i,$$

where $\omega_x(X_i) = K_{x,h}(X_i) / \sum_{i=1}^{n} K_{x,h}(X_i)$ and $K_{x,h}$ is one of the associated kernels defined below.

Beside the criterion of kernel support, we retain the root mean squared error (RMSE) and also the practical coefficient of determination $R^2$ defined respectively by

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \{y_i - \widehat{m}_n(x_i)\}^2}$$

and

$$R^2 = \frac{\sum_{i=1}^{n} \{\widehat{m}_n(x_i) - \bar{y}\}^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2},$$

where $\bar{y} = n^{-1}(y_1 + \ldots + y_n)$; see Kokonendji et al. (2009).

Given a data sample, the package allows to compute the density or p.m.f. and regression functions using one of the seven associated kernels: extended beta, lognormal, gamma, reciprocal inverse Gaussian for continuous data, DiracDU for categorical data, and binomial and discrete triangular for count data. The bandwidth parameter is computed using the cross-validation technique. When the associated kernel function is binomial, the bandwidth parameter is also computed using the local Bayesian procedure. The associated kernel functions are defined below. The first four kernels are for continuous data and the last three kernels are for discrete case.

**Extended beta kernel:** The extended beta kernel is defined on $S_{x,h,a,b} = [a, b] = T$ with $a < b < \infty$, $x \in T$ and $h > 0$:

$$BE_{x,h,a,b}(y) = \frac{(y-a)^{(x-a)/\{(b-a)h\}} (b-y)^{(b-x)/\{(b-a)h\}}}{(b-a)^{1+h^{-1}} B\left(1 + (x-a)/(b-a)h, 1 + (b-x)/(b-a)h\right)} 1_{S_{x,h,a,b}}(y),$$

where $B(r, s) = \int_0^1 t^{r-1}(1-t)^{s-1} dt$ is the usual beta function with $r > 0$, $s > 0$ and $1_A$ denotes the indicator function of A. For $a = 0$ and $b = 1$, it corresponds to the beta kernel which is the probability density function of the beta distribution with shape parameters $1+x/h$ and $(1-x)/h$; see Libengué (2013).

**Gamma kernel:** The gamma kernel is defined on $S_{x,h} = [0, \infty) = T$ with $x \in T$ and $h > 0$ by

$$GA_{x,h}(y) = \frac{y^{x/h}}{\Gamma(1 + x/h) h^{1+x/h}} exp\left(-\frac{y}{h}\right) 1_{S_{x,h}}(y),$$

where $\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt$ is the classical gamma function. The probability density function $GA_{x,h}$ is the gamma distribution with scale parameter $1 + x/h$ and shape parameter $h$; see Chen (2000).

**Lognormal kernel:** The lognormal kernel is defined on $S_{x,h} = [0, \infty) = T$ with $x \in T$ and $h > 0$ by

$$LN_{x,h}(y) = \frac{1}{yh\sqrt{2\pi}} exp\left\{ -\frac{1}{2}\left(\frac{1}{h}log(\frac{y}{x}) - h\right)^2 \right\} 1_{S_{x,h}}(y).$$

It is the probability density function of the classical lognormal distribution with parameters $log(x) + h^2$ and $h$; see Libengué (2013).

**Binomial kernel:** Let $x \in N := \{0, 1, \ldots\}$ and $S_x = \{0, 1, \ldots, x + 1\}$. The Binomial kernel is defined on the support $S_x$ by

$$B_{x,h}(y) = \frac{(x+1)!}{y!(x+1-y)!}\left(\frac{x+h}{x+1}\right)^y \left(\frac{1-h}{x+1}\right)^{(x+1-y)} 1_{S_x}(y),$$

where $h \in (0, 1]$. Note that $B_{x,h}$ is the p.m.f. of the binomial distribution with its number of trials $x + 1$ and its success probability $(x + h)/(x + 1)$; see Kokonendji and Senga Kiessé (2011).

**Discrete triangular kernel:** For fixed arm $a \in N$, we define $S_{x,a} = \{x - a, \ldots, x, \ldots, x + a\}$. The discrete triangular kernel is defined on $S_{x,a}$ by

$$DT_{x,h;a}(y) = \frac{(a+1)^h - |y-x|^h}{P(a,h)} 1_{S_{x,a}}(y),$$

where $x \in N$, $h > 0$ and $P(a, h) = (2a+1)(a+1)^h - 2(1+2^h+\cdots+a^h)$ is the normalizing constant. For $a = 0$, the Discrete Triangular kernel $DT_{x,h;0}$ corresponds to the Dirac kernel on $x$; see Kokonendji et al. (2007), and also Kokonendji and Zocchi (2010) for an asymmetric version of discrete triangular.

**DiracDU kernel:** For fixed number of categories $c \in \{2, 3, ...\}$, we define $S_c = \{0, 1, \ldots, c - 1\}$. The DiracDU kernel is defined on $S_c$ by

$$DU_{x,h;c}(y) = (1-h)1_{\{x\}}(y) + \frac{h}{c-1} 1_{S_c\setminus\{x\}}(y),$$

where $x \in S_c$ and $h \in (0, 1]$. See Kokonendji and Senga Kiessé (2011), and also Aitchison and Aitken (1976) for multivariate case.

Note that the global normalizing constant is 1 for DiracDU.

**The bandwidth selection:** Two functions are implemented to select the bandwidth: cross-validation and local Bayesian procedure. The cross-validation technique is used for all the associated kernels both in density and regression; see Kokonendji and Senga Kiessé (2011). The local Bayesian procedure is implemented to select the bandwidth in the estimation of p.m.f. when using binomial kernel; see Zougab et al. (2014).

In the coming versions of the package, adaptive Bayesian procedure will be included for bandwidth selection in density estimation when using gamma kernel. A global Bayesian procedure will also be implemented for bandwidth selection in regression when using binomial kernel.

### Author(s)

W. E. Wansouwé, S. M. Somé and C. C. Kokonendji

Maintainer: W. E. Wansouwé <ericwansouwe@gmail.com>

### References

Aitchison, J. and Aitken, C.G.G. (1976). Multivariate binary discrimination by the kernel method, *Biometrika* **63**, 413 - 420.

Chen, S. X. (1999). Beta kernels estimators for density functions, *Computational Statistics and Data Analysis* **31**, 131 - 145.

Chen, S. X. (2000). Probability density function estimation using gamma kernels, *Annals of the Institute of Statistical Mathematics* **52**, 471 - 480.

Igarashi, G. and Kakizawa, Y. (2015). Bias correction for some asymmetric kernel estimators, *Journal of Statistical Planning and Inference* **159**, 37 - 63.

Kokonendji, C.C. and Senga Kiessé, T. (2011). Discrete associated kernel method and extensions, *Statistical Methodology* **8**, 497 - 516.

Kokonendji, C.C., Senga Kiessé, T. and Demétrio, C.G.B. (2009). Appropriate kernel regression on a count explanatory variable and applications, *Advances and Applications in Statistics* **12**, 99 - 125.

Libengue, F.G. (2013). *Méthode Non-Paramétrique par Noyaux Associés Mixtes et Applications*, Ph.D. Thesis Manuscript (in French) to Université de Franche-Comté, Besançon, France and Université de Ouagadougou, Burkina Faso, June 2013, **LMB no. 14334**, Besançon.

Zougab, N., Adjabi, S. and Kokonendji, C.C. (2014). Bayesian approach in nonparametric count regression with binomial kernel, *Communications in Statistics - Simulation and Computation* **43**, 1052 - 1063.

---

## `dke.fun` *Function for density estimation*

---

### Description

The (S3) generic function `dkde.fun` computes the density. Its default method does so with the given kernel and bandwidth $h$.

### Usage

```
dke.fun(Vec, ...)
## Default S3 method:
dke.fun(Vec, h, type_data = c("discrete", "continuous"),
ker = c("BE", "GA", "LN", "RIG"), x = NULL, a0 = 0, a1 = 1, ... )
```

## Arguments

| | |
|---|---|
| `Vec` | The data sample from which the estimate is to be computed. |
| `h` | The bandwidth or smoothing parameter. |
| `type_data` | The data sample type. Data can be continuous or discrete (categorical or count). Here, in this function , we deal with continuous data. |
| `ker` | A character string giving the smoothing kernel to be used which is the associated kernel: "BE" extended beta, "GA" gamma, "LN" lognormal and "RIG" reciprocal inverse Gaussian. |
| `x` | The points of the grid at which the density is to be estimated. |
| `a0` | The left bound of the support used for extended beta kernel. Default value is 0 for beta kernel. |
| `a1` | The right bound of the support used for extended beta kernel. Default value is 1 for beta kernel. |
| `...` | Further arguments. |

## Details

The associated kernel estimator $\widehat{f}_n$ of $f$ is defined in the above sections. We recall that in general, the sum of the estimated values on the support is not equal to 1. In practice, we compute the global normalizing constant $C_n$ before computing the estimated density $\tilde{f}_n$; see e.g. Libengué (2013).

## Value

Returns a list containing:

| | |
|---|---|
| `data` | The data - same as input Vec. |
| `n` | The sample size. |
| `kernel` | The asssociated kernel used to compute the density estimate. |
| `h` | The bandwidth used to compute the density estimate. |
| `eval.points` | The coordinates of the points where the density is estimated. |
| `est.fn` | The estimated density values. |
| `C_n` | The global normalizing constant. |
| `hist` | The histogram corresponding to the observations. |

## Author(s)

W. E. Wansouwé, S. M. Somé and C. C. Kokonendji

## References

Libengué, F.G. (2013). *Méthode Non-Paramétrique par Noyaux Associés Mixtes et Applications*, Ph.D. Thesis Manuscript (in French) to Université de Franche-Comté, Besançon, France and Université de Ouagadougou, Burkina Faso, June 2013, **LMB no. 14334**, Besançon.

## Examples

```
## A sample data with n=100.
V<-rgamma(100,1.5,2.6)
##The bandwidth can be the one obtained by cross validation.
h<-0.052
## We choose Gamma kernel.

est<-dke.fun(V,h,"continuous","GA")
```

---

hbay.fun *Local Bayesian procedure for bandwidth selection*

---

## Description

The (S3) generic function `hbay.fun` computes the local Bayesian procedure for bandwidth selection.

## Usage

```
hbay.fun(Vec, ...)
## Default S3 method:
hbay.fun(Vec, x = NULL, ...)
```

## Arguments

| | |
|---|---|
| `Vec` | The data sample from which the estimate is to be computed. |
| `x` | The points of the grid where the density is to be estimated. |
| `...` | Further arguments for (non-default) methods. |

## Details

`hbay.fun` implements the choice of the bandwidth $h$ using the local Bayesian approach of a kernel density estimator.

## Value

Returns the bandwidth selected using the local Bayesian procedure.

## Author(s)

W. E. Wansouwé, S. M. Somé and C. C. Kokonendji

## References

Chen, S. X. (1999). Beta kernels estimators for density functions, *Computational Statistics and Data Analysis* **31**, 131 - 145.

Zougab, N., Adjabi, S. and Kokonendji, C.C. (2014). Bayesian approach in nonparametric count regression with binomial kernel, *Communications in Statistics - Simulation and Computation* **43**, 1052 - 1063.

hcvc.fun *Cross-validation function for bandwidth selection for continuous data*

## Description

The (S3) generic function `hcvc.fun` computes the cross-validation bandwidth selector.

## Usage

```
hcvc.fun(Vec,...)
## Default S3 method:
hcvc.fun(Vec, bw = NULL, type_data, ker, a0 = 0, a1 = 1, ...)
```

## Arguments

| | |
|---|---|
| `Vec` | The data sample from which the estimate is to be computed. |
| `bw` | The sequence of bandwidths where to compute the cross-validation. Default value is NULL. |
| `type_data` | The sample data type. |
| `ker` | The associated kernel. |
| `a0` | The left bound of the extended beta. Default value is 0. |
| `a1` | The right bound of the extended beta.Default value is 1. |
| `...` | Further arguments. |

## Details

`hcvc.fun` implements the choice of the bandwidth $h$ using the cross-validation approach of a kernel density estimator.

## Value

Returns a list containing:

| | |
|---|---|
| `hcv` | value of bandwidth parameter. |
| `CV` | the values of cross-validation function. |
| `seq_h` | the sequence of bandwidths where the cross validation is computed. |

## Author(s)

W. E. Wansouwé, S. M. Somé and C. C. Kokonendji

## References

Chen, S. X. (1999). Beta kernels estimators for density functions, *Computational Statistics and Data Analysis* **31**, 131 - 145.

Chen, S. X. (2000). Gamma kernels estimators for density functions, *Annals of the Institute of Statistical Mathematics* **52**, 471 - 480.

Libengué, F.G. (2013). *Méthode Non-Paramétrique par Noyaux Associés Mixtes et Applications*, Ph.D. Thesis Manuscript (in French) to Université de Franche-Comté, Besançon, France and Université de Ouagadougou, Burkina Faso, June 2013, **LMB no. 14334**, Besançon.

Igarashi, G. and Kakizawa, Y. (2015). Bias correction for some asymmetric kernel estimators, *Journal of Statistical Planning and Inference* **159**, 37 - 63.

## Examples

```
V=rgamma(100,1.5,2.6)
## Not run:
hcvc.fun(V,NULL,"continuous","GA")

## End(Not run)
```

---

`hcvd.fun` *Cross-validation function for bandwidth selection in p.m.f. estimation*

---

## Description

The (S3) generic function `hcvd.fun` computes the cross-validation bandwidth selector in p.m.f. estimation.

## Usage

```
hcvd.fun(Vec, ...)
## Default S3 method:
hcvd.fun(Vec, seq_bws = NULL, ker = c("bino", "triang", "dirDU"), a = 1, c = 2,...)
```

## Arguments

| | |
|---|---|
| `Vec` | The data sample from which the estimate is to be computed. |
| `seq_bws` | The sequence of bandwidths where to compute the cross-validation. Default value is NULL. |
| `ker` | The associated kernel |
| `a` | The arm of the discrete triangular kernel. Default value is 1. |
| `c` | The number of categories in DiracDU kernel.Default value is 2. |
| `...` | Further arguments. |

## Details

The `hcvd.fun` function implements the choice of the bandwidth $h$ using the cross-validation approach in p.m.f. estimate.

## Value

Returns a list containing:

| | |
|---|---|
| `hcv` | The optimal bandwidth parameter. |
| `CV` | The cross-validation function values. |
| `seq_h` | The sequence of bandwidths where the cross-validation is computed. |

## Author(s)

W. E. Wansouwé, S. M. Somé and C. C. Kokonendji

## References

Chen, S. X. (1999). Beta kernels estimators for density functions, *Computational Statistics and Data Analysis* **31**, 131 - 145.

Chen, S. X. (2000). Probability density function estimation using gamma kernels, *Annals of the Institute of Statistical Mathematics* **52**, 471 - 480.

Libengué, F.G. (2013). *Méthode Non-Paramétrique par Noyaux Associés Mixtes et Applications*, Ph.D. Thesis Manuscript (in French) to Université de Franche-Comté, Besançon, France and Université de Ouagadougou, Burkina Faso, June 2013, **LMB no. 14334**, Besançon.

Igarashi, G. and Kakizawa, Y. (2015). Bias correction for some asymmetric kernel estimators, *Journal of Statistical Planning and Inference* **159**, 37 - 63.

## Examples

```
## Data can be simulated data or real data
## We use real data
## and then compute the cross validation.
Vec<-c(10,0,1,0,4,0,6,0,0,0,1,1,1,2,4,4,5,6,6,6,6,7,1,7,0,7,7,
7,8,0,8,12,8,8,9,9,0,9,9,10,10,10,10,0,10,10,11,12,12,10,12,12,
13,14,15,16,16,17,0,12)
## Not run:
CV<-hcvd.fun(Vec,NULL,"bino")
CV$hcv

## End(Not run)
##The cross validation function can be also ploted.
## Not run:
plot.fun(CV$seq_bws,CV$CV, type="l")

## End(Not run)
```

hcvreg.fun *Cross-validation function for bandwidth selection in regresssion*

## Description

The (S3) generic function `hcvreg.fun` computes the bandwidth by cross-validation for the regression. Its default method does so. It allows to compute the optimal bandwidth using the cross-validation method. The associated kernels available are: "BE" extended beta, "GA" gamma, "LN" lognormal and "RIG" reciprocal inverse Gaussian, DiracDU, binomial and discrete triangular; see Kokonendji and Senga Kiessé (2011), and also Kokonendji et al. (2009).

## Usage

```
hcvreg.fun(Vec, ...)
## Default S3 method:
hcvreg.fun(Vec, y, type_data = c("discrete", "continuous"),
ker = c("bino", "triang", "dirDU", "BE", "GA", "LN", "RIG"),
 h = NULL, a0 = 0, a1 = 1, a = 1, c = 2, ...)
```

## Arguments

| | |
|---|---|
| `Vec` | The explanatory variable. |
| `y` | The response variable. |
| `type_data` | The data sample type. Data can be continuous or discrete. |
| `ker` | A character string giving the smoothing kernel to be used which is the associated kernel: "BE" extended beta, "GA" gamma, "LN" lognormal and "RIG" reciprocal inverse Gaussian, "dirDU" DiracDU,"bino" binomial, "triang" discrete triangular. |
| `h` | The bandwidth or smoothing parameter.the smoothing bandwidth to be used, can also be a character string giving a rule to choose the bandwidth. |
| `a0` | The left bound of the support used for extended beta kernel. Default value is 0 for beta kernel. |
| `a1` | The right bound of the support used for extended beta kernel. Default value is 1 for beta kernel. |
| `a` | The arm of the discrete triangular kernel |
| `c` | The number of categories |
| `...` | Further arguments |

## Details

The selection of the bandwidth parameter is always crucial. If the bandwidth is small, we will obtain an undersmoothed estimator, with high variability. On the contrary, if the value is big, the resulting estimator will be very smooth and farther from the function that we are trying to estimate. The cross-validation function defined in the above sections is used to compute the optimal bandwidth for the associated kernels.

### Value

Returns a list containing:

| | |
|---|---|
| `kernel` | The associated kernel used to compute the optimal bandwidth. |
| `hcv` | The optimal bandwidth parameter obtained by cross-validation. |
| `CV` | The values of the cross-validation. |
| `seq_bws` | A sequence of bandwidths where the cross-validation is computed. |

### Author(s)

W. E. Wansouwé, S. M. Somé and C. C. Kokonendji

### References

Kokonendji, C.C. and Senga Kiessé, T. (2011). Discrete associated kernel method and extensions, *Statistical Methodology* **8**, 497 - 516.

Kokonendji, C.C., Senga Kiessé, T. and Demétrio, C.G.B. (2009). Appropriate kernel regression on a count explanatory variable and applications, *Advances and Applications in Statistics* **12**, 99 - 125.

### Examples

```
## Data can be simulated data or real data
## We use real data
## and then compute the cross validation.
data(milk)
x=milk$week
y=milk$yield
hcvreg.fun(x,y,"discrete",ker="triang",a=1)
```

---

## `kef` — *Continuous and discrete associated kernel function*

---

### Description

This function computes the associated kernel function.

### Usage

```
kef(x, t, h, type_data = c("discrete", "continuous"),
ker = c("bino", "triang", "dirDU", "BE", "GA", "LN", "RIG"),
a0 = 0, a1 = 1, a = 1, c = 2)
```

### Arguments

| | |
|---|---|
| `x` | The target. |
| `t` | A single value or the grid where the associated kernel function is computed. |
| `h` | The bandwidth or smoothing parameter. |
| `type_data` | The sample data type |
| `ker` | The associated kernel:"bino" Binomial, "triang" discrete triangular kernel, "BE" extended beta, "GA" gamma, "LN" lognormal and "RIG" reciprocal inverse Gaussian,"dirDU" DiracDU. |
| `a0` | The left bound of the support used for extended beta kernel. Default value is 0 for beta kernel. |
| `a1` | The right bound of the support used for extended beta kernel. Default value is 1 for beta kernel. |
| `a` | The arm in discrete triangular kernel. The default value is 1. |
| `c` | The number of categories in DiracDU kernel. The default value is 2. |

### Details

The associated kernel is one of the those which have been defined in the sections above : extended beta, gamma, lognormal, reciprocal inverse Gaussian, DiracDU, binomial and discrete triangular; see Kokonendji and Senga Kiessé (2011), and also Kokonendji et al. (2007).

### Value

Returns the value of the associated kernel function at `t` according to the target and the bandwidth.

### Author(s)

W. E. Wansouwé, S. M. Somé and C. C. Kokonendji

### References

Chen, S. X. (1999). Beta kernels estimators for density functions, *Computational Statistics and Data Analysis* **31**, 131 - 145.

Chen, S. X. (2000). Probability density function estimation using gamma kernels, *Annals of the Institute of Statistical Mathematics* **52**, 471 - 480.

Igarashi, G. and Kakizawa, Y. (2015). Bias correction for some asymmetric kernel estimators, *Journal of Statistical Planning and Inference* **159**, 37 - 63.

Kokonendji, C.C. and Senga Kiessé, T. (2011). Discrete associated kernel method and extensions, *Statistical Methodology* **8**, 497 - 516.

Kokonendji, C.C., Senga Kiessé, T. and Zocchi, S.S. (2007). Discrete triangular distributions and non-parametric estimation for probability mass function, *Journal of Nonparametric Statistics* **19**, 241 - 254.

Libengué, F.G. (2013). *Méthode Non-Paramétrique par Noyaux Associés Mixtes et Applications*, Ph.D. Thesis Manuscript (in French) to Université de Franche-Comté, Besançon, France and Université de Ouagadougou, Burkina Faso, June 2013, **LMB no. 14334**, Besançon.

## Examples

```
x<-5
h<-0.2
t<-0:10
kef(x,t,h,"discrete","bino")
```

---

kern.fun *The associated kernel function*

---

## Description

The (S3) generic function kern.fun computes the value of the associated kernel function. Its default method does so with a given kernel and bandwidth $h$.

## Usage

```
kern.fun(x, ...)
## Default S3 method:
kern.fun(x, t, h, type_data = c("discrete", "continuous"),
 ker = c("bino", "triang", "dirDU", "BE", "GA", "LN", "RIG"),
a0 = 0, a1 = 1, a = 1, c = 2, ...)
```

## Arguments

| | |
|---|---|
| x | The target |
| t | A single value or the grid where the discrete associated kernel function is computed. |
| h | The bandwidth or smoothing parameter. |
| type_data | The sample data type |
| ker | The associated kernel: "dirDU" DiracDU,"bino" Binomial, "triang" Discrete Triangular kernel, "BE" extended beta, "GA" gamma, "LN" lognormal and "RIG" reciprocal inverse Gaussian. |
| a0 | The left bound of the support used for extended beta kernel. Default value is 0 for beta kernel. |
| a1 | The right bound of the support used for extended beta kernel. Default value is 0 for beta kernel. |
| a | The arm in Discrete Triangular kernel. The default value is 1. |
| c | The number of categories in DiracDU kernel. The default value is 2. |
| ... | Further arguments |

## Details

The associated kernel is one of the those which have been defined in the sections above : extended beta, gamma,lognormal, reciprocal inverse Gaussian, DiracDU, Binomial and Discrete Triangular; see Kokonendji and Senga Kiessé (2011), and also Kokonendji et al. (2007).

### Value

Returns the value of the discrete associated kernel function at t according to the target and the bandwidth.

### Author(s)

W. E. Wansouwé, S. M. Somé and C. C. Kokonendji

### References

Kokonendji, C.C. and Senga Kiessé, T. (2011). Discrete associated kernel method and extensions, *Statistical Methodology* **8**, 497 - 516.

Kokonendji, C.C., Senga Kiessé, T. and Zocchi, S.S. (2007). Discrete triangular distributions and non-parametric estimation for probability mass function, *Journal of Nonparametric Statistics* **19**, 241 - 254.

### Examples

```
x<-5
h<-0.2
t<-0:10
kern.fun(x,t,h,"discrete","bino")
```

---

## kpmfe.fun *Function for associated kernel estimation of p.m.f.*

---

### Description

The function estimates the p.m.f. in a single value or in a grid using discrete associated kernels. Three different associated kernels are available: DiracDU (for categorical data), binomial and discrete triangular (for count data).

### Usage

```
kpmfe.fun(Vec,...)
## Default S3 method:
kpmfe.fun(Vec, h, type_data = c("discrete", "continuous"),
        ker = c("bino", "triang", "dirDU"), x = NULL, a = 1, c = 2, ...)
```

### Arguments

| | |
|---|---|
| Vec | the data sample from which the estimate is to be computed. |
| h | The bandwidth or smoothing parameter. The smoothing bandwidth to be used, can also be a character string giving a rule to choose the bandwidth. |
| type_data | The data sample type. Data type is "discrete" (categorical or count). |
| ker | The associated kernel: "dirDU" DiracDU,"bino" binomial, "triang" discrete triangular. |

| | |
|---|---|
| x | The points of the grid at which the density is to be estimated. |
| a | The arm in discrete triangular kernel. The default value is 1. |
| c | The number of categories in DiracDU. The default value is 2. |
| ... | Further arguments. |

## Details

The associated kernel estimator $\widehat{f}_n$ of $f$ is defined in the above sections. We recall that in general, the sum of the estimated values on the support is not equal to 1. In practice, we compute the global normalizing constant $C_n$ before computing the estimated p.m.f. $\tilde{f}_n$; see Kokonendji and Senga Kiessé (2011).

The bandwidth parameter in the function is obtained using the cross-validation technique for the three associated kernels. For binomial kernel, the local Bayesian approach is also implemented and is recommanded to select the bandwidth; see Zougab et al. (2012).

## Value

Returns a list containing:

| | |
|---|---|
| data | The number of observations. |
| n | The number of observations. |
| eval.points | The support of the estimated p.m.f. |
| h | The bandwidth |
| C_n | The global normalizing constant. |
| ISE_0 | The integrated square error. |
| f_0 | A vector of (x,f0(x)). |
| f_n | A vector of (x,fn(x)). |
| f0 | The empirical p.m.f. |
| est.fn | The estimated p.m.f. containing estimated values after normalization. |

## Author(s)

W. E. Wansouwé, S. M. Somé and C. C. Kokonendji

## References

Kokonendji, C.C. and Senga Kiessé, T. (2011). Discrete associated kernel method and extensions, *Statistical Methodology* **8**, 497 - 516.

Kokonendji, C.C., Senga Kiessé, T. and Zocchi, S.S. (2007). Discrete triangular distributions and non-parametric estimation for probability mass function. *Journal of Nonparametric Statistics* **19**, 241 - 254.

Zougab, N., Adjabi, S. and Kokonendji, C.C. (2012). Binomial kernel and Bayes local bandwidth in discrete functions estimation. *Journal of Nonparametric Statistics* **24**, 783 - 795.

### Examples

```
## A sample data with n=60.
V<-c(10,0,1,0,4,0,6,0,0,0,1,1,1,2,4,4,5,6,6,6,6,7,1,7,0,7,7,
7,8,0,8,12,8,8,9,9,0,9,9,10,10,10,10,0,10,10,11,12,12,10,12,12,
13,14,15,16,16,17,0,12)


##The bandwidth can be the one obtained by cross validation.
h<-0.081
## We choose Binomial kernel.

est<-kpmfe.fun(Vec=V,h,"discrete","bino")
##To obtain the normalizing constant:
est
```

---

## milk *Average daily fat yields.*

---

### Description

This data is the average daily fat yields (kg/day) of milk from a single cow for each of 35 weeks; see Kokonendji et al. (2009).

### Usage

```
data(milk)
```

### Format

A data frame with 35 observations on the following 2 variables.

`week` Number of the week

`yield` The yield quantity

### Source

McCulloch, C.E. (2001). An Introduction to Generalized Linear Mixed Models, 46a Reuniao Anual da RBRAS - 9o SEAGRO, University of Sao Paulo - ESALQ, Piracicaba.

### References

Kokonendji, C.C., Senga Kiessé, T. and Demétrio, C.G.B. (2009). Appropriate kernel regression on a count explanatory variable and applications, *Advances and Applications in Statistics* **12**, 99 - 125.

### Examples

```
data(milk)
```

## plot.dke.fun *Plot of density function*

### Description

The plot.dke.fun is to plot the associated kernel density estimation.

### Usage

```
## S3 method for class 'dke.fun'
plot(x,main = NULL, sub = NULL, xlab = NULL,
ylab = NULL, type = "l", las = 1, lwd = 1, col = "blue", lty = 1, ...)
```

### Arguments

| | |
|---|---|
| x | An object class dke.fun |
| main | The main parameter |
| sub | The sub title |
| xlab, ylab | The axis label |
| type | the type parameter |
| las | Numeric in {0,1,2,3}; the style of axis labels. |
| lwd | The line width, a positive number, defaulting to 1. |
| col | A specification for the default plotting color. |
| lty | The line type. |
| ... | Futher arguments |

### Value

Plot of associated kernel density function is sent to graphics window.

### Author(s)

W. E. Wansouwé, S. M. Somé and C. C. Kokonendji

### References

Kokonendji, C.C. and Senga Kiessé, T. (2011). Discrete associated kernel method and extensions, *Statistical Methodology* **8**, 497 - 516.

Kokonendji, C.C., Senga Kiessé, T. and Zocchi, S.S. (2007). Discrete triangular distributions and non-parametric estimation for probability mass function. *Journal of Nonparametric Statistics* **19**, 241 - 254.

Zougab, N., Adjabi, S. and Kokonendji, C.C. (2012). Binomial kernel and Bayes local bandwidth in discrete functions estimation. *Journal of Nonparametric Statistics* **24**, 783 - 795.

---

`plot.hcvc.fun` *Plot of cross-validation function for bandwidth selection in density or p.m.f. estimation.*

---

## Description

The functions allows to plot the cross-validation both in discrete `plot.hcvd.fun` and continuous `plot.hcvc.fun` cases.

## Usage

```
## S3 method for class 'hcvc.fun'
plot(x,  ...)
## S3 method for class 'hcvd.fun'
plot(x,  ...)
```

## Arguments

| | |
|---|---|
| `x` | an object |
| `...` | Further arguments |

## Details

Plot a graphic for cross-validation function

## Value

returns a graphics

## Author(s)

W. E. Wansouwé, S. M. Somé and C. C. Kokonendji

## References

Kokonendji, C.C. and Senga Kiessé, T. (2011). Discrete associated kernel method and extensions, *Statistical Methodology* **8**, 497 - 516.

Kokonendji, C.C., Senga Kiessé, T. and Zocchi, S.S. (2007). Discrete triangular distributions and non-parametric estimation for probability mass function. *Journal of Nonparametric Statistics* **19**, 241 - 254.

Zougab, N., Adjabi, S. and Kokonendji, C.C. (2012). Binomial kernel and Bayes local bandwidth in discrete functions estimation. *Journal of Nonparametric Statistics* **24**, 783 - 795.

## plot.kern.fun *Plot of associated kernel function*

### Description

The plot.kern.fun function loops through calls to the kern.fun function.

### Usage

```
## S3 method for class 'kern.fun'
plot(x, ...)
```

### Arguments

| | |
|---|---|
| x | an object of class kern.fun (output from kern.fun). |
| ... | Other graphics parameters |

### Value

Plot of associated the kernel function is sent to graphics window.

### Author(s)

W. E. Wansouwé, S. M. Somé and C. C. Kokonendji

### References

Kokonendji, C.C. and Senga Kiessé, T. (2011). Discrete associated kernel method and extensions, *Statistical Methodology* **8**, 497 - 516.

Kokonendji, C.C., Senga Kiessé, T. and Demétrio, C.G.B. (2009). Appropriate kernel regression on a count explanatory variable and applications, *Advances and Applications in Statistics* **12**, 99 - 125.

## plot.kpmfe.fun *Plot of the function for associated kernel estimation of the p.m.f.*

### Description

The function plots the p.m.f. estimation in a single value or in a grid using discrete associated kernels. Three different associated kernels are available: DiracDU (for categorical data), binomial and discrete triangular (for count data).

### Usage

```
## S3 method for class 'kpmfe.fun'
plot(x, ...)
```

### Arguments

| | |
|---|---|
| `x` | An object of class `kpmfe.fun`. |
| `...` | Further arguments |

### Details

Plot a graphic

### Author(s)

W. E. Wansouwé, S. M. Somé and C. C. Kokonendji

### References

Kokonendji, C.C. and Senga Kiessé, T. (2011). Discrete associated kernel method and extensions, *Statistical Methodology* **8**, 497 - 516.

Kokonendji, C.C., Senga Kiessé, T. and Zocchi, S.S. (2007). Discrete triangular distributions and non-parametric estimation for probability mass function. *Journal of Nonparametric Statistics* **19**, 241 - 254.

Zougab, N., Adjabi, S. and Kokonendji, C.C. (2012). Binomial kernel and Bayes local bandwidth in discrete functions estimation. *Journal of Nonparametric Statistics* **24**, 783 - 795.

---

## `plot.reg.fun` *Plot for associated kernel regression*

---

### Description

Plot for associated kernel regression for univariate data. The `plot.reg.fun` function loops through calls to the `reg.fun` function.

### Usage

```
## S3 method for class 'reg.fun'
plot(x,  ...)
```

### Arguments

| | |
|---|---|
| `x` | An object of class `reg.fun` |
| `...` | other graphics parameters |

### Details

The function allows to plot the regression

### Value

Plot is sent to graphics window.

### Author(s)

W. E. Wansouwé, S. M. Somé and C. C. Kokonendji

### References

Kokonendji, C.C. and Senga Kiessé, T. (2011). Discrete associated kernel method and extensions, *Statistical Methodology* **8**, 497 - 516.

Kokonendji, C.C., Senga Kiessé, T. and Zocchi, S.S. (2007). Discrete triangular distributions and non-parametric estimation for probability mass function. *Journal of Nonparametric Statistics* **19**, 241 - 254.

Zougab, N., Adjabi, S. and Kokonendji, C.C. (2012). Binomial kernel and Bayes local bandwidth in discrete functions estimation. *Journal of Nonparametric Statistics* **24**, 783 - 795.

---

## `print.reg.fun` *Print for regression function*

---

### Description

The function allows to print the result of computation in regression as a data frame.

### Usage

```
## S3 method for class 'reg.fun'
print(x, digits = NULL, ...)
```

### Arguments

| | |
|---|---|
| `x` | object of class `reg.fun`. |
| `digits` | The number of digits |
| `...` | Further arguments |

### Details

The associated kernel estimator $\widehat{m}_n$ of $m$ is defined in the above sections; see Kokonendji and Senga Kiessé (2011). The bandwidth parameter in the function is obtained using the cross-validation technique for the associated kernels.

## Value

Returns a list containing:

| | |
|---|---|
| data | The explanatory variable, printed as a data frame |
| y | The response variable, printed as a data frame |
| n | The size of the sample |
| kernel | The associated kernel |
| h | The smoothing parameter |
| eval.points | The grid where the regression is computed, printed as data frame |
| m_n | The estimated values, printed as data frame |
| Coef_det | The Coefficient of determination |

## Author(s)

W. E. Wansouwé, S. M. Somé and C. C. Kokonendji

## References


Kokonendji, C.C. and Senga Kiessé, T. (2011). Discrete associated kernel method and extensions, *Statistical Methodology* **8**, 497 - 516.

Kokonendji, C.C., Senga Kiessé, T. and Demétrio, C.G.B. (2009). Appropriate kernel regression on a count explanatory variable and applications, *Advances and Applications in Statistics* **12**, 99 - 125.

Zougab, N., Adjabi, S. and Kokonendji, C.C. (2014). Bayesian approach in nonparametric count regression with Binomial Kernel, *Communications in Statistics - Simulation and Computation* **43**, 1052 - 1063.


## Examples

```
data(milk)
x=milk$week
y=milk$yield
##The bandwidth is the one obtained by cross validation.
h<-0.10
## We choose binomial kernel.
m_n<-reg.fun(x, y, "discrete",ker="bino", h)
print.reg.fun(m_n)
```

---

reg.fun *Function for associated kernel estimation of regression*

---

## Description

The function estimates the discrete and continuous regression in a single value or in a grid using associated kernels. Different associated kernels are available: extended beta, gamma, lognormal, reciprocal inverse Gaussian (for continuous data), DiracDU (for categorical data), binomial and also discrete triangular (for count data).

## Usage

```
reg.fun(Vec, ...)
## Default S3 method:
reg.fun(Vec, y, type_data = c("discrete", "continuous"),
ker = c("bino", "triang", "dirDU", "BE", "GA", "LN", "RIG"),
 h, x = NULL, a0 = 0, a1 = 1, a = 1, c = 2, ...)
```

## Arguments

| | |
|---|---|
| `Vec` | The explanatory variable. |
| `y` | The response variable. |
| `type_data` | The sample data type. |
| `ker` | The associated kernel: "dirDU" DiracDU,"bino" binomial, "triang" discrete triangular, etc. |
| `h` | The bandwidth or smoothing parameter. |
| `x` | The single value or the grid where the regression is computed. |
| `a0` | The left bound of the support used for extended beta kernel. Default value is 0 for beta kernel. |
| `a1` | The right bound of the support used for extended beta kernel. Default value is 0 for beta kernel. |
| `a` | The arm in Discrete Triangular kernel. The default value is 1. |
| `c` | The number of categories in DiracDU. The default value is 2. |
| `...` | Further arguments |

## Details

The associated kernel estimator $\widehat{m}_n$ of $m$ is defined in the above sections; see also Kokonendji and Senga Kiessé (2011). The bandwidth parameter in the function is obtained using the cross-validation technique for the seven associated kernels. For binomial kernel, the local Bayesian approach is also implemented; see Zougab et al. (2014).

## Value

Returns a list containing:

| | |
|---|---|
| `data` | The data sample, explanatory variable |
| `y` | The data sample, response variable |
| `n` | The size of the sample |
| `kernel` | The asociated kernel |
| `h` | The bandwidth |
| `eval.points` | The grid where the regression is computed |
| `m_n` | The estimated values |
| `Coef_det` | The coefficient of determination |

## Author(s)

W. E. Wansouwé, S. M. Somé and C. C. Kokonendji

## References

Kokonendji, C.C. and Senga Kiessé, T. (2011). Discrete associated kernel method and extensions, *Statistical Methodology* **8**, 497 - 516.

Kokonendji, C.C., Senga Kiessé, T. and Demétrio, C.G.B. (2009). Appropriate kernel regression on a count explanatory variable and applications, *Advances and Applications in Statistics* **12**, 99 - 125.

Zougab, N., Adjabi, S. and Kokonendji, C.C. (2014). Bayesian approach in nonparametric count regression with binomial kernel, *Communications in Statistics - Simulation and Computation* **43**, 1052 - 1063.

## Examples

```
data(milk)
x=milk$week
y=milk$yield
##The bandwidth is the one obtained by cross validation.
h<-0.10
## We choose binomial kernel.
## Not run:
m_n<-reg.fun(x, y, "discrete",ker="bino", h)

## End(Not run)
```

# Index