Package 'empirical'

December 3, 2018

Title Probability Distributions as Models of Data

Version 0.2.0

Date 2018-12-03

License GPL (>= 2)

Maintainer Abby Spurdle <spurdle.a@gmail.com>

Author Abby Spurdle

URL https://sites.google.com/site/asrpws

Description Computes continuous (not step) empirical (and nonparametric) probability density, cumulative distribution and quantile functions. Supports univariate, multivariate and conditional probability distributions, some kernel smoothing features and weighted data (possibly useful mixed with fuzzy clustering). Can compute multivariate and conditional probabilities. Also, can compute conditional medians, quantiles and modes.

Depends graphics, stats

Imports barsurf

Suggests intoo, bivariate, fclust, mgcv, gam, moments

NeedsCompilation no

Repository CRAN

Date/Publication 2018-12-02 23:30:04 UTC

R topics documented:

comb.prob *Compute Probabilities*

Description

Compute probabilities combinatorially.

Usage

```
comb.prob (F, a, b)
```

Arguments

F	An ecdfmv object.
a	Either a vector or a matrix. Each element (or column) gives the lower limits.
b	Either a vector or a matrix. Each element (or column) gives the upper limits.

See Also

ecdfmv

Examples

```
data (trees)
attach (trees)
ecdfmv.f = ecdfmv (cbind (Height, Volume) )
a = c (0, 0)
b = c (80, 30)
comb.prob (ecdfmv.f, a, b)
```

conditional *Nonparametric Conditional Probability Distributions*

Description

Empirical conditional probability density functions, empirical conditional cumulative distribution functions and empirical conditional quantile functions. Refer to the vignette for better examples.

Usage

```
epdfc (rv, conditions, x,
    restack.pdf=sbcpdf, rsp=0.5, bw,
    bind=TRUE, w=NA, is.string=FALSE, npoints=30)
ecdfc (rv, conditions, x,
    restack.pdf=sbcpdf, rsp=0.5, bw,
    bind=TRUE, w=NA, is.string=FALSE, npoints=30)
ecdfc.inverse (rv, conditions, x,
    restack.pdf=sbcpdf, rsp=0.5, bw,
    bind=TRUE, w=NA, is.string=FALSE, npoints=30)
```

Arguments

rv	A name of a random variable.
conditions	A named vector of conditions.
x	A vector of data points.
restack.pdf	A PDF used for restacking.
rsp	A restacking parameter. Refer to the vignette.
bw	A bandwidth parameter. Refer to the vignette.
bind	If true, add an extra two data points.
w	A vector of weights.
is.string	If true rv is expected to be a character type, otherwise, rv is expected to be an unquoted name of a variable.
npoints	Number of synthetic data points.

Value

These functions return functions.

See Also

epdfuv, ecdfuv, ecdfuv.inverse, epdfmv, ecdfmv

Examples

```
#construct an empirical multivariate probability density function
#and then evaluate it
data (trees)
attach (trees)
epdfc.f = epdfc (Volume, c (Height=80), cbind (Height, Volume) )
epdfc.f (30)
```

mode *Empirical Mode*

Description

Compute the mode of data using the empirical probability density function. Note that there are some problems using all=TRUE or warning=TRUE.

Usage

```
emode (f, include.boundaries=TRUE, all=FALSE, warning=FALSE)
```

Arguments

f	An epdfuv or epdfc object.
include.boundaries	Include the min and max data values.
all	If true, return all modal points, otherwise, return the max modal point only.
warning	If true generate a warning if there are no or multiple modal points.

See Also

epdfuv, epdfc

Examples

```
#compute the empirical mode
x = rnorm (30, 4) ^ 2
epdfuv.f = epdfuv (x)
emode (epdfuv.f)
```

multivariate *Nonparametric Multivariate Probability Distributions*

Description

Empirical multivariate probability density functions and empirical multivariate cumulative distribution functions. Refer to the vignette for better examples.

Usage

```
epdfmv (x,
    restack.pdf=sbcpdf, restack.cdf=sbccdf,
    rsp=0.5, bw, bind=TRUE, w=NA)
ecdfmv (x,
    restack.pdf=sbcpdf, restack.cdf=sbccdf,
    rsp=0.5, bw, bind=TRUE, w=NA)
```

Arguments

x	A vector of data points.
restack.pdf	A PDF used for restacking. ECDFs ignore this.
restack.cdf	A CDF used for restacking. EPDFs only use this if plotting bivariate models with all=TRUE.
rsp	A restacking parameter. Refer to the vignette.
bw	A bandwidth parameter. Refer to the vignette.
bind	If true, add an extra two data points.
w	A vector of weights.

Value

These functions return functions.

See Also

epdfuv, ecdfuv, ecdfuv.inverse, epdfc, ecdfc, ecdfc.inverse, epdfmv.f, ecdfmv.f

Examples

```
#construct an empirical multivariate probability density function
#and then evaluate it
data (trees)
attach (trees)
epdfmv.f = epdfmv (cbind (Height, Volume) )
epdfmv.f (c (80, 30) )
```

---

plots_bivariate *Plots Bivariate*

Description

Plots of multivariate probability distributions (with m=2).

Usage

```
## S3 method for class 'epdfmv'
plot(x, use.plot3d=FALSE,
    xlab="x1", ylab="x2", npoints=30, ..., all=FALSE)
## S3 method for class 'ecdfmv'
plot(x, use.plot3d=FALSE,
    xlab="x1", ylab="x2", npoints=30, ...)
```

Arguments

x	An epdfmv or ecdfmv object.
use.plot3d	If true, use a 3d plot, otherwise, use a 2d plot.
xlab	The x label.
ylab	The y label.
npoints	The number of grid points in each direction.
...	Other arguments.
all	If true, plot all combinations.

See Also

epdfmv, ecdfmv

Examples

```
#plot an empirical multivariate cumulative distribution function
#(with m=2)
data (trees)
attach (trees)
ecdfmv.f = ecdfmv (cbind (Height, Volume) )
plot (ecdfmv.f, TRUE)
```

---

plots_univariate *Plots Univariate*

Description

Plots of univariate probability distributions.

Usage

```
## S3 method for class 'epdfuv'
plot(x, plot.points=FALSE, ...)
## S3 method for class 'ecdfuv'
plot(x, plot.points=FALSE, ...)
## S3 method for class 'ecdfuv.inverse'
plot(x, plot.points=FALSE, ...)
## S3 method for class 'epdfuv'
lines(x, ...)
## S3 method for class 'ecdfuv'
lines(x, ...)
## S3 method for class 'ecdfuv.inverse'
lines(x, ...)
```

Arguments

x	An epdfuv, ecdfuv, ecdfuv.inverse, epdfc, ecdfc or ecdfc.inverse object.
plot.points	Plot the data points.
...	Other arguments.

See Also

epdfuv, ecdfuv, ecdfuv.inverse

Examples

```
#plot an empirical univariate cumulative distribution function
data (trees)
attach (trees)
ecdfuv.f = ecdfuv (Height)
plot (ecdfuv.f)
```

---

returned_functions *Returned Functions*

Description

Functions returned by functions. Note that you can name your functions whatever you like. (In the vignette I've named them f, F and F.inv). Also note that you don't call these functions, rather call epdfuv(), ecdfuv(), etc and then call the functions that are returned.

Usage

```
epdfuv.f (x)
ecdfuv.f (x)
ecdfuv.f.inverse (y)
epdfmv.f (x)
ecdfmv.f (x)
```

Arguments

x	A vector of x values.
y	A vector of y values.

See Also

epdfuv, ecdfuv, ecdfuv.inverse, epdfmv, ecdfmv, epdfc, ecdfc, ecdfc.inverse

Examples

```
#construct an empirical quantile function and then evaluate it
#(compute the 0.25 quantile)
x = rnorm (30, 4) ^ 2
ecdfuv.f.inverse = ecdfuv.inverse (x)
ecdfuv.f.inverse (0.25)
```

---

sbc *Simplified Bell Curves*

Description

These are kernels used for restacking.

Usage

```
sbcpdf (x)
sbccdf (x)
```

Arguments

x	A vector of x values.

Examples

```
sbcpdf (c (-1, -0.5, 0, 0.5, 1) )
```

---

univariate *Nonparametric Univariate Probability Distributions*

Description

Empirical univariate probability density functions, empirical univariate cumulative distribution functions and empirical univariate quantile functions. Refer to the vignette for better examples.

Usage

```
epdfuv (x, derandomize=TRUE, preserve="mean",
    drp, nhood, bind=TRUE, randomize=TRUE, w=NA)
ecdfuv (x, derandomize=TRUE, preserve="mean",
    drp, nhood, bind=TRUE, randomize=TRUE, w=NA)
ecdfuv.inverse (x, derandomize=TRUE, preserve="mean",
    drp, nhood, bind=TRUE, randomize=TRUE, w=NA)
```

Arguments

x	A vector of data points.
derandomize	If true, smooth the data points.
preserve	Either "mean" or "range". If derandomize and mean (the default), preserve the mean and variance. If derandomize and range, preserve the range.
drp	A smoothness (derandomization) parameter. Refer to the vignette.
nhood	A neighborhood size parameter. Refer to the vignette.
bind	If true, add an extra two data points.
randomize	If there a duplicated values, add a small amount of random variation.
w	A vector of weights.

Value

These functions return functions.

See Also

epdfmv, ecdfmv, epdfc, ecdfc, ecdfc.inverse, epdfuv.f, ecdfuv.f, ecdfuv.f.inverse

Examples

```
#construct an empirical univariate probability density function
#and then evaluate it
data (trees)
attach (trees)
epdfuv.f = epdfuv (Height)
epdfuv.f (80)
```

Index