Package 'rrp'

January 2, 2012

Type Package

Title Random Recursive Partitioning

Version 2.9

Date 2009-02-18

Depends rpart

Author S.M. Iacus

Maintainer S.M. Iacus <stefano.iacus@unimi.it>

Description Random Recursive Partitiong and Rank-based proximities for data matching, missing data imputation and nonparametric classification and prediction

License GPL (>= 2)

Repository CRAN

Date/Publication 2009-02-19 07:59:18

R topics documented:

`rrp-package` *Implementation of the Random Recursive Partitionig algorithm*

Description

This package implements the RRP algorithm introduced in Iacus-Porro (2009) and different functions to for the applications described in Iacus-Porro (2007). It also contains benchmark datasets in econometric average treatment effect estimation taken from the Lalonde (1986) paper.

Details

Package:	rrp
Type:	Package
Version:	1.0
Date:	2006-02-13
License:	GPL Version 2 or later.

The main function is `rrp.dist`

Author(s)

Stefano M. Iacus

Maintainer: Stefano M. Iacus <stefano.iacus@unimi.it>

References

Iacus, S.M., Porro, G. (2009) Random Recursive Partitioning: a matching method for the estimation of the average treatment effect, *Journal of Applied Econometrics*, 24, 163-185.

Iacus, S.M., Porro, G. (2007) Missing data imputation, matching and other applications of random recursive partitioning, *Computational Statistics and Data Analysis*, 52, 2, 773-789.

Lalonde, R. (1986) Evaluating the Econometric Evaluations of Training Programs, *American Economic Review*, 76, 604-620.

addDist *Adds constant to the elements of a dist object vector-wise*

Description

Adds constant to the elements of a dist object vector-wise. The idea is that a symmetrix matrix `M` can be represented as a triangular object `d`. With this function you set `M[x,x] = M[x,x] + k` but instead of working on `M` this function works on its `dist` representation `d`.

Usage

```
addDist(d, x, k)
```

Arguments

`d`	a dist object, upper diagonal, representing a symmetric matrix
`x`	a list of vectors of indices
`k`	a vector of constants to add, arguments are eventually recycled silently

Value

returns the modified dist object

Author(s)

Stefano M. Iacus

See Also

`dist`

Examples

```
M <- matrix(0,5,5)
d <- as.dist(M)
x <- list(1:3, 4:5)

addDist(d,x,c(-1,+1))

# which is the equivalent of (apart for the diagonal elements)
M[1:3,1:3] <- M[1:3,1:3] - 1
M[4:5,4:5] <- M[4:5,4:5] + 1

# Indeed, we have
d
as.dist(M)
```

addXPtr *Adds constant to the elements of a XPtr object vector-wise*

Description

Adds constant to the elements of a `XPtr` object vector-wise. The idea is that a symmetrix matrix `M` can be represented as a triangular object `d`. With this function you set `M[x,x] = M[x,x] + k` but instead of working on `M` this function works on its `XPtr` representation `d`.

Usage

```
addXPtr(d, x, k)
```

Arguments

`d`	a `XPtr` object, lower diagonal, representing a symmetric matrix
`x`	a list of vectors of indices
`k`	a vector of constants to add, arguments are eventually recycled silently

Value

returns the modified `XPtr` object

Author(s)

Stefano M. Iacus

See Also

`dist`

Examples

```
M <- matrix(0,5,5)
d <- newXPtr(5,0)
x <- list(1:3, 4:5)

addXPtr(d,x,c(-1,+1))

# which is the equivalent of (apart for the diagonal elements)
M[1:3,1:3] <- M[1:3,1:3] - 1
M[4:5,4:5] <- M[4:5,4:5] + 1

# Indeed, we have
(XPtrToDist(d))
as.dist(M)
```

applyXPtr *Apply facility for objects of class XPtr*

Description

This tool function allow to apply a function along some index and on a subset of elements of the XPtr object. Is the counterpart of apply for objects fo class matrix.

Usage

```
applyXPtr(d, idx, sub, f)
```

Arguments

d	an object of class XPtr
idx	a vector of integers, the indexes of the apply
sub	a vector of integers, the indexes of the subest
f	a function to be evaluated in the .GlobalEnv

Details

The object d is like a dist object, which is one-dimensional but usually associated to a symmetrix matrix M with 0's on the diagonal. This function allows to apply a function f on d as if it was applied this way on the matrix M: apply(M[idx,sub],1,f).

Value

val	an invisible object of type list

Author(s)

S.M. Iacus

References

Iacus, S.M., Porro, G. (2009) Random Recursive Partitioning: a matching method for the estimation of the average treatment effect, *Journal of Applied Econometrics*, 24, 163-185.

Iacus, S.M., Porro, G. (2007) Missing data imputation, matching and other applications of random recursive partitioning, *Computational Statistics and Data Analysis*, 52, 2, 773-789.

See Also

newXPtr

Examples

```
a <- newXPtr(10,1)
addXPtr(a, list(c(1,3,4,9), c(2,5), c(1,7,9)), c(-1, 5, 10))
f <- function(x) sum(x)
idx <- 1:5
sub <- 7:10
applyXPtr(a, idx, sub, f) -> l
(XPtrToDist(a))
l
```

DW *Dehejia-Wahba dataset*

Description

A subset of the Lalonde dataset (see cited reference).

Usage

```
data(DW)
```

Format

A data frame with 445 observations on the following 10 variables.

`treated` treated variable indicator

`age` age

`education` years of education

`black` race indicator variable

`married` marital status indicator variable

`nodegree` indicator variable of not possessing a degree

`re74` real earnings in 1974

`re75` real earnings in 1975

`re78` real earnings in 1978 (post treatment outcome)

`hispanic` ethnic indicator variable

Source

see references

References

Dehejia, R., Wahba, S. (1999) Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs, *Journal of the American Statistical Association*, 94, 1053-1062.

DWvsCPS *Dehejia-Wahba treated units versus CPS control individuals*

Description

The Lalonde subset of DW treated units versus CPS (Current Population Survey) control individuals

Usage

```
data(DWvsCPS)
```

Format

A data frame with 16177 observations on the following 10 variables.

`treated` treated variable indicator

`age` age

`education` years of education

`black` race indicator variable

`married` marital status indicator variable

`nodegree` indicator variable of not possessing a degree

`re74` real earnings in 1974

`re75` real earnings in 1975

`re78` real earnings in 1978 (post treatment outcome)

`hispanic` ethnic indicator variable

Details

These two sets of treated and control units can be hardly matched.

Source

see references

References

Dehejia, R., Wahba, S. (1999) Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs, *Journal of the American Statistical Association*, 94, 1053-1062.

DWvsPSID *Dehejia-Wahba treated units versus PSID control individuals*

Description

The Lalonde subset of DW treated units versus PSID (Panel Study of Income Dynamics) control individuals

Usage

```
data(DWvsPSID)
```

Format

A data frame with 2675 observations on the following 10 variables.

`treated` treated variable indicator

`age` age

`education` years of education

`black` race indicator variable

`married` marital status indicator variable

`nodegree` indicator variable of not possessing a degree

`re74` real earnings in 1974

`re75` real earnings in 1975

`re78` real earnings in 1978 (post treatment outcome)

`hispanic` ethnic indicator variable

Details

These two sets of treated and control units can be hardly matched.

Source

see references

References

Dehejia, R., Wahba, S. (1999) Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs, *Journal of the American Statistical Association*, 94, 1053-1062.

LL *Lalonde dataset*

Description

Lalonde experimental dataset (see cited reference).

Usage

```
data(LL)
```

Format

A data frame with 722 observations on the following 10 variables.

`treated` treated variable indicator

`age` age

`education` years of education

`black` race indicator variable

`married` marital status indicator variable

`nodegree` indicator variable of not possessing a degree

`re74` real earnings in 1974

`re75` real earnings in 1975

`re78` real earnings in 1978 (post treatment outcome)

`hispanic` ethnic indicator variable

Source

see references

References

Lalonde, R. (1986) Evaluating the Econometric Evaluations of Training Programs, *American Economic Review*, 76, 604-620.

LLvsCPS *Lalonde treated units versus CPS control individuals*

Description

The Lalonde set of treated units versus CPS (Current Population Survey) control individuals

Usage

```
data(LLvsCPS)
```

Format

A data frame with 16289 observations on the following 10 variables.

`treated` treated variable indicator

`age` age

`education` years of education

`black` race indicator variable

`married` marital status indicator variable

`nodegree` indicator variable of not possessing a degree

`re74` real earnings in 1974

`re75` real earnings in 1975

`re78` real earnings in 1978 (post treatment outcome)

`hispanic` ethnic indicator variable

Details

These two sets of treated and control units can be hardly matched.

Source

see references

References

Lalonde, R. (1986) Evaluating the Econometric Evaluations of Training Programs, *American Economic Review*, 76, 604-620.

LLvsPSID *Lalonde treated units versus PSID control individuals*

Description

The Lalonde set of treated units versus PSID (Panel Study of Income Dynamics) control individuals

Usage

```
data(LLvsPSID)
```

Format

A data frame with 2787 observations on the following 10 variables.

`treated` treated variable indicator

`age` age

`education` years of education

`black` race indicator variable

`married` marital status indicator variable

`nodegree` indicator variable of not possessing a degree

`re74` real earnings in 1974

`re75` real earnings in 1975

`re78` real earnings in 1978 (post treatment outcome)

`hispanic` ethnic indicator variable

Details

These two sets of treated and control units can be hardly matched.

Source

see references

References

Lalonde, R. (1986) Evaluating the Econometric Evaluations of Training Programs, *American Economic Review*, 76, 604-620.

`mulXPtr` *Multiplies the elements of a XPtr object by a constant vector-wise*

Description

Multiplies the elements of a `XPtr` object by a constant vector-wise. The idea is that a symmetrix matrix `M` can be represented as a triangular object `d`. With this function you set `M[x,x] = M[x,x] * k` but instead of working on `M` this function works on its `XPtr` representation `d`.

Usage

```
mulXPtr(d, x, k)
```

Arguments

`d`	a `XPtr` object, lower diagonal, representing a symmetric matrix
`x`	a list of vectors of indices
`k`	a vector of multiplicative constants, arguments are eventually recycled silently

Value

returns the modified `XPtr` object

Author(s)

Stefano M. Iacus

See Also

`dist`

Examples

```
M <- matrix(2,5,5)
d <- newXPtr(5,2)
x <- list(1:3, 4:5)

mulXPtr(d,x,c(-1,2))

# which is the equivalent of (apart for the diagonal elements)
M[1:3,1:3] <- M[1:3,1:3] * (-1)
M[4:5,4:5] <- M[4:5,4:5] * 2

# Indeed, we have
(XPtrToDist(d))
as.dist(M)
```

newXPtr *Creates a new XPtr object*

Description

Creates a new XPtr of size n and initialize it with the constant k. Objects of class XPtr are just externalptr with attributes. They are used instead of classical dist to avoid copies. This makes RRP to work on bigger data sets.

Usage

```
newXPtr(n, k = 0)
```

Arguments

n	the size of the object
k	the initializing constant

Details

You should think about XPtr as dist objects, i.e. a representation of a symmetric matrix with 0's on the diagonals. XPtr are stored as lower diagonal matrices. The size n is the dimension of the corresponding matrix, say M, hence M is intended as a n * n matrix. The real length of the XPtr is n*(n-1)/2.

Value

returns a new XPtr if there is enough memory.

Author(s)

S.M. Iacus

References

Iacus, S.M., Porro, G. (2006) Random Recursive Partitioning and its applications to missing data imputation, classification and average treatment effect estimation, *submitted*.

See Also

rrp.dist, XPtrToDist

Examples

```
a <- newXPtr(10, 1)
(XPtrToDist(a))
as.dist(matrix(1,10,10))
```

NSW *National Supported Work (NSW) Demonstration data*

Description

National Supported Work (NSW) Demonstration data, of a job training program implemented during the Seventies in the United States and analyzed by Lalonde. This data set also contains data from the CPS (Current Population Survey) and PSID (Panel Study of Income Dynamics) datasets.

Usage

```
data(NSW)
```

Format

A data frame with 19204 observations on the following 13 variables.

`treated` treated variable indicator

`age` age

`education` years of education

`black` race indicator variable

`married` marital status indicator variable

`nodegree` indicator variable of not possessing a degree

`dw` indicator of Dehejia-Wahba selected units

`re74` real earnings in 1974

`re75` real earnings in 1975

`re78` real earnings in 1978 (post treatment outcome)

`hispanic` ethnic indicator variable

`st` indicator variable of Smith-Todd selected units

`sample` number of reference sample: 1 = Lalonde, 2 = CPS and 3 = PSID

Details

Lalonde dataset has 722 experimental treated and control units. CPS has 15992 non experimental control units and PSID has 2490 non experimental control individuals.

Source

see references

References

Lalonde, R. (1986) Evaluating the Econometric Evaluations of Training Programs, *American Economic Review*, 76, 604-620.

rank.dist *rank-based dissimilarity matrix*

Description

Algorithm which generates the rank based proximity matrix.

Usage

```
rank.dist(X, msplit = 2, cut.in = 0,
   thr=0.75, weights, asdist=FALSE, verbose = 0)
```

Arguments

X	a data.frame object
msplit	minimum split parameter
cut.in	number of breaks in which to cut continuous variables
thr	a threshold to keep only the proximities greater than the threshold
weights	a vector of weights
asdist	if TRUE returns an object of class dist otherwise a list of the same length of the data frame
verbose	if greater than 1 some information is printed

Details

This algorithm allows for missing data in X.

Value

an object of class list or dist

Author(s)

S.M. Iacus

See Also

rrp.dist

`rrp.class` *Nonparametric nearest neighbor classification using RRP dissimilarity*

Description

This method is a simple nearest neighbor based on the RRP dissimilarity matrix. It can be used also as a supervised method, i.e. using the class variable in the construction of the RRP dissimilarity matrix and by imputing missing values to the test set.

Usage

```
rrp.class(x, cl, train, test, k = 1)
```

Arguments

`x`	a `XPtr` object
`cl`	class vector, coerced to be of type `factor`
`train`	the vecotr of indexes of the training set
`test`	the vector of training indexes of test set
`k`	number of nearest to consider

Details

From version 1.6 of the package the RRP matrix is stored as an external pointer to avoid duplications. This allow to work on bigger datasets. Hence this function no longer accepts `dist` objects.

Value

a vector of type `factor` with predicted classes.

Author(s)

S.M. Iacus

References

Iacus, S.M., Porro, G. (2009) Random Recursive Partitioning: a matching method for the estimation of the average treatment effect, *Journal of Applied Econometrics*, 24, 163-185.

Iacus, S.M., Porro, G. (2007) Missing data imputation, matching and other applications of random recursive partitioning, *Computational Statistics and Data Analysis*, 52, 2, 773-789.

See Also

`rrp.dist`

Examples

```
data(iris)
X <- iris[,-5]
n <- dim(X)[1]

set.seed(123)
test <- sample(1:n, 10)
train <- (1:n)[-test]

## unsupervised
D <- rrp.dist(X)
pred <- rrp.class(D, iris[train,5], train, test)
table(pred, iris[test, 5])

# supervised
X <- iris
X[test,5] <- NA
D <- rrp.dist(X)
pred <- rrp.class(D, iris[train,5], train, test)
table(pred, iris[test, 5])
```

`rrp.dist` *RRP dissimilarity matrix*

Description

Main piece of the RRP algorithm which generates the RRP dissimilarity matrix.

Usage

```
rrp.dist(X, treated = NULL, msplit = 10, Rep = 250, cut.in = 15,
 check.bal = FALSE, plot = FALSE, asdist = FALSE, verbose = 0)
```

Arguments

`X`	a `data.frame` object
`treated`	optional class indicator variable
`msplit`	minimum split parameter in the `rpart` algorithm
`Rep`	number of RRP replications
`cut.in`	number of breaks in which to cut continuous variables
`check.bal`	indicator function. If `TRUE` balance check using hyper-rectangles will be used inside leaves
`plot`	wheter to plot the porximity matrix as image
`asdist`	if `TRUE` returns an object of class `dist`
`verbose`	if greater than 1 some information is printed

Details

This algorithm allows for missing data in `X`. From version 1.6 of the package the RRP matrix is stored as an external pointer to avoid duplications. This allow to work on bigger datasets.

Value

an object of class `externalptr, XPtr` or `dist`

Author(s)

S.M. Iacus

References

Iacus, S.M., Porro, G. (2009) Random Recursive Partitioning: a matching method for the estimation of the average treatment effect, *Journal of Applied Econometrics*, 24, 163-185.

Iacus, S.M., Porro, G. (2007) Missing data imputation, matching and other applications of random recursive partitioning, *Computational Statistics and Data Analysis*, 52, 2, 773-789.

See Also

`rpart`

---

rrp.impute *Nearest neighbor hot-deck imputation using RRP dissimilarity matrix*

Description

This function performs a simple nearest neighbor hot-deck imputation method using the RRP dissimilarity matrix.

Usage

```
rrp.impute(data, D = NULL, k = 1, msplit = 10, Rep = 250, cut.in = 15)
```

Arguments

`data`	a `data.frame` containing missing data on some covariates
`D`	`NULL` or an object of class `XPtr`
`k`	number of nearest neighbors to use
`msplit`	minimum split parameter in the `rpart` algorithm
`Rep`	number of RRP replications
`cut.in`	number of breaks used to cut continuous covariates

Details

If missing data are on a continuous covariate, the missing value is imputed as the average of the covariate values of the nearest neighbors, otherwise the majority of the 'votes' determines the class of the missing observation on the basis of nearest available data.

If `D` is `NULL` a RRP-dissimilarity matrix is created.

From version 1.6 of the package the RRP matrix is stored as an external pointer to avoid duplications. This allow to work on bigger datasets. Hence this function no longer accepts `dist` objects.

Value

A list

`new.data`	a copy of the data `data` with missing data imputed
`dist`	an object of class `XPtr` used to search for nearest neighbors

Author(s)

S.M. Iacus

References

Iacus, S.M., Porro, G. (2009) Random Recursive Partitioning: a matching method for the estimation of the average treatment effect, *Journal of Applied Econometrics*, 24, 163-185.

Iacus, S.M., Porro, G. (2007) Missing data imputation, matching and other applications of random recursive partitioning, *Computational Statistics and Data Analysis*, 52, 2, 773-789.

See Also

`rrp.dist`, `rrp.class`

Examples

```
data(iris)

X <- iris
n <- dim(X)[1]

set.seed(123)
miss <- sample(1:n, 10)
for(i in miss)
 X[i, sample(1:5, 2)] <- NA

X[miss,]

## unsupervised
x <- rrp.impute(X)

x$new.data[miss,]
iris[miss,]
```

`rrp.predict` *Nonparametric nearest neighbor predictor using RRP dissimilarity*

Description

This method is a simple nearest neighbor based on the RRP dissimilarity matrix. It can be used also as a supervised method, i.e. using the outcome variable in the construction of the RRP dissimilarity matrix and by imputing missing values to the test set.

Usage

```
rrp.predict(x, y, train, test, k = 1)
```

Arguments

`x`	a `XPtr` object
`y`	numeric vector of train outcomes
`train`	the vecotr of indexes of the training set
`test`	the vector of training indexes of test set
`k`	number of nearest to consider

Details

From version 1.6 of the package the RRP matrix is stored as an external pointer to avoid duplications. This allow to work on bigger datasets. Hence this function no longer accepts `dist` objects.

Value

a vector of type `numeric` with predicted outcomes.

Author(s)

S.M. Iacus

References

Iacus, S.M., Porro, G. (2009) Random Recursive Partitioning: a matching method for the estimation of the average treatment effect, *Journal of Applied Econometrics*, 24, 163-185.

Iacus, S.M., Porro, G. (2007) Missing data imputation, matching and other applications of random recursive partitioning, *Computational Statistics and Data Analysis*, 52, 2, 773-789.

See Also

`rrp.dist`, `rrp.class`

Examples

```
require(MASS)
attach(birthwt)

race <- factor(race, labels= c("white", "black", "other"))
ptd <- factor(ptl>0)
ftv <- factor(ftv)
levels(ftv)[-(1:2)] <- "2+"
table(ftv)

bwt <- data.frame(bwt, age, lwt, race, smoke=(smoke>0), ptd,
ht=(ht>0), ui = (ui>0), ftv)

detach()
rm(race, ptd, ftv)

set.seed(123)
n <- dim(bwt)[1]
test <- sample(1:n, 15)
train <- (1:n)[-test]

D <- rrp.dist(bwt[,-1])
true.wht <- bwt$bwt[test]
pred.wht <- rrp.predict(D, bwt$bwt[train], train, test)
mean(pred.wht-true.wht)
sd(pred.wht-true.wht)

mod <- lm(bwt ~ ., data=bwt[train,])
pred.wht <- predict(mod, newdata = bwt[test,])
mean(pred.wht-true.wht)
sd(pred.wht-true.wht)
```

`setDist` *Sets the elements of a dist object to a constant vector-wise*

Description

Sets the elements of a dist object to a constant vector-wise. The idea is that a symmetrix matrix `M` can be represented as a triangular object `d`. With this function you set `M[x,x] = k` but instead of working on `M` this function works on its `dist` representation `d`.

Usage

```
setDist(d, x, k)
```

Arguments

`d`	a dist object, upper diagonal, representing a symmetric matrix
`x`	a list of vectors of indices
`k`	a vector of constants to add, arguments are eventually recycled silently

Value

returns the modified dist object

Author(s)

S. M. Iacus

See Also

dist

Examples

```
M <- matrix(0,5,5)
d <- as.dist(M)
x <- list(1:3,4:5)

setDist(d,x,c(1,2))
# which is the equivalent of (apart for the diagonal elements)
M[1:3,1:3] <- 1
M[4:5,4:5] <- 2


# Indeed, we have
d
as.dist(M)
```

---

setXPtr *Sets the elements of a XPtr object to a constant vector-wise*

Description

Sets the elements of a XPtr object to a constant vector-wise. The idea is that a symmetrix matrix `M` can be represented as a triangular object `d`. With this function you set `M[x,x] = k` but instead of working on `M` this function works on its `dist` representation `d`.

Usage

```
setXPtr(d, x, k)
```

Arguments

d	a `XPtr` object, lower diagonal, representing a symmetric matrix
x	a list of vectors of indices
k	a vector of constants to add, arguments are eventually recycled silently

Value

returns the modified dist object

Author(s)

S. M. Iacus

See Also

dist

Examples

```
M <- matrix(0,5,5)
d <- newXPtr(5, 0)
x <- list(1:3,4:5)

setXPtr(d,x,c(1,2))
# which is the equivalent of (apart for the diagonal elements)
M[1:3,1:3] <- 1
M[4:5,4:5] <- 2


# Indeed, we have
(XPtrToDist(d))
as.dist(M)
```

ST *Smith-Todd dataset*

Description

A subset of the Lalonde dataset (see cited reference).

Usage

```
data(ST)
```

Format

A data frame with 250 observations on the following 10 variables.

`treated` treated variable indicator

`age` age

`education` years of education

`black` race indicator variable

`married` marital status indicator variable

nodegree indicator variable of not possessing a degree

re74 real earnings in 1974

re75 real earnings in 1975

re78 real earnings in 1978 (post treatment outcome)

hispanic ethnic indicator variable

Source

see references

References

Smith, J., Todd, P. (2005a) Does Matching Overcome Lalonde's Critique of Nonexperimental Estimators?, *Journal of Econometrics*, 125(1-2), 305-353.

STvsCPS *Smith-Todd treated units versus CPS control individuals*

Description

The Smith-Todd subset of Lalonde's treated units versus CPS (Current Population Survey) control individuals

Usage

```
data(STvsCPS)
```

Format

A data frame with 16100 observations on the following 10 variables.

treated treated variable indicator

age age

education years of education

black race indicator variable

married marital status indicator variable

nodegree indicator variable of not possessing a degree

re74 real earnings in 1974

re75 real earnings in 1975

re78 real earnings in 1978 (post treatment outcome)

hispanic ethnic indicator variable

Details

These two sets of treated and control units can be hardly matched.

Source

see references

References

Smith, J., Todd, P. (2005a) Does Matching Overcome Lalonde's Critique of Nonexperimental Estimators?, *Journal of Econometrics*, 125(1-2), 305-353.

`STvsPSID` *Smith-Todd treated units versus PSID control individuals*

Description

The Smith-Todd subset of Lalonde's treated units versus PSID (Panel Study of Income Dynamics) control individuals

Usage

```
data(STvsPSID)
```

Format

A data frame with 2598 observations on the following 10 variables.

`treated` treated variable indicator

`age` age

`education` years of education

`black` race indicator variable

`married` marital status indicator variable

`nodegree` indicator variable of not possessing a degree

`re74` real earnings in 1974

`re75` real earnings in 1975

`re78` real earnings in 1978 (post treatment outcome)

`hispanic` ethnic indicator variable

Details

These two sets of treated and control units can be hardly matched.

Source

see references

References

Smith, J., Todd, P. (2005a) Does Matching Overcome Lalonde's Critique of Nonexperimental Estimators?, *Journal of Econometrics*, 125(1-2), 305-353.

XPtrToDist *Coercion from an XPtr object into a dist one*

Description

Coercies an `XPtr` object into a `dist` object which can be used in other algorithms.

Usage

```
XPtrToDist(d)
```

Arguments

`d`	an `XPtr` object

Details

This routine does not duplicate but just makes available as a `dist` the data contained in `d`

Value

returns an invisible object of class `dist`

Author(s)

S.M. Iacus

References

Iacus, S.M., Porro, G. (2009) Random Recursive Partitioning: a matching method for the estimation of the average treatment effect, *Journal of Applied Econometrics*, 24, 163-185.

Iacus, S.M., Porro, G. (2007) Missing data imputation, matching and other applications of random recursive partitioning, *Computational Statistics and Data Analysis*, 52, 2, 773-789.

See Also

`newXPtr`

Examples

```
a <- newXPtr(10,5)
(XPtrToDist(a))
as.dist(matrix(5,10,10))
```

Index