Package 'persDx'

May 30, 2022

Type Package

Title Estimating Personalized Diagnostics Rules

Version 0.1.0

Author Yunro Chung [aut, cre]

Maintainer Yunro Chung <yunro.chung@asu.edu>

Description Recommend the optimal biomarker (or test) for disease screening or diagnosis based on patients' individual characteristics.

Depends R (>= 4.2.0), pROC

License GPL (>= 2)

Encoding UTF-8

NeedsCompilation no

Repository CRAN

Date/Publication 2022-05-30 08:30:02 UTC

R topics documented:

---

persDx-package   *Estimating Personalized Diagnostics Rules*

Description

Recommend the optimal biomarker (or test) for disease screening or diagnosis based on patients' individual characteristics.

Details

1

Package:	persDx
Type:	Package
Version:	0.1.0
Date:	2022-05-26
License:	GPL (>= 2)

Author(s)

Yunro Chung [aut,cre] Maintainer: Yunro Chung <yunro.chung@asu.edu>

References

Yaliang Zhang and Yunro Chung, Estimating individualized diagnostics rules via linear grid search with applications to subgroup-specific biomarker discovery (in progress)

---

lin.persDx	*Estimating linear personalized diagnostic rules.*

Description

Estimate personalized diagnostics rule that recommends biomarker (or test) A versus B by maximizing area under the receiver operating (ROC) curve (AUC) based on linear combinatino of predictors.

Usage

```
lin.persDx(D,YA,YB,X,dirA,dirB,eps,PLOT)
```

Arguments

D	Binary outcome with D=1 for disease (or case) and D=0 for non-diseased (or control) (n X 1 vector).
YA	Biomarker A, measured on a continuous or orinal scale (n X 1 vector).
YB	Biomarker B, measured on a continuous or ordinal scale (n X 1 vector).
X	Predictors (n x p matrix).
dirA	Direction of YA to D, where dirA="<" (or dirA=">") indicates higher (or lower) YA is assoicated with Pr(D=1)). Default is dirA="<".
dirB	Direction of YB to D, where dirB="<" (or dirB=">") indicates higher (or lower) YB is assoicated with Pr(D=1)). Default is dirB="<".
eps	Tuning parameter for predictor selections. Default is eps=0.01.
PLOT	TRUE of FALSE for showing plot.

Details

The lin.persDx function estimates the personalized diagnostics rule $\tau(X)$, where $\tau(X)$=A recommends YA if $\theta_0 + \theta_1 X_1 + \cdots + \theta_p X_p > 0$ or $\tau(X)$=B recommends YB otherwise by maximizing AUC. It guarantess to increase AUC compared to the one-size-fits all strategh. That is, AUC is always greater than maximum of AUC.A and AUC.B, where AUC.A is an empirical AUC value using (D,YA), AUC.B is an empirical AUC value using (D,YB), AUC is an empirical AUC value using (D,YC), where YC=YA if $\tau(X)$=A and YC=YB if $\tau(X)$=B. Since the empirical AUC is not a continuous function on $(\theta_0, \theta_1, \ldots, \theta_p)$, a grid-based search algorithm is used. Specifically, for a univariate predictor $X = X_1$, a grid search algorithm is used to estimate (θ_0, θ_1), and for multiple predictors $X = (X_1, X_2, \ldots, X_p)$, a forward selection algorithm is further implemented that sequentially adds each of the predictors to $\tau(X)$ that increases the AUC most. The stopping criteria is AUC increasement $\leq$ eps. It allows eps=0, which searches all X, while some of θ can be estimated to zeros when corresponding predictors have no contribution to increase the AUC.

Value

A list of class lin.persDx:

df	Data frame with YA, YB, X, and estimated subgroup (A or B).
AUC.A	AUC for YA.
AUC.B	AUC for YB.
AUC	AUC for YA or YB based on $\tau(X)$.
theta	$\theta_0, \theta_1, \ldots, \theta_p$.

Author(s)

Yunro Chung [aut, cre]

References

Estimation of linear decision rules for personalized diagnostics using area under the receiver operating characteristics curve (in progress)

Examples

```
set.seed(1)
n1=n0=100
n=n1+n0
D=c(rep(1,n1),rep(0,n0))
X1=runif(n,0,1)
X2=runif(n,0,1)
X3=runif(n,0,1)
X=data.frame(X1,X2,X3)
muA=2
muB=1
mu0=0
sigma=1

tau=rep("B",n)
```

```
tau[X1+X2>=1]="A"

YA=D*(rnorm(n,muA,sigma)*(tau=="A")+rnorm(n,mu0,sigma)*(tau=="B"))+(1-D)*rnorm(n,mu0,sigma)
YB=D*(rnorm(n,muB,sigma)*(tau=="B")+rnorm(n,mu0,sigma)*(tau=="A"))+(1-D)*rnorm(n,mu0,sigma)

fit=lin.persDx(D, YA, YB, X)
fit
```

Index