# Package ‘MCID’

**Type** Package

**Title** Estimating the Minimal Clinically Important Difference

**Version** 0.1.0

**Date** 2021-09-07

**Description** Apply the marginal classification method to achieve the purpose of providing the point and interval estimates for the minimal clinically important difference based on the classical anchor-based method. For more details of the methodology, please see Zehua Zhou, Leslie J. Bisson and Jiwei Zhao (2021) <arXiv:2108.11589>.

**License** GPL (>= 2)

**Encoding** UTF-8

**Imports** stats

**RoxygenNote** 7.1.0

**NeedsCompilation** no

**Author** Zehua Zhou [cre, aut], Jiwei Zhao [aut]

**Maintainer** Zehua Zhou <zehuazho@buffalo.edu>

**Repository** CRAN

**Date/Publication** 2021-09-10 11:20:12 UTC

## R topics documented:

<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>cv.imcid</td>
<td>2</td>
</tr>
<tr>
<td>cv.pmcid</td>
<td>3</td>
</tr>
<tr>
<td>imcid</td>
<td>4</td>
</tr>
<tr>
<td>pmcid</td>
<td>5</td>
</tr>
</tbody>
</table>

**Index** 7
cv.imcid

Selection of the tuning parameters for determining the MCID at the individual level

Description

cv.imcid returns the optimal tuning parameter $\delta$ and $\lambda$ selected from a given grid by using k-fold cross-validation. The tuning parameters are selected for determining the MCID at the individual level.

Usage

cv.imcid(x, y, z, lamseq, delseq, k = 5, maxit = 100, tol = 0.01)

Arguments

- x: a continuous variable denoting the outcome change of interest
- y: a binary variable denoting the patient-reported outcome derived from the anchor question
- z: a vector or matrix denoting the patient’s clinical profiles
- lamseq: a vector containing the candidate values for the tuning parameter $\lambda$, where $\lambda$ is the coefficient of the penalty term, used for avoiding the issue of model overfitting
- delseq: a vector containing the candidate values for the tuning parameter $\delta$, where $\delta$ is used to control the difference between the 0-1 loss and the surrogate loss. We recommend selecting the possible values from the neighborhood of the standard deviation of x
- k: the number of groups into which the data should be split to select the tuning parameter $\delta$ by cross-validation. Defaults to 5
- maxit: the maximum number of iterations. Defaults to 100
- tol: the convergence tolerance. Defaults to 0.01

Value

A list including the combinations of the selected tuning parameters and the value of the corresponding target function

Examples

n <- 500
lambdaseq <- 10 ^ seq(-3, 3, 0.1)
deltaseq <- seq(0.1, 0.3, 0.1)
a <- 0.1
b <- 0.55
c <- -0.1
\begin{verbatim}
d <- 0.45
set.seed(721)
p <- 0.5
y <- 2 * rbinom(n, 1, p) - 1
z <- rnorm(n, 1, 0.1)
y_1 <- which(y == 1)
y_0 <- which(y == -1)
x <- 
\end{verbatim}

\begin{verbatim}
x[y_1] <- a + z[y_1] * b + rnorm(length(y_1), 0, 0.1)
x[y_0] <- c + z[y_0] * d + rnorm(length(y_0), 0, 0.1)
\end{verbatim}

\begin{verbatim}
 sel <- cv.pmcid(x = x, y = y, delseq = lambdaseq,
                   delseq = deltaseq, k = 5, maxit = 100, tol = 1e-02)
 sel$'Selected lambda'
 sel$'Selected delta'
\end{verbatim}

---

**cv.pmcid**

*Selection of the tuning parameter for determining the MCID at the population level*

**Description**

`cv.pmcid` returns the optimal tuning parameter $\delta$ selected from a given grid by using k-fold cross-validation. The tuning parameter is selected for determining the MCID at the population level

**Usage**

\[
\text{cv.pmcid}(x, y, delseq, k = 5, maxit = 100, tol = 0.01)
\]

**Arguments**

- `x`: a continuous variable denoting the outcome change of interest
- `y`: a binary variable indicating the patient-reported outcome derived from the anchor question
- `delseq`: a vector containing the candidate values for the tuning parameter $\delta$, where $\delta$ is used to control the difference between the 0-1 loss and the surrogate loss. We recommend selecting the possible values from the neighborhood of the standard deviation of x
- `k`: the number of groups into which the data should be split to select the tuning parameter $\delta$ by cross-validation. Defaults to 5
- `maxit`: the maximum number of iterations. Defaults to 100
- `tol`: the convergence tolerance. Defaults to 0.01
Value

a list including the selected tuning parameter and the value of the corresponding target function

Examples

```r
n <- 500
deltaseq <- seq(0.1, 1, 0.1)
a <- 0.2
b <- -0.1
p <- 0.5

set.seed(115)
y <- 2 * rbinom(n, 1, p) - 1
y_1 <- which(y == 1)
y_0 <- which(y == -1)
x <- c()
x[y_1] <- rnorm(length(y_1), a, 0.1)
x[y_0] <- rnorm(length(y_0), b, 0.1)

sel <- cv.pmcid(x = x, y = y, delseq = deltaseq, k = 5,
               maxit = 100, tol = 1e-02)
sel$'Selected delta'
sel$'Function value'
```

imcid

Point and interval estimation for the MCID at the individual level

Description

We formulate the individualized MCID as a linear function of the patients’ clinical profiles. `imcid` returns the point estimate for the linear coefficients of the MCID at the individual level

Usage

`imcid(x, y, z, n, lambda, delta, maxit = 100, tol = 0.01, alpha = 0.05)`

Arguments

- `x`: a continuous variable denoting the outcome change of interest
- `y`: a binary variable indicating the patient-reported outcome derived from the anchor question
- `z`: a vector or matrix denoting the patient’s clinical profiles
- `n`: the sample size
- `lambda`: the selected tuning parameter $\lambda$, can be returned by `cv.imcid`
- `delta`: the selected tuning parameter $\delta$, can be returned by `cv.imcid`
- `maxit`: the maximum number of iterations. Defaults to 100
pmcid

**tol**  
the convergence tolerance. Defaults to 0.01

**alpha**  
nominal level of the confidence interval. Defaults to 0.05

**Value**

a list including the point estimates for the linear coefficients of the individualized MCID and their standard errors, and the corresponding confidence intervals based on the asymptotic normality

**Examples**

```r
n <- 500
lambdaseq <- 10 ^ seq(-3, 3, 0.1)
deltaseq <- seq(0.1, 0.3, 0.1)
a <- 0.1
b <- 0.55
c <- -0.1
d <- 0.45
### True linear coefficients of the individualized MCID: ###
### beta0=0, beta1=0.5 ###
set.seed(115)
p <- 0.5
y <- 2 * rbinom(n, 1, p) - 1
z <- rnorm(n, 1, 0.1)
y_1 <- which(y == 1)
y_0 <- which(y == -1)
x <- c()
x[y_1] <- a + z[y_1] * b + rnorm(length(y_1), 0, 0.1)
x[y_0] <- c + z[y_0] * d + rnorm(length(y_0), 0, 0.1)
result <- imcid(x = x, y = y, z = z, n = n, lambda = lamsel, delta = delsel, maxit = 100, tol = 1e-02, alpha = 0.05)
```

**pmcid**

Point and interval estimation for the MCID at the population level

**Description**

pmcid returns the point estimate for the MCID at the population level
Usage

\texttt{pmcid(x, y, n, delta, maxit = 100, tol = 0.01, alpha = 0.05)}

Arguments

- \texttt{x}: a continuous variable denoting the outcome change of interest
- \texttt{y}: a binary variable indicating the patient-reported outcome derived from the anchor question
- \texttt{n}: the sample size
- \texttt{delta}: the selected tuning parameter $\delta$, can be returned by \texttt{cv.pmcid}
- \texttt{maxit}: the maximum number of iterations. Defaults to 100
- \texttt{tol}: the convergence tolerance. Defaults to 0.01
- \texttt{alpha}: nominal level of the confidence interval. Defaults to 0.05

Value

A list including the point estimate of the population MCID and its standard error, and the confidence interval based on the asymptotic normality.

Examples

```r
n <- 500
deltaseq <- seq(0.1, 1, 0.1)
a <- 0.2
b <- -0.1
p <- 0.5
### True MCID is 0.5 ###
set.seed(115)
y <- 2 * rbinom(n, 1, p) - 1
y_1 <- which(y == 1)
y_0 <- which(y == -1)
x <- c()
x[y_1] <- rnorm(length(y_1), a, 0.1)
x[y_0] <- rnorm(length(y_0), b, 0.1)

sel <- cv.pmcid(x = x, y = y, delseq = deltaseq, k = 5,
                maxit = 100, tol = 1e-02)
delsel <- sel$'Var Selected delta'

result <- pmcid(x = x, y = y, n = n, delta = delsel,
                maxit = 100, tol = 1e-02, alpha = 0.05)
result$'Var Point estimate'
result$'Var Standard error'
result$'Var Confidence interval'
```
Index

cv.imcid, 2

cv.pmcid, 3

imcid, 4

pmcid, 5