Package ‘mri’

July 22, 2015

Type Package

Title Modified Rand Index (1 and 2.1 and 2.2) and Modified Adjusted Rand Index (1 and 2.1 and 2.2)

Version 0.1.1

Date 2015-07-20

Author Marjan Cugmas

Maintainer Marjan Cugmas <marjan.cugmas@fdv.uni-lj.si>

Description Provides three Modified Rand Indices and three Modified Adjusted Rand Indices for comparing two partitions, which are usually obtained on two different sets of units, where one is a subset of another set of units. Splitting and merging of clusters have a different affects on the value of the indices.

License GPL (>= 2)

LazyData TRUE

NeedsCompilation no

Repository CRAN

Date/Publication 2015-07-22 16:15:48

R topics documented:

mari1 .......................................................... 2
mari21 ........................................................ 4
mari22 ........................................................ 6
mri1 .......................................................... 8
mri21 ......................................................... 10
mri22 ......................................................... 12

Index 15
mari1

*Modified Adjusted Rand Index 1*

**Description**

mari1 is used to compute the value of the Modified Adjusted Rand Index 1. It considers two partitions $U$ and $V'$, which are usually obtained on two sets of units $S$ and $T$, where $T$ is a subset of $S$. Because two vectors $U$ and $V$ do not have the same length, the cluster of units, which are not present in the partition $V$, need to be added to the partition $V$ (denoted as $V''$). This can be done manually or with function `merge`.

**Usage**

```r
mari1(U, V, outgoing.name, k = 500)
```

**Arguments**

- `U`: Partition $U$.
- `V`: Partition $V'$.
- `outgoing.name`: The label of added cluster (see Description). Do not specify this argument if there is no added values.
- `k`: Number of permutations.

**Value**

The function returns the value of the Modified Adjusted Rand Index 1. The expected value of the index in the case of two random and independant partitions is 0. The maximum value of the index is 1. Higher value indicates more similar (stable) partitions. The splitting of clusters lowers the value of the index, but the merging of clusters does not affect the value of the index.

**Author(s)**

Marjan Cugmas

**References**


**See Also**

Other types of Modified Rand Index: `mriR1`, `mriRR`.
Modified Adjusted Rand Index: `mari1`, `mari21`, `mari22`. 
Examples

# THIS FUNCTION CONVERT THE CONTINGENCY TABLE INTO TWO VECTORS
fromTableToVectors <- function(cont.table){
  dat <- matrix(ncol = 2)
  for (j in 1:dim(cont.table)[2]){
    for (i in 1:dim(cont.table)[1]){
      if (cont.table[i, j] > 1){  
        for (t in 1:cont.table[i,j]){  
          imei <- rownames(cont.table)[i]  
          imej <- colnames(cont.table)[j]  
          dat <- rbind(dat, c(imei, imej))  
        }
      }
    }
  }
  dat <- dat[-1,-1]
  return(dat)
}

# DATA
example.Model <- list(NULL)

# example 1
cont.table <- rbind(c(0, 10, 0, 0, 0),
                    c(10, 0, 0, 0, 0),
                    c(0, 0, 10, 0, 0),
                    c(0, 0, 0, 5, 5))
rownames(cont.table) <- c(1:4)
colnames(cont.table) <- c(1:5)
example.Model[[1]] <- cont.table

# example 2
cont.table <- rbind(c(0, 10, 0, 0, 0),
                    c(0, 10, 0, 0, 0),
                    c(0, 0, 10, 0, 0),
                    c(0, 0, 0, 5, 5))
rownames(cont.table) <- c(1:4)
colnames(cont.table) <- c(1:5)
example.Model[[2]] <- cont.table

# example 3
cont.table <- rbind(c(0, 0, 0, 0, 0),
                    c(10, 10, 0, 0, 0),
                    c(0, 0, 10, 0, 0),
                    c(0, 0, 0, 5, 5))
rownames(cont.table) <- c(1:4)
colnames(cont.table) <- c(1:5)
example.Model[[3]] <- cont.table

# COMPUTE MRI
# example 1
U <- fromTableToVectors(example.Model[[1]]),[1]
V <- fromTableToVectors(example.Model[[1]]),[2]
mari1(U = U, V = V, outgoing.name = 4, k = 500)

# example 2
U <- fromTableToVectors(example.Model[[2]]),[1]
V <- fromTableToVectors(example.Model[[2]]),[2]
mari1(U = U, V = V, outgoing.name = 4, k = 500)

# example 3
U <- fromTableToVectors(example.Model[[3]]),[1]
V <- fromTableToVectors(example.Model[[3]]),[2]
mari1(U = U, V = V, outgoing.name = 4, k = 500)

Description

mari21 is used to compute the value of the Modified Adjusted Rand Index 2.1. It considers two partitions $U'$ and $V$, which are usually obtained on two sets of units $S$ and $T$, where $S$ is a subset of $T$. Because two vectors $U$ and $V$ do not have the same length, the cluster of units, which are not present in the partition $U$, need to be added to the partition $U$ (denoted as $U'$). This can be done manually or with function merge.

Usage

mari21(U, V, incoming.name, k = 500)

Arguments

- **U**: Partition $U'$.
- **V**: Partition $V$.
- **incoming.name**: The label of added cluster (see Description). Do not specify this argument if there is no added values.
- **k**: Number of permutations.

Value

The function returns the value of the Modified Adjusted Rand Index 2.1. The expected value of the index in the case of two random and independant partitions is 0. The maximum value of the index is 1. Higher value indicates more similar (stable) partitions. The splitting of clusters lowers the value of the index, but the merging of clusters does not affect the value of the index. The difference in the number of units between two sets of units $S$ and $T$ lower the value of the index.

Author(s)

Marjan Cugmas
References


See Also

Other types of Modified Rand Index: *mri1, mri21, mri22.*

Modified Adjusted Rand Index: *mari1, mari22.*

Examples

```r
# THIS FUNCTION CONVERT THE CONTINGENCY TABLE INTO TWO VECTORS
fromTableToVectors <- function(cont.table){
  dat <- matrix(ncol = 2)
  for (j in 1:dim(cont.table)[2]){  
    for (i in 1:dim(cont.table)[1]){  
      if (cont.table[i, j] > 1){
        for (t in 1:cont.table[i,j]){  
          imei <- rownames(cont.table)[i]
          imej <- colnames(cont.table)[j]
          dat <- rbind(dat, c(imei, imej))
        }
      }
    }
  }
  dat <- dat[-1,]
  return(dat)
}

# DATA
example.Model <- list(NULL)

tabela <- rbind(c(0, 10, 0, 0, 0),
  c(10, 0, 0, 0, 0),
  c(0, 0, 10, 0, 0),
  c(0, 0, 0, 5, 5))
rownames(tabela) <- c("1", "2", "3", "4")
colnames(tabela) <- c("1", "2", "3", "4", "5")
example.Model[[1]] <- t(tabela)

tabela <- rbind(c(0, 10, 0, 0, 0),
  c(0, 10, 0, 0, 0),
  c(0, 0, 10, 0, 0),
  c(0, 0, 0, 5, 5))
rownames(tabela) <- c("1", "2", "3", "4")
colnames(tabela) <- c("1", "2", "3", "4", "5")
example.Model[[2]] <- t(tabela)

tabela <- rbind(c(0, 0, 0, 0, 0),
  c(10, 10, 0, 0, 0),
  c(0, 0, 5, 0, 5),
  c(0, 0, 0, 5, 5))
rownames(tabela) <- c("1", "2", "3", "4")
```


colnames(tabela) <- c("1", "2", "3", "4", "5")
example.Model[[3]] <- t(tabela)

# COMPUTE MRI22

# example 1
U <- fromTableToVectors(example.Model[[1]]),[,1]
V <- fromTableToVectors(example.Model[[1]]),[,2]
mari22(U = U, V = V, incoming.name = 4, k = 500)

# example 2
U <- fromTableToVectors(example.Model[[2]]),[,1]
V <- fromTableToVectors(example.Model[[2]]),[,2]
mari22(U = U, V = V, incoming.name = 4, k = 500)

# example 3
U <- fromTableToVectors(example.Model[[3]]),[,1]
V <- fromTableToVectors(example.Model[[3]]),[,2]
mari22(U = U, V = V, incoming.name = 5, k = 500)

mari22

**Modified Adjusted Rand Index 2.2**

**Description**

mari22 is used to compute the value of the Modified Adjusted Rand Index 2.2. It considers two partitions $U'$ and $V$, which are usually obtained on two sets of units $S$ and $T$, where $S$ is a subset of $T$. Because two vectors $U$ and $V$ do not have the same length, the cluster of units, which are not present in the partition $U$, need to be added to the partition $U$ (denoted as $U'$). This can be done manually or with function `merge`.

**Usage**

mari22(U, V, incoming.name, k = 500)

**Arguments**

- **U**
  - Partition $U'$.
- **V**
  - Partition $V$.
- **incoming.name**
  - The label of added cluster (see Description). Do not specify this argument if there is no added values.
- **k**
  - Number of permutations.

**Value**

The function returns the value of the Modified Adjusted Rand Index 2.2. The expected value of the index in the case of two random and independent partitions is 0. The maximum value of the index is 1. Higher value indicates more similar (stable) partitions. While the splitting of clusters lowers the value of the index, the merging of clusters does not affect the value of the index. The difference in the number of units between two sets of units $S$ and $T$ does not lower the value of the index.
Author(s)
Marjan Cugmas

References

See Also
Other types of Modified Rand Index: mri1, mri21, mri22.
Modified Adjusted Rand Index: mari1, mari21.

Examples

# THIS FUNCTION CONVERT THE CONTINGENCY TABLE INTO TWO VECTORS
fromTableToVectors <- function(cont.table){
  dat <- matrix(ncol = 2)
  for (j in 1:dim(cont.table)[2]){  
    for (i in 1:dim(cont.table)[1]){ 
      if (cont.table[i, j] > 1){ 
        for (t in 1:cont.table[i,j]){ 
          imei <- rownames(cont.table)[i]
          imej <- colnames(cont.table)[j]
          dat <- rbind(dat, c(imei, imej))
        }
      }
    }
  }
  dat <- dat[-1,]
  return(dat)
}

# DATA
example.Model <- list(NULL)

tabela <- rbind(c(0, 10, 0, 0, 0),
    c(10, 0, 0, 0, 0),
    c(0, 0, 10, 0, 0),
    c(0, 0, 0, 5, 5))
rownames(tabela) <- c("1", "2", "3", "4")
colnames(tabela) <- c("1", "2", "3", "4", "5")
example.Model[[1]] <- t(tabela)

tabela <- rbind(c(0, 10, 0, 0, 0),
    c(0, 10, 0, 0, 0),
    c(0, 0, 10, 0, 0),
    c(0, 0, 0, 5, 5))
rownames(tabela) <- c("1", "2", "3", "4")
colnames(tabela) <- c("1", "2", "3", "4", "5")
example.Model[[2]] <- t(tabela)
```r
# COMPUTE MRI12

# example 1
U <- fromTableToVectors(example.Model[[1]]), , 1
V <- fromTableToVectors(example.Model[[1]]), , 2
mari1(U = U, V = V, incoming.name = 4, k = 500)

# example 2
U <- fromTableToVectors(example.Model[[2]]), , 1
V <- fromTableToVectors(example.Model[[2]]), , 2
mari1(U = U, V = V, incoming.name = 4, k = 500)

# example 3
U <- fromTableToVectors(example.Model[[3]]), , 1
V <- fromTableToVectors(example.Model[[3]]), , 2
mari1(U = U, V = V, incoming.name = 5, k = 500)
```

---

**mri1**  
*Modified Rand Index 1*

**Description**

*mri1* is used to compute the value of the Modified Rand Index 1. It considers two partitions *U* and *V′*, which are usually obtained on two sets of units *S* and *T*, where *T* is a subset of *S*. Because two vectors *U* and *V* do not have the same length, the cluster of units, which are not present in the partition *V*, need to be added to the partition *V* (denoted as *V′*). This can be done manually or with function *merge*.

**Usage**

*mri1*(U, V, outgoing.name)

**Arguments**

- **U**: Partition *U*.
- **V**: Partition *V′*.
- **outgoing.name**: The label of added cluster (see Description). Do not specify this argument if there is no added values.
Value

The function returns the value of the Modified Rand Index. The index can take the value on the interval $[0, 1]$, where higher value indicates more similar (stable) partitions. While the splitting of clusters lowers the value of the index, the merging of clusters does not affect the value of the index.

Author(s)

Marjan Cugmas

References


See Also

Other types of Modified Rand Index: mri21, mri22.
Modified Adjusted Rand Index: mari1, mari21, mari22.

Examples

```r
# THIS FUNCTION CONVERT THE CONTINGENCY TABLE INTO TWO VECTORS
fromTableToVectors <- function(cont.table){
  dat <- matrix(ncol = 2)
  for (j in 1:dim(cont.table)[2]){
    for (i in 1:dim(cont.table)[1]){
      if (cont.table[i, j] > 1){
        for (t in 1:cont.table[i,j]){imatrix}}
      imei <- rownames(cont.table)[i]
      imej <- colnames(cont.table)[j]
      dat <- rbind(dat, c(imei, imej))
    }
  }
  dat <- dat[-1]  # remove the last row
  return(dat)
}

# DATA
example.Model <- list(NULL)

# example 1
cont.table <- rbind(c(0, 10, 0, 0, 0),
  c(0, 0, 0, 0, 0),
  c(0, 0, 10, 0, 0),
  c(0, 0, 0, 0, 5))
rownames(cont.table) <- c(1:4)
colnames(cont.table) <- c(1:5)
example.Model[[1]] <- cont.table

# example 2
\[ \text{cont.table} \leftarrow \text{rbind}(c(0, 10, 0, 0, 0), \\
 c(0, 10, 0, 0, 0), \\
 c(0, 0, 10, 0, 0), \\
 c(0, 0, 0, 5, 5)) \]

\[ \text{rownames(cont.table)} \leftarrow c(1:4) \]

\[ \text{colnames(cont.table)} \leftarrow c(1:5) \]

\[ \text{example.Model}[2] \leftarrow \text{cont.table} \]

\[ \text{# example 3} \]
\[ \text{cont.table} \leftarrow \text{rbind}(c(0, 0, 0, 0, 0), \\
 c(10, 10, 0, 0, 0), \\
 c(0, 0, 10, 0, 0), \\
 c(0, 0, 0, 5, 5)) \]

\[ \text{rownames(cont.table)} \leftarrow c(1:4) \]

\[ \text{colnames(cont.table)} \leftarrow c(1:5) \]

\[ \text{example.Model}[3] \leftarrow \text{cont.table} \]

\[ \text{# COMPUTE MRI1} \]
\[ \text{# example 1} \]
\[ U \leftarrow \text{fromTableToVectors(example.Model}[1])[,1] \]
\[ V \leftarrow \text{fromTableToVectors(example.Model}[2])[,2] \]
\[ \text{mri1}(U = U, V = V, \text{outgoing.name} = 4) \]

\[ \text{# example 2} \]
\[ U \leftarrow \text{fromTableToVectors(example.Model}[2])[,1] \]
\[ V \leftarrow \text{fromTableToVectors(example.Model}[2])[,2] \]
\[ \text{mri1}(U = U, V = V, \text{outgoing.name} = 4) \]

\[ \text{# example 3} \]
\[ U \leftarrow \text{fromTableToVectors(example.Model}[3])[,1] \]
\[ V \leftarrow \text{fromTableToVectors(example.Model}[3])[,2] \]
\[ \text{mri1}(U = U, V = V, \text{outgoing.name} = 4) \]

---

\textbf{mri21} \hspace{1cm} \textit{Modified Rand Index 2.1}

\textbf{Description}

\texttt{mri21} is used to compute the value of the Modified Rand Index 2.1. It considers two partitions \(U'\) and \(V\), which are usually obtained on two sets of units \(S\) and \(T\), where \(S\) is a subset of \(T\). Because two vectors \(U\) and \(V\) do not have the same length, the cluster of units, which are not present in the partition \(U\), need to be added to the partition \(U\) (denoted as \(U'\)). This can be done manually or with function \text{merge}.

\textbf{Usage}

\texttt{mri21(U, V, incoming.name) \hspace{1cm} \textit{Modified Rand Index 2.1}}
Arguments

U  Partition U'.

V  Partition V.

incoming.name  The label of added cluster (see Description). Do not specify this argument if there is no added values.

Value

The function returns the value of the Modified Rand Index 2.1. The index can take the value on the interval \([0, 1]\), where higher value indicates more similar (stable) partitions. While the splitting of clusters lowers the value of the index, the merging of clusters does not affect the value of the index. The difference in the number of units between two sets of units \(S\) and \(T\) lower the value of the index.

Author(s)

Marjan Cugmas

References


See Also

Other types of Modified Rand Index: mri1, mri22.

Modified Adjusted Rand Index: mari1, mari21, mari22.

Examples

```r
# THIS FUNCTION CONVERT THE CONTINGENCY TABLE INTO TWO VECTORS
fromTableToVectors <- function(cont.table){
  dat <- matrix(ncol = 2)
  for (j in 1:dim(cont.table)[2]){  
    for (i in 1:dim(cont.table)[1]){  
      if (cont.table[i, j] >= 1) {  
        for (t in 1:cont.table[i,j]){  
          imei <- rownames(cont.table)[i]  
          imej <- colnames(cont.table)[j]  
          dat <- rbind(dat, c(imei, imej))  
        }
      }
    }
  }
  dat <- dat[-1,]
  return(dat)
}

# DATA
eexample.Model <- list(NULL)
```
tabela <- rbind(c(0, 10, 0, 0, 0),
c(10, 0, 0, 0, 0),
c(0, 0, 10, 0, 0),
c(0, 0, 0, 5, 5))
rownames(tabela) <- c("1", "2", "3", "4")
colnames(tabela) <- c("1", "2", "3", "4", "5")
exemplo.Model[[1]] <- t(tabela)

# COMPUTE MRI22

# example 1
U <- fromTableToVectors(exemplo.Model[[1]][,1])
V <- fromTableToVectors(exemplo.Model[[1]][,2])
mri22(U = U, V = V, incoming.name = 4)

# example 2
U <- fromTableToVectors(exemplo.Model[[2]][,1])
V <- fromTableToVectors(exemplo.Model[[2]][,2])
mri22(U = U, V = V, incoming.name = 4)

# example 3
U <- fromTableToVectors(exemplo.Model[[3]][,1])
V <- fromTableToVectors(exemplo.Model[[3]][,2])
mri22(U = U, V = V, incoming.name = 5)

---

**Modified Rand Index 2.2**

**Description**

mri22 is used to compute the value of the Modified Rand Index 2.2. It considers two partitions \(U'\) and \(V\), which are usually obtained on two sets of units \(S\) and \(T\), where \(S\) is a subset of \(T\). Because two vectors \(U\) and \(V\) do not have the same length, the cluster of units, which are not present in the partition \(U\), need to be added to the partition \(U\) (denoted as \(U'\)). This can be done manually or with function `merge`. 
Usage

mri22(U, V, incoming.name)

Arguments

U Partition U’.
V Partition V.
incoming.name The label of added cluster (see Description). Do not specify this argument if there is no added values.

Value

The function return the value of the Modified Rand Index 2.2. The index can take the value on the interval \([0, 1]\), where higher value indicates more similar (stable) partitions. While the splitting of clusters lowers the value of the index, the merging of clusters does not affect the value of the index. The difference in the number of units between two sets of units \(S\) and \(T\) does not lower the value of the index.

Author(s)

Marjan Cugmas

References


See Also

Other types of Modified Rand Index: mri1, mri21.
Modified Adjusted Rand Index: mari1, mari21, mari22.

Examples

# THIS FUNCTION CONVERT THE CONTINGENCY TABLE INTO TWO VECTORS
fromTableToVectors <- function(cont.table){
  dat <- matrix(ncol = 2)
  for (j in 1:dim(cont.table)[2]){
    for (i in 1:dim(cont.table)[1]){
      if (cont.table[i, j] >= 1){
        for (t in 1:cont.table[i,j]){
          imei <- rownames(cont.table)[i]
          imej <- colnames(cont.table)[j]
          dat <- rbind(dat, c(imei, imej))
        }
      }
    }
  }
  dat <- dat[-1,]
  return(dat)
}
# DATA
example.Model <- list(NULL)

tabela <- rbind(c(0, 10, 0, 0, 0),
                 c(10, 0, 0, 0, 0),
                 c(0, 0, 10, 0, 0),
                 c(0, 0, 0, 5, 5))
rownames(tabela) <- c("1", "2", "3", "4")
colnames(tabela) <- c("1", "2", "3", "4", "5")
example.Model[[1]] <- t(tabela)

tabela <- rbind(c(0, 10, 0, 0, 0),
                 c(0, 10, 0, 0, 0),
                 c(0, 0, 10, 0, 0),
                 c(0, 0, 0, 5, 5))
rownames(tabela) <- c("1", "2", "3", "4")
colnames(tabela) <- c("1", "2", "3", "4", "5")
example.Model[[2]] <- t(tabela)

tabela <- rbind(c(0, 0, 0, 0, 0),
                 c(10, 10, 0, 0, 0),
                 c(0, 0, 0, 5, 5),
                 c(0, 0, 0, 5, 5))
rownames(tabela) <- c("1", "2", "3", "4")
colnames(tabela) <- c("1", "2", "3", "4", "5")
example.Model[[3]] <- t(tabela)

# COMPUTE MRI22

# example 1
U <- fromTableToVectors(example.Model[[1]]),[,1]
V <- fromTableToVectors(example.Model[[1]]),[,2]
mri22(U = U, V = V, incoming.name = 4)

# example 2
U <- fromTableToVectors(example.Model[[2]]),[,1]
V <- fromTableToVectors(example.Model[[2]]),[,2]
mri22(U = U, V = V, incoming.name = 4)

# example 3
U <- fromTableToVectors(example.Model[[3]]),[,1]
V <- fromTableToVectors(example.Model[[3]]),[,2]
mri22(U = U, V = V, incoming.name = 5)
Index

mari1, 2, 2, 5, 7, 9, 11, 13
mari21, 2, 4, 7, 9, 11, 13
mari22, 2, 5, 6, 9, 11, 13
merge, 2, 4, 6, 8, 10, 12
mri1, 5, 7, 8, 11, 13
mri21, 2, 5, 7, 9, 10, 13
mri22, 2, 5, 7, 9, 11, 12