Package ‘PhViD’

November 3, 2016

Type Package

Title PharmacoVigilance Signal Detection

Version 1.0.8

Date 2016-11-03

Author Ismail Ahmed & Antoine Poncet

Maintainer Ismail Ahmed <ismail.ahmed@inserm.fr>

Depends R (>= 2.10), LBE, MCMCpack

Description A collection of several pharmacovigilance signal detection methods extended to the multiple comparison setting.

License GPL-2

Encoding UTF-8

NeedsCompilation no

SystemRequirements Tcl/Tk package TkTable.

Repository CRAN

Date/Publication 2016-11-03 12:14:08

R topics documented:

PhViD-package .................................................. 2
as.PhViD ...................................................... 3
BCPNN ......................................................... 3
GPS .......................................................... 5
PhViD.search .................................................. 7
PhViDdata.frame ............................................. 8
PRR .......................................................... 9
RFET .......................................................... 10
ROR .......................................................... 12

Index 14
**Description**

The PhViD-package proposes the main pharmacovigilance signal detection methods extended to the multiple comparison setting. For the frequentist methods, the package requires the LBE procedure that is stored in the Bioconductor website [http://bioconductor.org/](http://bioconductor.org/). LBE can be installed by entering

```r
source("http://bioconductor.org/biocLite.R")
biocLite("LBE")
```

in the R console.

**Author(s)**

Ismaïl Ahmed & Antoine Poncet

Maintainer: Ismaïl Ahmed <ismail.ahmed@inserm.fr>

**References**


DuMouchel W. Bayesian Data Mining in Large Frequency Tables, with an Application to the FDA Spontaneous Reporting System. The American Statistician. 1999, 53. 177-190.

Evans SJ, Waller PC, Davis S, Use of Proportional Reporting Ratios (PRRs) for Signal Generation from Spontaneous Adverse Drug Reaction Reports Pharmacoepidemiology and Drug Safety, 2001, 10, 483-486.

Noren, GN, Bate A, Orre R, Edwards IR. Extending the methods used to screen the WHO drug safety database towards analysis of complex associations and improved accuracy for rare events Statistics in Medicine, 2006, 25, 3740-3757.

**as.PhViD**

_data.frame to PhViD data_

### Description

as.PhViD is a function that converts a data.frame into an object that can be used in the signal detection method functions.

### Usage

```r
as.PhViD(DATA.FRAME, MARGIN.THRES = 1)
```

### Arguments

- **DATA.FRAME**
  - The data.frame has to be structured as follows
    - 1st column: label of the drugs
    - 2nd column: label of the adverse events
    - 3rd column: Number of spontaneous reports of the corresponding couple drug-adverse event.

- **MARGIN.THRES**
  - This option can be used to eliminate the drugs and the adverse events for which the marginal counts are less than MARGIN.THRES.

### Value

- L
  - data.frame that contains the labels of the drugs and the adverse events.

- N
  - sum of the spontaneous reports counts.

- data
  - data.frame that contains the number of spontaneous reports (n11) and the corresponding marginal counts as well (n1 and n.1).

### Author(s)

Ismail Ahmed & Antoine Poncet

---

**BCPNN**

_Bayesian confidence propagation neural network_

### Description

Bayesian confidence propagation neural network (Bate et al. 1998, Noren et al. 2006) extended to the multiple comparison framework.

### Usage

```r
BCPNN(DATABASE, RR0 = 1, MIN.n11 = 1, DECISION = 1, DECISION.THRES = 0.05, RANKSTAT = 1, MC = FALSE, NB.MC = 10000)
```
Arguments

**DATABASE**
Object returned by the function as PhViD.

**RR0**
Value of the tested risk. By default, RR0=1.

**MIN. n11**
Minimum number of notifications for a couple to be potentially considered as a signal. By default, MIN. n11 = 1.

**DECISION**
Decision rule for the signal generation based on
1 = FDR (Default value)
2 = Number of signals
3 = Ranking statistic. See RANKSTAT

**DECISION.THRES**
Threshold for DECISION. Ex 0.05 for FDR (DECISION=1).

**RANKSTAT**
Statistic used for ranking the couples:
1 = Posterior probability of the null hypothesis
2 = 2.5% quantile of the posterior distribution of IC.

**MC**
If MC=TRUE, the statistic of interest (see RANKSTAT) is calculated by Monte Carlo simulations which can be very long. If MC=FALSE, IC is approximated by a normal distribution (which can be very crude for small counts).

**NB.MC**
If MC=TRUE, NB.MC indicates the number of Monte Carlo simulations to be done.

Details

The BCPNN method is based on the calculation of the Information Component IC. If MC = FALSE, the bayesian model used is the beta-binomial proposed by Bate et al. (1998). The statistic of interest (see RANKSTAT) is calculated by the normal approximation made in Bate et al. (1998) with the use of the exact expectation and variance proposed by Gould (2003). If MC = TRUE, the model is based on the Dirichlet-multinomial model proposed more recently in Noren et al. (2006). In this case, the statistic of interest is calculated by Monte Carlo simulations.

Value

**ALLSIGNS**
Data.frame summarizing the results of all couples with at least MIN. n11 notifications ordered by RANKSTAT. It contains notably the labels, the cell counts, the expected counts \((n_1 \times n.1 / N, \text{ see } \text{as.PhViD})\), RANKSTAT, the ratios(count/expected count), the marginal counts and the estimations of FDR, FNR, Se et Sp. If RANKSTAT!=1, the last column is the posterior probability of the null hypothesis.

**SIGNS**
Same Data.frame as ALLSIGNS but restricted to the list of generated signals.

**NB.SIGNS**
Number of generated signals.

**INPUT.PARAM**
Parameters entered in the function.

Author(s)

Ismail Ahmed & Antoine Poncet
References


Gould AL, Practical Pharmacovigilance Analysis Strategies Pharmacoepidemiology and Drug Safety, 2003, 12, 559-574

Noren, GN, Bate A, Orre R, Edwards IR. Extending the methods used to screen the WHO drug safety database towards analysis of complex associations and improved accuracy for rare events Statistics in Medicine, 2006, 25, 3740-3757.

Examples

```r
## start
data(PhViD.data.frame)
PhViDdata <- as.PhViD(PhViD.data.frame)
# res <- BCPNN(PhViDdata)
## end
```

GPS

**Gamma Poisson Shrinkage**

Description

Gamma Poisson Shrinkage model proposed by DuMouchel (1999) extended to the multiple comparison framework.

Usage

```r
GPS(DATABASE, RR0 = 1, MIN.n11 = 1, DECISION = 1, DECISION.THRES = 0.05,
RANKSTAT = 1, TRONC = FALSE, TRONC.THRES = 1,
PRIOR.INIT = c(alpha1 = 0.2, beta1 = 0.06, alpha2 = 1.4,
beta2 = 1.8, w = 0.1), PRIOR.PARAM = NULL)
```

Arguments

- **DATABASE**
  - Object returned by the function as.PhViD.
- **RR0**
  - Value of the tested risk. By default, RR0=1.
- **MIN.n11**
  - Minimum number of notifications for a couple to be potentially considered as a signal. This option does not affect the calculation of the hyper parameters. By default, MIN.n11 = 1.
DECISION  Decision rule for the signal generation based on
1 = FDR (Default value)
2 = Number of signals
3 = Ranking statistic. See RANKSTAT

DECISION.THRES  Threshold for DECISION. Ex 0.05 for FDR (DECISION=1).

RANKSTAT  Statistic used for ranking the couples:
1 = Posterior probability of the null hypothesis
2 = 5% quantile of the posterior distribution of \( \lambda \)
3 = Posterior Expectation of \( \log_2(\lambda) \)

TRONC  If TRUE, only the data with at least TRONC.THRES notifications are considered in
the calculation of the hyper parameters and the likelihood is a product of mixture
of two negative binomial truncated by TRONC.THRES-1. By default, TRONC=F

TRONC.THRES  See TRONC

PRIOR.INIT  Vector of initialization of the prior parameters \((\alpha_1, \beta_1, \alpha_2, \beta_2, w)\). By default,
PRIOR.INIT = c(\(\alpha_1 = 0.2, \beta_1 = 0.06, \alpha_2 = 1.4, \beta_2 = 1.8, w = 0.1\)), ie the
prior parameters found in DuMouchel (1999).

PRIOR.PARAM  Chosen hyper parameters. By default, PRIOR.PARAM = NULL which means that
the hyperparameters are calculated by maximising the marginal likelihood.

Details

Each observed count \(n_{11}\) is assumed to be drawn from a Poisson distribution with parameters \(e_{11}\)
where \(e_{11}\) is the expected count under the hypothesis of independence between the adverse events
and the drugs \((n_1 \times n.1/N, \text{see } asPhViD)\). \(\lambda\) is a priori assumed to be distributed according to a
mixture of two gamma distributions: \(\lambda \sim w \Gamma(\alpha_1, \beta_1) + (1 - w) \Gamma(\alpha_2, \beta_2)\).

Value

ALLSIGNS  Data.frame summarizing the results of all couples with at least MIN.n11 notifications ordered by RANKSTAT. It contains notably the labels, the cell counts, the expected counts, RANKSTAT, the ratios(count/expected count), the marginal counts and the estimations of FDR, FNR, Se et Sp. If RANKSTAT!=1, the last column is the posterior probability of the null hypothesis.

SIGNS  Same Data.frame as ALLSIGNS but restricted to the list of generated signals.

NB.SIGNS  Number of generated signals.

INPUT.PARAM  Parameters entered in the function.

PARAM  A list that contains the prior hyper parameters (PRIOR.PARAM). Additionally if
PRIOR.PARAM=NULL, it also contains the prior hyper parameters initialization
(PRIOR.INIT) and the convergence code (see nlm()).

Author(s)

Ismail Ahmed & Antoine Poncet
References


Examples

```r
## start
#data(PhViD.data.frame)

#PhViD.data <- as.PhViD(PhViD.data.frame)
#res <- GPS(PhViD.data)

#List of signals generated by the decision rule proposed
#by Szarfman et al. (2002)
#res2 <- GPS(PhViD.data, DECISION = 3, DECISION.THRES = 2, RANKSTAT = 2)
## end
```

PhViD.search

Description

This function makes possible to extract some information from the output of the PhViD functions for a given couple adverse event-drug, for a drug or for an adverse event.

Usage

```r
PhViD.search(RESULT, DRUG = NULL, EVENT = NULL)
```

Arguments

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RESULT</td>
<td>RESULT must be the output of one the signal detection method functions (ROR, PRR, RFET, GPS or BCPNN)</td>
</tr>
<tr>
<td>DRUG</td>
<td>The label of the drug. By default, DRUG=FALSE.</td>
</tr>
<tr>
<td>EVENT</td>
<td>The label of the adverse event. By default, EVENT=FALSE.</td>
</tr>
</tbody>
</table>
**PhViD.data.frame**

**Value**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRUG</td>
<td>Recalls the label of the drug.</td>
</tr>
<tr>
<td>EVENT</td>
<td>Recalls the label of the event.</td>
</tr>
<tr>
<td>EXIST_DRUG</td>
<td>Indicates if the label of the drug exists in the database.</td>
</tr>
<tr>
<td>EXIST_EVENT</td>
<td>Indicates if the label of the adverse event exists in the database.</td>
</tr>
<tr>
<td>EXIST_COUPLE</td>
<td>Indicates if the couple is present in the database.</td>
</tr>
<tr>
<td>LIST</td>
<td>It is a dataframe that contains the labels, the counts, the expected counts,</td>
</tr>
<tr>
<td></td>
<td>the value of the statistic of interest, the rank and the estimated FDR for</td>
</tr>
<tr>
<td></td>
<td>each couple.</td>
</tr>
</tbody>
</table>

**Author(s)**

Antoine Poncet & Ismaïl Ahmed

---

**PhViD.data.frame**

*Simulated Pharmacovigilance data*

**Description**

This is a simulated data set aiming at mimicking the French database coded in ATC5 for the drugs and HLT for the adverse events. The simulation procedure is described in Ahmed et al.

**Usage**

`data(PhViD.data.frame)`

**Format**

A data frame with 102483 observations on the following 3 variables.

- **Drug lab** a factor indicating the label of the 634 drugs.
- **AE lab** a factor indicating the label of the 756 adverse events.
- **n11** a numeric vector indicating the number of spontaneous reports of the corresponding couple.

**Author(s)**

Ismail Ahmed & Antoine Poncet
**Description**

Proportional Reporting Ratio proposed by Evans et al. (2001) extended to the multiple comparison framework. Note that the computed variance is different from the one used in van Puijenbroek et al. (2002)

**Usage**

\[
\text{PRR} (\text{DATABASE, RR} = 1, \text{MIN.
n} = 1, \text{DECISION} = 1, \\
\text{DECISION.THRES} = 0.05, \text{RANKSTAT} = 1)
\]

**Arguments**

- DATABASE: Object returned by the function \text{as.PhViD}.
- RR: Value of the tested relative risk. By default, \text{RR}=1.
- MIN.n: Minimum number of notifications for a couple to be potentially considered as a signal. By default, \text{MIN.n} = 1.
- DECISION: Decision rule for the signal generation based on 1 = FDR (Default value)
  2 = Number of signals
  3 = Ranking statistic. See RANKSTAT
- DECISION.THRES: Threshold for DECISION. Ex 0.05 for FDR (DECISION=1).
- RANKSTAT: Statistic used for ranking the couples:
  1 = P-value
  2 = Lower bound of the 95% two sided confidence interval of \log(\text{PRR}).

**Details**

The FDR is estimated with the LBE procedure proposed by Dalmasso et al. (2005). Note that the FDR can only be estimated if the statistic of interest is the P-value.

**Value**

- ALLSIGNALS: Data.frame summarizing the results of all couples with at least MIN.n notifications ordered by RANKSTAT. It contains notably the labels, the cell counts, the expected counts \((n1 \times n1/N, \text{see as.PhViD})\), RANKSTAT, the observed relative risks (PRR), the marginal counts and the estimations of FDR (when RANKSTAT=1.)
- SIGNALS: Same Data.frame as ALLSIGNALS but restricted to the list of generated signals.
- NB.SIGNALS: Number of generated signals.
- INPUT.PARAM: Parameters entered in the function.
Author(s)

Ismail Ahmed & Antoine Poncet

References


Evans SJ, Waller PC, Davis S, Use of Proportional Reporting Ratios (PRRs) for Signal Generation from Spontaneous Adverse Drug Reaction Reports Pharmacoepidemiology and Drug Safety, 2001, 10, 483-486.


Examples

```r
## start
data(asPhViD(PhViDdata.frame))
PhViDdata <- asPhViD(PhViDdata.frame)
# res <- PRR(PhViDdata)
## end
```

---

**RFET** Reporting Fisher's Exact Test

**Description**

This function proposes the Fisher’s Exact Test as an alternative to the PRR and ROR methods. The statistic of interest is the P-value or the mid-P-value resulting from the test (Ahmed et al., Biometrics).

**Usage**

```r
RFET(DATABASE, OR0 = 1, MIN.n11 = 1, DECISION = 1, DECISION.THRES = 0.05, MID.PVAL = FALSE)
```

**Arguments**

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DATABASE</td>
<td>Object returned by the function as PhViD.</td>
</tr>
<tr>
<td>OR0</td>
<td>Value of the tested odds ratio. By default, OR0=1.</td>
</tr>
<tr>
<td>MIN.n11</td>
<td>Minimum number of notifications for a couple to be potentially considered as a signal. By default, MIN.n11 = 1.</td>
</tr>
</tbody>
</table>
**Decision**

Decision rule for the signal generation based on
1 = FDR (Default value)
2 = Number of signals
3 = P-values or mid-P-values. See **MID.PVAL**

**Decision.THRES**

Threshold for **DECISION**. Ex 0.05 for FDR (**DECISION**=1).

**MID.PVAL**

if **MID.PVAL**=TRUE, the statistic of interest becomes the mid-P-values instead of the P-values resulting from the Fisher’s exact test. By default **MID.PVAL**=FALSE.

**Details**

The FDR is estimated with the LBE procedure proposed by Dalmasso et al. (2005).

**Value**

**ALLSIGNALS**

Data.frame summarizing the results of all couples with at least MIN.n11 notifications ordered by **RANKSTAT**. It contains notably the labels, the cell counts, the expected count \((n1 \times n.1/N)\), see **as.PhiID**), **RANKSTAT**, the observed odds ratio (ROR), the marginal counts and the estimation of FDR.

**SIGNALS**

Same Data.frame as **ALLSIGNALS** but restricted to the list of generated signals.

**NB.SIGNALS**

Number of generated signals.

**INPUT.PARAM**

Parameters entered in the function.

**Author(s)**

Ismaïl Ahmed & Antoine Poncet

**References**


**Examples**

```r
## start
data(PhViIDdata.frame)
#PhViIDdata <- as.PhiID(PhViIDdata.frame)
#res <- RFET(PhViIDdata)
## end```
Reporting Odds Ratio

Description

Reporting Odds Ratio proposed by van Puijenbroak et al. (2002) extended to the multiple comparison framework.

Usage

\[
\text{ROR(DATABASE, OR0 = 1, MIN.n11 = 1, DECISION = 1,}
\text{ DECISION.THRES = 0.05, RANKSTAT = 1)}
\]

Arguments

- **DATABASE**: Object returned by the function as .PhViD.
- **OR0**: Value of the tested odds ratio. By default, OR0=1.
- **MIN.n11**: Minimum number of notifications for a couple to be potentially considered as a signal. By default, MIN.n11 = 1.
- **DECISION**: Decision rule for the signal generation based on
  - 1 = FDR (Default value)
  - 2 = Number of signals
  - 3 = Ranking statistic. See RANKSTAT
- **DECISION.THRES**: Threshold for DECISION. Ex 0.05 for FDR (DECISION=1).
- **RANKSTAT**: Statistic used for ranking the couples:
  - 1 = P-value
  - 2 = Lower bound of the 95% two sided confidence interval of log(ROR).

Details

The FDR is estimated with the LBE procedure proposed by Dalmasso et al. (2005). Note that the FDR can only be estimated if the statistic of interest is the P-value.

Value

- **ALLSIGNALS**: Data.frame summarizing the results of all couples with at least MIN.n11 notifications ordered by RANKSTAT. It contains notably the labels, the cell counts, the expected counts \(n1 \times n1/N\), see as .PhViD, RANKSTAT, the observed odds ratios (ROR), the marginal counts and the estimations of FDR (when RANKSTAT=1.)
- **SIGNALS**: Same Data.frame as ALLSIGNALS but restricted to the list of generated signals.
- **NB.SIGNALS**: Number of generated signals.
- **INPUT.PARAM**: Parameters entered in the function.
Author(s)
Ismaïl Ahmed & Antoine Poncet

References

Examples
```r
## start
data(PhViDdata.frame)
PhViDdata <- as.PhViD(PhViDdata.frame)
res <- ROR(PhViDdata, MIN.n11 = 3)

# Decision rule proposed by van Puijenbroek et al. (2002)
# res2 <- ROR(PhViDdata, MIN.n11 = 1, DECISION=3, DECISION.THRES=0, RANKSTAT=2)
## end
```
Index

+Topic **datasets**
  PhViDdata.frame, 8
+Topic **htest**
  BCPNN, 3
  GPS, 5
  PhViD-package, 2
  PRR, 9
  RFET, 10
  ROR, 12
+Topic **manip**
  as.PhViD, 3
  PhViD.search, 7
  as.PhViD, 3, 4, 6, 9, 11, 12

BCPNN, 3
GPS, 5
PhViD (PhViD-package), 2
PhViD-package, 2
PhViD.search, 7
PhViDdata.frame, 8
PRR, 9
RFET, 10
ROR, 12