

Package ‘BioStatR’

November 28, 2022

Type Package

Title Initiation à La Statistique Avec R

Version 4.0.0

Date 2022-11-28

Depends R (>= 2.10)

Imports ggplot2

Author Frederic Bertrand [cre, aut] (<<https://orcid.org/0000-0002-0837-8281>>),
Myriam Maumy-Bertrand [aut] (<<https://orcid.org/0000-0002-4615-1512>>)

Maintainer Frederic Bertrand <frederic.bertrand@utt.fr>

Description Datasets and functions for the book ``Initiation à la Statistique avec R'', F. Bertrand and M. Maumy-Bertrand (2022, ISBN:978-2100782826 Dunod, 4ème edition).

LazyLoad yes

LazyData yes

License GPL-3

Encoding UTF-8

Classification/MSC

URL <https://fbertran.github.io/BioStatR/>,
<https://github.com/fbertran/BioStatR/>

BugReports <https://github.com/fbertran/BioStatR/issues/>

RoxygenNote 7.2.1

NeedsCompilation no

Repository CRAN

Date/Publication 2022-11-28 12:50:02 UTC

R topics documented:

binom.ci	2
BioStatR	3
cvar	4
eta2	5
Europe	6
Extrait_Taille	6
gg_qqplot	7
Mesures	9
Mesures5	9
panel.hist	10
plotcdf2	11
poi.ci	13
Quetelet	14
Index	15

binom.ci	<i>Intervalles de confiance pour une proportion</i>
----------	---

Description

Cette fonction permet de calculer plusieurs types d'intervalles de confiance pour une proportion.

Usage

```
binom.ci(x, n, conf.level = 0.95, method = c("Wilson", "exact", "Wald", "all"))
```

Arguments

x	Nombre de succès
n	Nombre d'essais
conf.level	Niveau de confiance recherché pour l'intervalle
method	Type d'intervalle de confiance à calculer : intervalle de "Wilson", intervalle "exact" de Clopper-Pearson, intervalle asymptotique de "Wald" ou tous les trois "all"

Value

matrix	Limites des intervalles de confiance demandés.
--------	--

Author(s)

Frédéric Bertrand
<frederic.bertrand@utt.fr>
<http://www-irma.u-strasbg.fr/~fbertran/>
Maumy-Bertrand
<myriam.maumy@utt.fr>
<http://www-irma.u-strasbg.fr/~mmaumy/>

References

F. Bertrand, M. Maumy-Bertrand, Initiation à la Statistique avec R, Dunod, 3ème edition, 2018.

See Also

[binom.test](#), [binom.ci](#), [poi.ci](#)

Examples

```
binom.ci(5,10,method="all")
```

BioStatR

BioStatR

Description

Motivation: Package compagnon du livre Initiation à la statistique avec R. Il contient les codes des chapitres du livre ainsi que les solutions des exercices mais aussi d'autres compléments à découvrir.

References

F. Bertrand, M. Maumy-Bertrand, Initiation à la Statistique avec R, 3ème edition, ISBN:9782100782826, Dunod, Paris, 2018.

Initiation à la Statistique avec R, Frédéric Bertrand, Myriam Maumy-Bertrand, 2018, , <https://www.dunod.com/sciences-techniques/initiation-statistique-avec-r-cours-exemples-exercices-et-problem>
<https://github.com/fbertran/BioStatR/> et <https://fbertran.github.io/BioStatR/>

Examples

```
set.seed(314)
```

cvar	<i>Coefficient de variation</i>
------	---------------------------------

Description

Calcule coefficient de variation d'une série statistique

Usage

```
cvar(x)
```

Arguments

x Un vecteur numérique

Details

Le coefficient de variation est égal à l'écart-type corrigé divisé par la moyenne. Il est exprimé en pourcents.

Value

num Valeur du coefficient de variation exprimé en pourcents

Author(s)

Frédéric Bertrand
<frederic.bertrand@utt.fr>
<http://www-irma.u-strasbg.fr/~fbertran/>
Maumy-Bertrand
<myriam.maumy@utt.fr>
<http://www-irma.u-strasbg.fr/~mmaumy/>

References

F. Bertrand, M. Maumy-Bertrand, Initiation à la Statistique avec R, Dunod, 3ème edition, 2018.

See Also

[mean](#), [sd](#)

Examples

```
data(Europe)  
cvar(Europe[,2])
```

`eta2`*Calcul du rapport de corrélation eta carré*

Description

Cette fonction calcule le rapport de corrélation η^2 qui est une mesure d'association importante entre une variable quantitative et une variable qualitative.

Usage`eta2(x, y)`**Arguments**

<code>x</code>	Un vecteur associé à la variable quantitative
<code>y</code>	Un facteur associé à la variable qualitative

Value

<code>num</code>	La valeur du rapport de corrélation empirique
------------------	---

Author(s)

Frédéric Bertrand
<frederic.bertrand@utt.fr>
<http://www-irma.u-strasbg.fr/~fbertran/>
Maumy-Bertrand
<myriam.maumy@utt.fr>
<http://www-irma.u-strasbg.fr/~mmaumy/>

References

F. Bertrand, M. Maumy-Bertrand, Initiation à la Statistique avec R, Dunod, 3ème edition, 2018.

Examples`eta2(Mesures5$taille, Mesures5$espece)`

Europe

Durées de travail en Europe

Description

This dataset provide mean weekly cumulated work durations for several European countries.

Format

A data frame with 25 observations on the following 2 variables.

Pays a factor with the some of the European countries as levels

Duree weekly cumulative work duration

Details

The duration is given in hours

References

F. Bertrand, M. Maumy-Bertrand, *Initiation à la Statistique avec R*, Dunod, 3ème edition, 2018.

Examples

```
data(Europe)
```

Extrait_Taille

Mesures de fruits d'arbustes

Description

This dataset measurements of several features of the fruits of small trees such as their mass (in g) or their length (in cm).

Format

A data frame with 252 observations on the following 5 variables.

masse a numeric vector

taille a numeric vector

espece a factor with levels bignone, glycine blanche, glycine violette and lauriers roses

Details

This dataset was made during the summer 2009 in the south of France. It provides measurements of several features of the fruits of small trees such as their mass or their length.

References

F. Bertrand, M. Maumy-Bertrand, Initiation à la Statistique avec R, Dunod, 3ème edition, 2018.

Examples

```
data(Extrait_Taille)
```

gg_qqplot

Graphique des quantiles (qqplot) et droite interquartile

Description

Dessine le graphique des quantiles ou `qqplot` et la droite interquartile (passant par le premier et le troisième quartile à la manière de la fonction `qqline`) avec la bibliothèque graphique `ggplot2`.

Usage

```
gg_qqplot(  
  df,  
  var,  
  qdist = qnorm,  
  params = list(),  
  qq.line = TRUE,  
  color = "red",  
  alpha = 0.5  
)
```

Arguments

<code>df</code>	Un jeu de données (dataframe)
<code>var</code>	Le nom d'une variable de <code>df</code>
<code>qdist</code>	La fonction quantile d'une (famille de) distribution. Par défaut celle de la famille des lois normales.
<code>params</code>	Une liste de paramètres pour spécifier la loi à utiliser. Par défaut la loi normale centrée et réduite. Les paramètres peuvent être estimés avec la fonction <code>fitdistr</code> de la bibliothèque MASS.
<code>qq.line</code>	Une valeur logique. Affiche ou masque la droite interquartile.
<code>color</code>	Le nom d'une couleur. Spécifie la couleur à utiliser pour la droite interquartile.
<code>alpha</code>	Indice de transparence. Spécifie la transparence à utiliser pour représenter les valeurs de l'échantillon.

Value

ggplot Un graphique utilisant la bibliothèque ggplot2. Affiche les valeurs des quartiles théoriques par lesquels passe la droite ainsi que son ordonnée à l'origine et sa pente si le tracé de celle-ci est demandé.

Author(s)

Frédéric Bertrand
 <frederic.bertrand@utt.fr>
<http://www-irma.u-strasbg.fr/~fbertran/>
 Maumy-Bertrand
 <myriam.maumy@utt.fr>
<http://www-irma.u-strasbg.fr/~mmaumy/>

References

F. Bertrand, M. Maumy-Bertrand, Initiation à la Statistique avec R, Dunod, 3e, 2018.

See Also

[qqplot](#), [qqline](#)

Examples

```
glycine.blanche<-subset(Mesures,subset=(Mesures$espece=="glycine blanche"))
gg_qqplot(glycine.blanche,"taille")

#bonus ajustement avec une autre loi (ici Student (car dist = qt) dont on estime les ddl)
lauriers.roses<-subset(Mesures,subset=(Mesures$espece=="laurier rose"))
shapiro.test(lauriers.roses$taille)
#pas issu d'une loi normale au risque alpha=5%
gg_qqplot(lauriers.roses,"taille")
gg_qqplot(lauriers.roses,"taille",qq.line=FALSE)
#essayons un qqplot avec une loi de Student
## Not run:
require(MASS)
params <- as.list(fitdistr(lauriers.roses$taille, "t")$estimate)
#avec la droite
gg_qqplot(lauriers.roses,"taille",qt,params)
#essayons un qqplot avec une loi gamma
params <- as.list(fitdistr(lauriers.roses$taille,"gamma")$estimate)
#avec la droite
gg_qqplot(lauriers.roses,"taille",qgamma,params)
#essayons un qqplot avec une loi du chi-deux
params <- list(df=fitdistr(lauriers.roses$taille,"chi-squared",start=list(df=5),
method="Brent",lower=1,upper=40)$estimate)
#avec la droite
gg_qqplot(lauriers.roses,"taille",qchisq,params)

## End(Not run)
```

Mesures

Mesures de fruits d'arbustes

Description

This dataset measurements of several features of the fruits of small trees such as their mass or their length.

Format

A data frame with 252 observations on the following 3 variables.

masse a numeric vector

taille a numeric vector

espece a factor with levels bignone, glycine blanche, glycine violette and lauriers roses

Details

This dataset was made during the summer 2009 in the south of France. It provides measurements of several features of the fruits of small trees such as their mass or their length.

References

F. Bertrand, M. Maumy-Bertrand, Initiation à la Statistique avec R, Dunod, 3ème edition, 2018.

Examples

```
data(Mesures)
```

Mesures5

Mesures de fruits d'arbustes

Description

This dataset measurements of several features of the fruits of small trees such as their mass or their length.

Format

A data frame with 252 observations on the following 5 variables.

masse a numeric vector

taille a numeric vector

graines a numeric vector

masse_sec a numeric vector

espece a factor with levels bignone, glycine blanche, glycine violette and lauriers roses

Details

This dataset was made during the summer 2009 in the south of France. It provides measurements of several features of the fruits of small trees such as their mass or their length.

References

F. Bertrand, M. Maumy-Bertrand, Initiation à la Statistique avec R, Dunod, 3ème edition, 2018.

Examples

```
data(Mesures5)
```

panel.hist

Histogrammes

Description

Sert à représenter des histogrammes dans les graphiques matriciels

Usage

```
panel.hist(x, ...)
```

Arguments

x Un vecteur numérique

... Des arguments à transmettre à la fonction qui créé les histogrammes

Details

Cette fonction s'utilise avec la fonctions graphique pairs.

Author(s)

Frédéric Bertrand
<frederic.bertrand@utt.fr>
<http://www-irma.u-strasbg.fr/~fbertran/>
Maumy-Bertrand
<myriam.maumy@utt.fr>
<http://www-irma.u-strasbg.fr/~mmaumy/>

References

F. Bertrand, M. Maumy-Bertrand, Initiation à la Statistique avec R, Dunod, 3ème edition, 2018.

See Also

[pairs](#), [hist](#)

Examples

```
data(Mesures5)
pairs(Mesures5,diag.panel="panel.hist")
```

plotcdf2

Représentation bivariée des variables discrètes ou des variables continues groupées en classes.

Description

Cette fonction construit un stéréogramme permettant de juger de l'association entre deux variables discrètes ou groupées en classes.

Usage

```
plotcdf2(  
  x,  
  y,  
  f,  
  xaxe,  
  yaxe,  
  col = NULL,  
  border = FALSE,  
  Nxy = 200,  
  theme = "0"  
)
```

Arguments

x	Valeurs observées ou modalités de la première variable discrète
y	Valeurs observées ou modalités de la seconde variable discrète
f	Si $f=0$ (donc $\text{length}(f)=0$), x et y sont deux séries statistiques. Si $\text{length}(f)>1$, f est un tableau de fréquences et x et y les noms des lignes et des colonnes de f.
xaxe	Nom de l'axe des abscisses
yaxe	Nom de l'axe des ordonnées
col	Couleur du stéréogramme
border	Le maillage du graphique doit-il être affiché ?
Nxy	Pas du maillage pour chaque axe
theme	Le thème détermine la palette de couleurs utilisées. Il y a quatre choix possibles en couleurs "0", "1", "2", "3" et un en nuances de gris "bw"

Value

Un stéréogramme des deux séries statistiques groupées ou des deux variables discrètes étudiées.

Author(s)

Frédéric Bertrand
 <frederic.bertrand@utt.fr>
<http://www-irma.u-strasbg.fr/~fbertran/>
 Maumy-Bertrand
 <myriam.maumy@utt.fr>
<http://www-irma.u-strasbg.fr/~mmaumy/>

References

F. Bertrand, M. Maumy-Bertrand, Initiation à la Statistique avec R, Dunod, 3ème édition, 2018.

Examples

```
xx=c(1.83,1.72,1.65,1.70,2.05,1.92,1.85,1.70,1.75,1.9)
yy=c(75,70,70,60,90,92,75,68,71,87)
plotcdf2(xx,yy,f=0,"taille en m","poids en kg")

xx=seq(2,12)
yy=seq(1,6)
p=c(1/36,0,0,0,0,0,
2/36,0,0,0,0,0,
2/36,1/36,0,0,0,0,
2/36,2/36,0,0,0,0,
2/36,2/36,1/36,0,0,0,
2/36,2/36,2/36,0,0,0,
0,2/36,2/36,1/36,0,0,
0,0,2/36,2/36,0,0,
0,0,0,2/36,1/36,0,
```

```
0,0,0,0,2/36,0,  
0,0,0,0,0,1/36)  
p=matrix(p,byrow=TRUE,ncol=6)  
plotcdf2(xx,yy,p,"somme des dés","valeur du plus petit")
```

poi.ci

Intervalle de confiance pour le paramètre d'une loi de Poisson

Description

Créé un intervalle de confiance pour le paramètre d'une loi de Poisson.

Usage

```
poi.ci(x, conf.level = 0.95)
```

Arguments

x	Un vecteur de données
conf.level	Niveau de confiance de l'intervalle

Value

matrix	Limites des intervalles de confiance demandés.
--------	--

Author(s)

Frédéric Bertrand
<frederic.bertrand@utt.fr>
<http://www-irma.u-strasbg.fr/~fbertran/>
Maumy-Bertrand
<myriam.maumy@utt.fr>
<http://www-irma.u-strasbg.fr/~mmaumy/>

References

F. Bertrand, M. Maumy-Bertrand, Initiation à la Statistique avec R, Dunod, 3ème edition, 2018.

See Also

[binom.test](#), [binom.ci](#), [poi.ci](#)

Examples

```
poi.ci(rpois(20,10))
```

Quetelet

Indices de Quetelet

Description

Ce jeu de données contient des mesures de masse et de taille pour permettre le calcul de l'indice de masse corporelle (aussi dit de Quetelet).

Format

Un data frame avec 66 observations de 3 variables.

sexe un facteur donnant le sexe de l'individu

poids le poids de l'individu

taille la hauteur de l'individu

Details

Le poids est exprimée en kg et la hauteur en cm

References

F. Bertrand, M. Maumy-Bertrand, Initiation à la Statistique avec R, Dunod, 3ème edition, 2018.

Examples

```
data(Quetelet)
```

Index

* datasets

Europe, 6
Extrait_Taille, 6
Mesures, 9
Mesures5, 9
Quetelet, 14

* univar

binom.ci, 2
cvar, 4
eta2, 5
gg_qqplot, 7
panel.hist, 10
plotcdf2, 11
poi.ci, 13

binom.ci, 2, 3, 13

binom.test, 3, 13

BioStatR, 3

cvar, 4

eta2, 5

Europe, 6

Extrait_Taille, 6

fitdistr, 7

gg_qqplot, 7

hist, 11

mean, 4

Mesures, 9

Mesures5, 9

pairs, 11

panel.hist, 10

plotcdf2, 11

poi.ci, 3, 13, 13

qqline, 7, 8

qqplot, 7, 8

Quetelet, 14

sd, 4