

# Package ‘ModStatR’

October 12, 2022

**Type** Package

**Title** Statistical Modelling in Action with R

**Version** 1.3.1

**Date** 2021-03-13

**Depends** R (>= 3.5.0)

**Imports** stats, boot, jmuOutlier, ellipse, hypergeo, gsl

**Suggests** BioStatR, ggplot2

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

**Maintainer** Frederic Bertrand <[frederic.bertrand@math.unistra.fr](mailto:frederic.bertrand@math.unistra.fr)>

**Description** Datasets and functions for the book “Modélisation statistique par la pratique avec R”, F. Bertrand, E. Claeys and M. Maumy-Bertrand (2019, ISBN:9782100793525, Dunod, Paris). The first chapter of the book is dedicated to an introduction to the R statistical software. The second chapter deals with correlation analysis: Pearson, Spearman and Kendall simple, multiple and partial correlation coefficients. New wrapper functions for permutation tests or bootstrap of matrices of correlation are provided with the package. The third chapter is dedicated to data exploration with factorial analyses (PCA, CA, MCA, MDA) and clustering. The fourth chapter is dedicated to regression analysis: fitting and model diagnostics are detailed. The exercises focus on covariance analysis, logistic regression, Poisson regression, two-way analysis of variance for fixed or random factors. Various example datasets are shipped with the package: for instance on pokemon, world of warcraft, house tasks or food nutrition analyses.

**LazyLoad** yes

**LazyData** yes

**License** GPL-3

**Encoding** UTF-8

**Classification/MS**C

**URL** <https://fbertran.github.io/homepage/>,  
<https://github.com/fbertran/ModStatR/>

**BugReports** <https://github.com/fbertran/ModStatR/issues/>

**RoxygenNote** 7.1.1

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2021-03-14 15:10:08 UTC

## R topics documented:

boot.mcor.ic . . . . .	3
CancerSein . . . . .	3
chal . . . . .	4
corrdist . . . . .	4
corrdistapprox . . . . .	5
corrdistapprox2 . . . . .	6
data_event . . . . .	6
d_hotels . . . . .	8
d_hotels_n . . . . .	9
d_macdo . . . . .	9
d_pres2002 . . . . .	11
d_pres2007 . . . . .	11
d_TM . . . . .	12
d_vac . . . . .	13
d_wow . . . . .	13
ecole2 . . . . .	14
ecole3 . . . . .	15
Gauss2F1 . . . . .	15
Gauss2F1gsl . . . . .	16
ModStatR . . . . .	16
my.confidence.region . . . . .	17
parasites . . . . .	17
perm.cor.mtest . . . . .	18
poke . . . . .	19
polypes . . . . .	20
ref.cor.mtest . . . . .	20
ref.cor.test . . . . .	21
resistance . . . . .	22
rho . . . . .	23
rho.mult . . . . .	23
SidaChat . . . . .	24
vitamines . . . . .	25

**Index**

**26**

---

boot.mcor.ic	<i>Intervalles de confiance bootstrap pour une matrice de corrélation de Bravais-Pearson</i>
--------------	--

---

**Description**

Intervalles de confiance bootstrap pour une matrice de corrélation de Bravais-Pearson

**Usage**

```
boot.mcor.ic(mat, boot.mcor.res, conplevel = 0.95)
```

**Arguments**

mat	Matrice des données
boot.mcor.res	Résultat du bootstrap de la matrice de corrélation de Bravais-Pearson
conplevel	Niveau de confiance pour les intervalles

**Value**

Liste de quatre éléments : matrice des limites inférieures des intervalles de confiance bootstrap percentile, matrice des limites supérieures des intervalles de confiance bootstrap percentile, matrice des limites inférieures des intervalles de confiance bootstrap BCa, matrice des limites supérieures des intervalles de confiance bootstrap BCa

**Examples**

```
data(Mesures5,package="BioStatR")
Mes5_red_lr = subset(Mesures5[, -5], subset=Mesures5$espece=="laurier rose")
library(boot)
boot.mcor <- boot(Mes5_red_lr[,c("masse", "taille", "masse_sec")], rho.mult, R=1000)
boot.mcor
boot.mcor.ic.res <- boot.mcor.ic(Mes5_red_lr[,c("masse", "taille", "masse_sec")], boot.mcor)
boot.mcor.ic.res
```

---

CancerSein	<i>Cancer du sein</i>
------------	-----------------------

---

**Description**

Jeu de données cancer du sein.

**Usage**

```
CancerSein
```

**Format**

Un objet `data.frame` avec trois variables et 62 observations :

**Traitement** Factor w/ 3 levels, Type de traitement administré

**Age** Numerical vector, Âge de la patiente au début du traitement

**Survie** Numerical vector, Durée de survie de la patiente

**Source**

*Modèle linéaire : Comparaison de groupes et régression* de B. Prum aux Éditions de l'INSERM, 1996.

---

chal	<i>Challenger</i>
------	-------------------

---

**Description**

Jeu de données navette spatiale.

**Usage**

chal

**Format**

Un objet `data.frame` avec deux variables et 24 observations :

**Temperature** Integer vector, Température au moment du décollage

**Défaillance** Integer vector, Défaillance de l'un des *o-ring*

---

corrdist	<i>Densité de rho chapeau, estimateur du coefficient de corrélation de Bravais-Pearson</i>
----------	--

---

**Description**

Densité de rho chapeau, estimateur du coefficient de corrélation de Bravais-Pearson

**Usage**

`corrdist(rho, rho_0, n)`

**Arguments**

rho	Valeur en laquelle la densité est évaluée
rho_0	Valeur de référence pour le coefficient de corrélation de Bravais-Pearson
n	Effectif de l'échantillon

**Value**

Valeur numérique, densité au point rho de l'estimateur, construit à partir d'un échantillon de taille n, du coefficient de corrélation de Bravais-Pearson, de valeur théorique égale à rho\_0, sous hypothèse de normalité multivariée

**Examples**

```
corrdist(.7, .8, 30)
```

---

corrdistapprox	<i>Approximation de la densité de rho chapeau, estimateur du coefficient de corrélation de Bravais-Pearson</i>
----------------	--

---

**Description**

Approximation de la densité de rho chapeau, estimateur du coefficient de corrélation de Bravais-Pearson

**Usage**

```
corrdistapprox(rho, rho_0, n)
```

**Arguments**

rho	Valeur en laquelle la densité est évaluée
rho_0	Valeur de référence pour le coefficient de corrélation de Bravais-Pearson
n	Effectif de l'échantillon

**Value**

Valeur numérique, approximation de la densité au point rho de l'estimateur, construit à partir d'un échantillon de taille n, du coefficient de corrélation de Bravais-Pearson, de valeur théorique égale à rho\_0, sous hypothèse de normalité multivariée

**Examples**

```
corrdistapprox(.7, .8, 30)
```

---

corrdistapprox2	<i>Meilleure approximation de la densité de rho chapeau, estimateur du coefficient de corrélation de Bravais-Pearson</i>
-----------------	--

---

**Description**

Meilleure approximation de la densité de rho chapeau, estimateur du coefficient de corrélation de Bravais-Pearson

**Usage**

```
corrdistapprox2(rho, rho_0, n)
```

**Arguments**

rho	Valeur en laquelle la densité est évaluée
rho_0	Valeur de référence pour le coefficient de corrélation de Bravais-Pearson
n	Effectif de l'échantillon

**Value**

Valeur numérique, approximation de la densité au point rho de l'estimateur, construit à partir d'un échantillon de taille n, du coefficient de corrélation de Bravais-Pearson, de valeur théorique égale à rho\_0, sous hypothèse de normalité multivariée

**Examples**

```
corrdistapprox2(.7, .8, 30)
```

---

data_event	<i>Matches de football joués en France de 2011/2012 à 2016/2017</i>
------------	---

---

**Description**

Ce fichier de données fournit une vue granulaire de 208 446 matchs de football joués en France depuis la saison 2011/2012 à la saison 2016/2017.

**Usage**

```
data_event
```

**Format**

Un objet data.frame avec 208446 lignes et 40 variables :

X Integer vector, identifiant unique de ligne

id\_odsp Factor w/ 2076 levels, identifiant unique de match

id\_event Factor w/ 208446, unique identifier of event (id\_odsp + sort\_order)

sort\_order Integer vector, chronological sequence of events in a game

time Integer vector, minute of the game

text Factor w/ 79629 levels, text commentary

event\_type Integer vector, primary event. 11 unique events (1-Attempt(shot), 2-Corner, 3-Foul, 4-Yellow Card, 5-Second yellow card, 6-(Straight) red card, 7-Substitution, 8-Free kick won, 9-Offside, 10-Hand Ball, 11-Penalty conceded)

event\_type2 Integer vector, secondary event. 4 unique events (12 - Key Pass, 13 - Failed through ball, 14-Sending off, 15-Own goal)

side Integer vector, 1-Home, 2-Away

event\_team Factor w/ 30 levels, Équipe de football qui est à l'origine de l'événement. In case of Own goals, event team is the team that benefited from the own goal

opponent Factor w/ 30 levels, team that the event happened against

player Factor w/ 1609, name of the player involved in main event (converted to lowercase and special chars were removed)

player2 Factor w/ 1498, name of player involved in secondary event

player\_in Factor w/ 1277, player that came in (only applies to substitutions)

player\_out Factor w/ 1204, player substituted (only applies to substitutions)

shot\_place Integer vector, placement of the shot (13 possible placement locations, available in the dictionary, only applies to shots)

shot\_outcome Integer vector, 4 possible outcomes (1-On target, 2-Off target, 3-Blocked, 4-Hit the post)

is\_goal Integer vector, binary variable if the shot resulted in a goal (own goals included)

location Integer vector, location on the pitch where the event happened (19 possible locations, available in the dictionary)

bodypart Integer vector, (1- right foot, 2-left foot, 3-head)

assist\_method Integer vector, in case of an assisted shot, 5 possible assist methods (details in the dictionary)

situation Integer vector, 4 types: 1-Open Play, 2-Set piece (excluding Direct Free kicks), 3-Corner, 4-Free kick

fast\_break Integer vector, binary

link\_odsp Factor w/ 2076 levels lien vers la page oddsportal

adv\_stats Logical vector, boolean if the game has detailed event data

date Factor w/ 592 levels, Date of game

league Factor w/ 1 level, Club League

season Integer vector, Year Played  
 country Factor w/ 1 level, Host Nation of League  
 ht Factor w/ 30 levels, home team  
 at Factor w/ 30 levels, away team  
 fthg Integer vector, full time home goals  
 ftag Integer vector, full time away goals  
 odd\_h Numerical vector, highest home win market odds  
 odd\_d Numerical vector, highest draw market odds  
 odd\_a Numerical vector, highest away market odds  
 odd\_over Numerical vector, highest over 2.5 market odds  
 odd\_under Numerical vector, highest under 2.5 market odds  
 odd\_bts Numerical vector, highest both teams to score market odds  
 odd\_bts\_n Numerical vector, highest both teams NOT to score market odds

### Details

Ces données sont une version « nettoyée » d'un fichier original, events\_France.csv, qu'il est possible de télécharger depuis la plate-forme Kaggle : <https://www.kaggle.com/secareanualin/football-events>. Certains matchs contiennent cependant des données manquantes (environ 10

### Source

Kaggle.

---

d_hotels	<i>Hotels</i>
----------	---------------

---

### Description

Jeu de données hotels (nom des hotels dans la variable NOM).

### Usage

d\_hotels

### Format

Un objet data.frame avec neuf variables et 39 observations :

**NOM** Factor w/ 39 levels, Nom de l'hôtel  
**PAYS** Factor w/ 5 levels, Pays où est situé l'hôtel  
**ETOILE** Integer vector, Catégorie de l'hôtel  
**CONFORT** Integer vector, Note de confort



**CHAMBRE** Integer vector, Nombre de chambres  
**CUISINE** Integer vector, Note de la cuisine  
**SPORT** Integer vector, Note des instructures sprotives  
**PLAGE** Integer vector, Note de la plage  
**PRIX** Integer vector, Prix moyen pour une nuit

---

d\_hotels\_n                      *Hotels*

---

### Description

Jeu de données hotels (nom des hotels en rownames).

### Usage

d\_hotels\_n

### Format

Un objet data . frame avec huit variables et 39 observations :

**PAYS** Factor w/ 5 levels, Pays où est situé l'hôtel  
**ETOILE** Integer vector, Catégorie de l'hôtel  
**CONFORT** Integer vector, Note de confort  
**CHAMBRE** Integer vector, Nombre de chambres  
**CUISINE** Integer vector, Note de la cuisine  
**SPORT** Integer vector, Note des instructures sprotives  
**PLAGE** Integer vector, Note de la plage  
**PRIX** Integer vector, Prix moyen pour une nuit

---

d\_macdo                              *MacDonald*

---

### Description

Jeu de données sur des aliments vendus au McDonald.

### Usage

d\_macdo

**Format**

Un objet `data.frame` avec 24 variables et 260 observations :

**Category** Factor w/ 9 levels, Type d'aliment

**Item** Factor w/ 260 levels, Nom de l'aliment

**Serving.Size** Factor w/ 107 levels, Taille

**Calories** Integer vector, Calories

**Calories.from.Fat** Integer vector, Calories dues à la graisse

**Total.Fat** Numerical vector, Graisse totale

**Total.Fat....Daily.Value.** Integer vector, % de la quantité de graisse totale attendue par jour

**Saturated.Fat** Numerical vector, Graisse saturée

**Saturated.Fat....Daily.Value.** Integer vector, % de la quantité de graisse saturée attendue par jour

**Trans.Fat** Numerical vector, Trans.Fat

**Cholesterol** Integer vector, Cholesterol

**Cholesterol....Daily.Value.** Integer vector, % de la quantité de cholestérol attendue par jour

**Sodium** Integer vector, Sel

**Sodium....Daily.Value.** Integer vector, % de la quantité de sel attendue par jour

**Carbohydrates** Integer vector, Carbohydrates

**Carbohydrates....Daily.Value.** Integer vector, % de la quantité de carbohydrates attendue par jour

**Dietary.Fiber** Integer vector, Fibres

**Dietary.Fiber....Daily.Value.** Integer vector, % de la quantité de fibres attendue par jour

**Sugars** Integer vector, Sucres

**Protein** Integer vector, Protéines

**Vitamin.A....Daily.Value.** Integer vector, % de la quantité de vitamine A attendue par jour

**Vitamin.C....Daily.Value.** Integer vector, % de la quantité de vitamine C attendue par jour

**Calcium....Daily.Value.** Integer vector, % de la quantité de calcium attendue par jour

**Iron....Daily.Value.** Integer vector, % de la quantité de fer attendue par jour

**Details**

Ce jeu de données a été extrait de kaggle où il est possible de trouver une description complémentaire des variables.

**Source**

Kaggle.

---

d\_pres2002

*Elections presidentielles 2002 : premier tour*

---

### Description

Jeu de données sur le premier tour des élections présidentielles de 2002. Les noms des lignes indiquent les régions.

### Usage

d\_pres2002

### Format

Un objet data.frame avec 12 variables et 8 observations :

**Arlette.Laguiller** Integer vector, Nombre de votes par région

**Dominique.Voynet** Integer vector, Nombre de votes par région

**François.Bayrou** Integer vector, Nombre de votes par région

**Frédéric.Nihous** Integer vector, Nombre de votes par région

**Gérard.Schivardi** Integer vector, Nombre de votes par région

**Jean.Marie.Le.Pen** Integer vector, Nombre de votes par région

**José.Bové** Integer vector, Nombre de votes par région

**Marie.George.Buffet** Integer vector, Nombre de votes par région

**Nicolas.Sarkozy** Integer vector, Nombre de votes par région

**Olivier.Besancenot** Integer vector, Nombre de votes par région

**Philippe.de.Villiers** Integer vector, Nombre de votes par région

**Ségolène.Royal** Integer vector, Nombre de votes par région

---

d\_pres2007

*Elections presidentielles 2007 : premier tour*

---

### Description

Jeu de données sur le premier tour des élections présidentielles de 2002. Les noms des lignes indiquent les régions.

### Usage

d\_pres2007

**Format**

Un objet `data.frame` avec 12 variables et 23 observations :

**Sarkozy** Integer vector, Nombre de votes par région  
**Bayrou** Integer vector, Nombre de votes par région  
**Royal** Integer vector, Nombre de votes par région  
**Le.Pen** Integer vector, Nombre de votes par région  
**Besanc.** Integer vector, Nombre de votes par région  
**Villiers** Integer vector, Nombre de votes par région  
**Voinet** Integer vector, Nombre de votes par région  
**Laguiller** Integer vector, Nombre de votes par région  
**Bove** Integer vector, Nombre de votes par région  
**Buffet** Integer vector, Nombre de votes par région  
**Nihous** Integer vector, Nombre de votes par région  
**Schivardi** Integer vector, Nombre de votes par région

---

d\_TM

*Tâches ménagères*

---

**Description**

Répartition des tâches ménagères dans différents foyers. Les noms des lignes correspondent aux tâches ménagères.

**Usage**

d\_TM

**Format**

Un objet `data.frame` avec 13 lignes et 4 variables :

**Wife** Integer vector, Nombre de fois que la tâche est réalisée exclusivement par une femme  
**Alternating** Integer vector, Nombre de fois que la tâche est réalisée alternativement par une femme et un homme  
**Husband** Integer vector, Nombre de fois que la tâche est réalisée exclusivement par un homme  
**Jointly** Integer vector, Nombre de fois que la tâche est réalisée en même temps par une femme et un homme

---

d\_vac *Lieux de vacances et occupation*

---

### Description

Ce jeu de données croise le type de séjours de vacances avec l'occupation des individus. L'occupation des individus correspond au nom des lignes du jeu de données.

### Usage

d\_vac

### Format

Un objet `data.frame` avec 8 lignes et 8 variables :

Hotel Integer vector, Séjour à l'hôtel

Location Integer vector, Séjour en location

Res.Second Integer vector, Séjour dans une résidence secondaire

Parents Integer vector, Séjour chez les parents

Amis Integer vector, Séjour chez des amis

Camping Integer vector, Séjour au camping

Sej.org Integer vector, Séjour organisé

Autres Integer vector, Autres

---

d\_wow *Zones de World of Warcraft*

---

### Description

Ce jeu de données est constitué des caractéristiques de zones du jeu World of Warcraft.

### Usage

d\_wow

**Format**

Un objet `data.frame` avec 160 lignes et 12 variables :

`Continent` Factor w/ 7 levels, Nom du continent

`Area` Factor w/ 12 levels, Nom de la région

`Zone` Factor w/ 80 levels, Nom de la zone

`Subzone` Factor w/ 38 levels, Nom de la sous-zone

`Type` Factor w/ 8 levels, Type de zone

`Size` Integer vector, Taille

`Controlled` Factor w/ 5 levels, Faction contrôlant la zone

`Min_req_level` Integer vector, Niveau minimum requis pour pouvoir accéder à la zone

`Min_rec_level` Integer vector, Niveau minimum recommandé pour s'aventurer dans la zone

`Max_rec_level` Integer vector, Niveau maximum recommandé pour s'aventurer dans la zone

`Min_bot_level` Integer vector, Niveau minimum pour pouvoir utiliser un bot dans la zone

`Max_bot_level` Integer vector, Niveau maximum pour pouvoir utiliser un bot dans la zone

**Details**

Ce jeu de données a été extrait de kaggle où il est possible de trouver une description complémentaire des variables.

**Source**

Kaggle.

---

ecole2

*Jeu de données école (2 variables)*

---

**Description**

Jeu de données école avec les résultats des tests de mathématiques et de sport. Ces données ont été simulées.

**Usage**

ecole2

**Format**

Un objet `data.frame` avec 119 lignes et 2 variables :

`Maths` Numerical vector, Note en mathématiques

`Sport` Numerical vector, Note en sport

---

ecole3                      *Jeu de données école (3 variables)*

---

**Description**

Jeu de données école avec non seulement les résultats des tests de mathématiques et de sport mais aussi l'âge des élèves. Ces données ont été simulées.

**Usage**

ecole3

**Format**

Un objet data.frame avec 119 lignes et 3 variables :

Maths Numerical vector, Note en mathématiques

Sport Numerical vector, Note en sport

Age Numerical vector, Âge de l'élève

---

Gauss2F1                      *Fonction hypergéométrique de Gauss (hypergeo)*

---

**Description**

Fonction hypergéométrique de Gauss (hypergeo)

**Usage**

Gauss2F1(a, b, c, x)

**Arguments**

a	Paramètre pour Gauss2F1
b	Paramètre pour Gauss2F1
c	Paramètre pour Gauss2F1
x	Argument principal, nombre réel ou complexe

**Value**

Valeur numérique, l'évaluation de la fonction hypergéométrique de Gauss pour les paramètres a, b, c et x

**Examples**

Gauss2F1(1/2, 1/2, (40-2)/2, 1-.75^2)

---

Gauss2F1gsl	<i>Fonction hypergéométrique de Gauss (gsl)</i>
-------------	---

---

**Description**

Fonction hypergéométrique de Gauss (gsl)

**Usage**

```
Gauss2F1gsl(a, b, c, x)
```

**Arguments**

a	Paramètre pour Gauss2F1
b	Paramètre pour Gauss2F1
c	Paramètre pour Gauss2F1
x	Argument principal, nombre réel ou complexe

**Value**

Valeur numérique, l'évaluation de la fonction hypergéométrique de Gauss pour les paramètres a, b, c et x

**Examples**

```
Gauss2F1gsl(1/2, 1/2, (40-2)/2, 1-.75^2)
```

---

ModStatR	<i>ModStatR</i>
----------	-----------------

---

**Description**

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

**References**

*Modélisation statistique par la pratique avec R*, Frédéric Bertrand, Emmanuelle Claeys, Myriam Maumy-Bertrand, 2019, ISBN:9782100793525, Dunod, Paris, <https://www.dunod.com/sciences-techniques/modelisation-statistique-par-pratique-avec-r-cours-et-exercices-corriges>, <https://github.com/fbertran/ModStatR/> et <https://fbertran.github.io/ModStatR/>

**Examples**

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



---

my.confidence.region *Ellipse ou intervalles de confiance pour une paire de paramètres d'un modèle linéaire*

---

### Description

Ellipse ou intervalles de confiance pour une paire de paramètres d'un modèle linéaire

### Usage

```
my.confidence.region(g, a = 2, b = 3, which = 0, col = "pink")
```

### Arguments

g	Modele linéaire
a	Premier paramètre de l'ellipse
b	Second paramètre de l'ellipse
which	Type de région de confiance : 1 (ellipse autour des deux paramètres), 2 (rectangle autour du premier paramètre, axe des x) et 3 (rectangle autour du second paramètre, axes des y)
col	Couleur de remplissage de la région de confiance

### Exemples

```
data(Mesures, package="BioStatR")
Mes.B = subset(Mesures, Mesures$espece=="bignone")
model2<-lm(masse~taille+I(taille^2), data=Mes.B)
my.confidence.region(model2, which=1)
my.confidence.region(model2, which=2)
my.confidence.region(model2, which=3)
```

---

parasites *Parasites*

---

### Description

Jeu de données sur létalité de la cyperméthrine sur des parasites.

### Usage

```
parasites
```

**Format**

Un objet data.frame avec 12 lignes et 4 variables :

Total Integer vector, Nombre total de parasites

N.morts Integer vector, Nombre de parasites morts

Niveau.de.dose Integer vector, Niveau de dose utilisé

Sexe Factor w/ 2 levels, Sexe de l'hôte

---

perm.cor.mtest	<i>Test par permutation d'une matrice de corrélations de Bravais-Pearson</i>
----------------	--

---

**Description**

Test par permutation d'une matrice de corrélations de Bravais-Pearson

**Usage**

```
perm.cor.mtest(
  mat,
  alternative = "two.sided",
  method = "pearson",
  num.sim = 20000,
  ...
)
```

**Arguments**

mat	Matrice des données
alternative	Type d'hypothèses bilatéral, unilatéral inférieur ou supérieur
method	Méthode de calcul de corrélation Pearson ou Spearman
num.sim	Nombre de simulations
...	Paramètre supplémentaires transmis à la fonction cor

**Value**

Liste de deux éléments : matrice p.mat (matrice des p-valeurs des tests) et matrice cor.mat (matrice des valeurs observées des coefficients de corrélation de Bravais-Pearson)

**Examples**

```
data(Mesures5, package="BioStatR")
Mes5_red_gv = subset(Mesures5[, -5], subset=Mesures5$espece=="glycine violette")
perm.cor.mtest(Mes5_red_gv, num.sim=100)
```

---

poke

*Pokemons*

---

### Description

Jeu de données contenant les caractéristiques de pokemons. Les pokemons concernés apparaissent dans les jeux et non dans les cartes pokemon ou Pokemon Go..

### Usage

poke

### Format

Un objet data.frame avec 800 lignes et 13 variables :

*X*. Integer vector, PokeDex index number

*Name* Factor w/ 800 levels, Name of the Pokemon

*Type.1* Factor w/ 18 levels, Type of pokemon

*Type.2* Factor w/ 18 levels, Other Type of Pokemon

*Total* Integer vector, Sum of Attack, Sp. Atk, Defense, Sp. Def, Speed and HP

*HP* Integer vector, Hit Points

*Attack* Integer vector, Attack Strength

*Defense* Integer vector, Defensive Strength

*Sp..Atk* Integer vector, Special Attack Strength

*Sp..Def* Integer vector, Special Defensive Strength

*Speed* Speed

*Generation* Integer vector, Number of generation

*Legendary* Factor w/ 2 levels, True if Legendary Pokemon False if not (more revision on mythical vs legendary needed)

### Details

Ce jeu de données a été extrait de kaggle où il est possible de trouver une description complémentaire des variables.

### Source

Kaggle.

---

polypes

*Polypes*

---

### Description

Nombre de polypes chez des sujets lors d'un essai clinique.

### Usage

polypes

### Format

Un objet `data.frame` avec 20 lignes et 3 variables :

`nombre` Integer vector, Nombre de polypes après 12 mois

`traitement` Factor w/ 2 levels, Bras de l'essai, un facteur avec deux niveaux placebo et medication

`age` Integer vector, Âge du patient

### Source

M. Giardiello, S. R. Hamilton, A. J. Krush, S. Piantadosi, L. M. Hylind, P. Celano, S. V. Booker, C. R. Robinson and G. J. A. Offerhaus (1993), Treatment of colonic and rectal adenomas with sulindac in familial adenomatous polyposis. *New England Journal of Medicine*, 328(18), 1313–1316. S. Piantadosi (1997), *Clinical Trials: A Methodologic Perspective*. John Wiley & Sons, New York.

---

ref.cor.mtest

*Test exact matriciel du corrélation de Bravais-Pearson avec une référence non nécessairement nulle*

---

### Description

Test exact matriciel du corrélation de Bravais-Pearson avec une référence non nécessairement nulle

### Usage

`ref.cor.mtest(mat, matrho_0)`

### Arguments

`mat` Matrice des données

`matrho_0` Matrice des valeurs de référence pour chacun des coefficients de corrélation de Bravais-Pearson

**Value**

Liste comportant trois matrices : la matrice des p-valeurs, la matrice des coefficients de corrélations observés et la matrice des effectifs ayant servis au calcul de ces coefficients de corrélation

**Examples**

```
data(Mesures5,package="BioStatR")
Mes5_red_lr = subset(Mesures5[,-5],subset=Mesures5$espece=="laurier rose")
ref.cor.mtest(Mes5_red_lr[,c("masse","taille","masse_sec")],0.7)
```

---

ref.cor.test	<i>Test exact du coefficient de corrélation de Bravais-Pearson avec une référence non nécessairement nulle</i>
--------------	--

---

**Description**

Test exact du coefficient de corrélation de Bravais-Pearson avec une référence non nécessairement nulle

**Usage**

```
ref.cor.test(corobs, rho_0, n)
```

**Arguments**

corobs	Valeur observée du coefficient de corrélation de Bravais-Pearson
rho_0	Valeur de référence pour le coefficient de corrélation de Bravais-Pearson
n	Effectif de l'échantillon

**Value**

Valeur numérique, p\$ valeur calculée de manière exacte du test avec la référence rho0 du coefficient de corrélation de Bravais-Pearson sous hypothèse de normalité multivariée

**Examples**

```
ref.cor.test(corobs=.7,rho_0=.8,n=30)
```

---

resistance	<i>Résistance d'un ciment</i>
------------	-------------------------------

---

### Description

Jeu de données contenant les résultats d'une expérience évaluant la résistance d'un ciment.

### Usage

resistance

### Format

Un objet `data.frame` avec 36 lignes et 3 variables :

Melangeur Factor w/ 3, COLUMN\_DESCRIPTION

Casseur Factor w/ 3, COLUMN\_DESCRIPTION

Resistance Integer vector, résistance en livres par pouces carrés

### Details

Davies et Goldsmith ont récolté les données d'une expérience dont le but était d'étudier les différentes sources de variabilité possibles de la résistance d'un ciment fabriqué à Portland. L'expérience s'est déroulée ainsi : plusieurs petits prélèvements d'un même type de ciment ont été mélangés à de l'eau et travaillés par trois personnes différentes, les « mélangeurs ». On a alors formé douze cubes à l'aide de chacune des préparations des « mélangeurs ». Puis on a donné ces 36 cubes à trois personnes chargées d'évaluer leur résistance, les « casseurs ». La répartition des 36 cubes entre ces « casseurs » a été faite de telle sorte que chaque « casseur » reçoive quatre cubes provenant de chacune des préparations des « mélangeurs » soit douze cubes au total. Tous les tests de résistance ont été faits sur la même machine. L'objectif principal de cette expérience était d'étudier et de quantifier l'importance de la variabilité dans les tests de résistance qui pouvait provenir des différences individuelles entre les « mélangeurs » et les « casseurs ». Les données ci-dessous, exprimées dans les unités d'origine c'est-à-dire .

### Source

Davies, O.L. et Goldsmith, P.L. (Eds.), *Statistical Methods in Research and Production*, 4th edition, Oliver and Boyd, Edinburgh, 1972.

---

rho *Corrélation de Bravais-Pearson pour bootstrap ou permutation*

---

**Description**

Corrélation de Bravais-Pearson pour bootstrap ou permutation

**Usage**

```
rho(x, y, indices)
```

**Arguments**

x	un vecteur numérique
y	un vecteur numérique
indices	un vecteur d'indices de même longueur que x et y

**Value**

Valeur numérique, le coefficient de corrélation de Bravais-Pearson calculé pour la permutation des vecteurs x et y spécifiée par le vecteur indices

**Examples**

```
set.seed(1133)  
rho(rnorm(30), rnorm(30), sample(30))
```

---

rho.mult *Matrice de corrélation de Bravais-Pearson, bootstrap ou permutation*

---

**Description**

Matrice de corrélation de Bravais-Pearson, bootstrap ou permutation

**Usage**

```
rho.mult(mat, indices)
```

**Arguments**

mat	Matrice des données
indices	Vecteur d'indices dont la longueur est égale au nombre de lignes de la matrice

**Value**

Matrice des corrélations de Bravais-Pearson des données permutées

**Examples**

```
data(Mesures5,package="BioStatR")
Mes5_red_gv = subset(Mesures5[, -5], subset=Mesures5$espece=="glycine violette")
set.seed(1133)
rho.mult(Mes5_red_gv[,c("masse", "taille", "masse_sec")], sample(nrow(Mes5_red_gv)))
```

---

SidaChat

*Sida du chat*

---

**Description**

Taux de leucocytes

**Usage**

SidaChat

**Format**

Un objet `data.frame` avec 33 lignes et 3 variables :

Sexe Factor w/ 2 levels, Sexe de l'animal

Jours Integer vector, Nombre de jours après l'inoculation

LnT4 Numerical vector, Logarithme népérien du taux de leucocytes T4

**Details**

Le taux de leucocytes T4 chez le chat a été mesuré plusieurs jours (valeur de la variable Jours) après avoir inoculé à l'animal le virus FeLV, analogue au HIV. Nous appelons LnT4 le logarithme népérien de ce taux de leucocytes T4.

**Source**

*Modèle linéaire : Comparaison de groupes et régression* de B. Prum aux Éditions de l'INSERM, 1996.



---

vitamines	<i>Vitamines</i>
-----------	------------------

---

**Description**

Influence de différents régimes alimentaires sur des rats de laboratoire.

**Usage**

vitamines

**Format**

Un objet data.frame avec 32 lignes et 3 variables :

Calorie Integer vector, COLUMN\_DESCRIPTION

Vitamine Integer vector, COLUMN\_DESCRIPTION

Poids Integer vector, COLUMN\_DESCRIPTION

**Details**

Le gain de poids des rats est désigné par la variable Poids, exprimée en grammes, les deux facteurs sont les variables Calorie et Vitamine. La variable Calorie vaut 1 si les rats n'ont pas suivi un régime hypercalorique et 2 s'ils ont suivi un tel régime hypercalorique. La variable Vitamine vaut 1 si les rats n'ont pas reçu de compléments vitaminés et 2 s'ils ont reçu de tels compléments.

**Source**

D'après B. Falissard. *Comprendre et utiliser les statistiques dans les sciences de la vie*. Masson, 2005.

# Index

## \* datasets

CancerSein, 3  
chal, 4  
d\_hotels, 8  
d\_hotels\_n, 9  
d\_macdo, 9  
d\_pres2002, 11  
d\_pres2007, 11  
d\_TM, 12  
d\_vac, 13  
d\_wow, 13  
data\_event, 6  
ecole2, 14  
ecole3, 15  
parasites, 17  
poke, 19  
polypes, 20  
resistance, 22  
SidaChat, 24  
vitamines, 25

boot.mcor.ic, 3

CancerSein, 3  
chal, 4  
corrdist, 4  
corrdistapprox, 5  
corrdistapprox2, 6

d\_hotels, 8  
d\_hotels\_n, 9  
d\_macdo, 9  
d\_pres2002, 11  
d\_pres2007, 11  
d\_TM, 12  
d\_vac, 13  
d\_wow, 13  
data\_event, 6

ecole2, 14

ecole3, 15

Gauss2F1, 15  
Gauss2F1gsl, 16

ModStatR, 16  
my.confidence.region, 17

parasites, 17  
perm.cor.mtest, 18  
poke, 19  
polypes, 20

ref.cor.mtest, 20  
ref.cor.test, 21  
resistance, 22  
rho, 23  
rho.mult, 23

SidaChat, 24

vitamines, 25