Package ‘cccrm’

September 27, 2024

Title Concordance Correlation Coefficient for Repeated (and Non-Repeated) Measures

Version 3.0.2

Date 2024-9-27

Depends R (>= 4.0)

Imports nlme, dplyr, Deriv, tidyselect, progressr, furrr, nlmeU, parallelly, purrr, tidyr, lifecycle, future

Description Estimates the Concordance Correlation Coefficient to assess agreement. The scenarios considered are non-repeated measures, non-longitudinal repeated measures (replicates) and longitudinal repeated measures. It also includes the estimation of the one-way intraclass correlation coefficient also known as reliability index. The estimation approaches implemented are variance components and U-statistics approaches. Description of methods can be found in Fleiss (1986) <doi:10.1002/9781118032923> and Carrasco et al. (2013) <doi:10.1016/j.cmpb.2012.09.002>.

License GPL (>= 2)

RoxygenNote 7.3.2

Encoding UTF-8

LazyData true

NeedsCompilation no

Author Josep Lluis Carrasco [aut, cre], Gonzalo Peon Pena [aut]

Maintainer Josep Lluis Carrasco <jlcarrasco@ub.edu>

Repository CRAN

Date/Publication 2024-09-27 10:10:03 UTC

Contents

`bdaw` *Blood draw data*

Description

Plasma cortisol area under curve (AUC) was calculated from the trapezoidal rule over the 12-h period of the hourly blood draws. The subjects were required to repeat the process in five visits. The aim of the agreement study was to assess how well the plasma cortisol AUC from hourly measurements agreed with plasma cortisol AUC that was measured every two hours.

Usage

```
bdaw
```

Format

A data frame with the following columns:

SUBJ Subject identifier

VNUM Visit number

AUC Area under the curve

MET Device identifier

`bfat` *Body fat data*

Description

Percentage body fat was estimated from skinfold calipers and DEXA on a cohort of 90 adolescent girls. Skinfold caliper and DEXA measurements were taken at ages 12.5, 13 and 13.5. The objective was to determine the amount of agreement between the skinfold caliper and DEXA measurements of percentage body fat.

Usage

```
bfat
```

Format

A data frame with the following columns:

SUBJECT Subject identifier

VISITNO Visit number

BF Percentage body fat

MET Device identifier

`bpres` *Blood pressure data*

Description

Systolic and diastolic blood pressure was measured in a sample of 384 subjects using a handle mercury sphygmomanometer device and an automatic device. The blood pressure was simultaneously measured twice by each instrument, thus every subject had four measurements, two by each method.

Usage

```
bpres
```

Format

A data frame with the following columns:

ID Subject identifier

SIS Systolic blood pressure in mmHg

DIA Diastolic blood pressure in mmHg

METODE Device identifier

NM Identifier of replicates

ALTURA Height in cm

EDAD Age in years

FRECUENC Heart rate

INFOR_AR Have the subject been informed about he is hypertense?

PESO Weight in Kg

SEXO Gender. 1 for Male. 2 for Female

TA Was the subject's blood pressure measured the last year? 1=Yes, 2=No, 9=Unknown

TNSI_MED Does the subject receive treatment for hypertension? 1=Yes, 2=No, 3=Doubtful, 8=not applicable, 9=Insufficient data.

cccUst — *Repeated Measures Concordance Correlation Coefficient estimated by U-statistics*

Description

Estimation of the concordance correlation coefficient for repeated measurements using the U-statistics approach. The function is also applicable for the non-repeated measurements scenario.

Usage

```
cccUst(dataset, ry, rmet, rtime = NULL, Dmat = NULL, delta = 1, cl = 0.95)
```

Arguments

`dataset`	An object of class `data.frame`.
`ry`	Character string. Name of the outcome in the data set.
`rmet`	Character string. Name of the method variable in the data set.
`rtime`	Character string. Name of the time variable in the data set.
`Dmat`	Matrix of weights.
`delta`	Power of the differences. A value of 0 provides an estimate that is comparable to a repeated measures version of kappa index.
`cl`	Confidence level.

Value

A vector that includes the point estimate, confidence interval and standard error of the CCC. Additionally the Fisher's Z-transformation value and its standard error are also provided.

References

King, TS and Chinchilli, VM. (2001). A generalized concordance correlation coefficient for continuous and categorical data. Statistics in Medicine, 20, 2131:2147.

King, TS; Chinchilli, VM; Carrasco, JL. (2007). A repeated measures concordance correlation coefficient. Statistics in Medicine, 26, 3095:3113.

Carrasco, JL; Phillips, BR; Puig-Martinez, J; King, TS; Chinchilli, VM. (2013). Estimation of the concordance correlation coefficient for repeated measures using SAS and R. Computer Methods and Programs in Biomedicine, 109, 293-304.

Examples

```
# Non-longitudinal scenario
newdat=bpres[bpres$NM==1,]
estccc=cccUst(newdat,"DIA","METODE")
estccc

estccc=cccUst(bdaw,"AUC","MET","VNUM")
estccc

estccc=cccUst(bfat,"BF","MET","VISITNO",Dmat=diag(c(2,1,1)))
estccc
```

`ccc_sim_data` *Data simulation using fixed and random effects*

Description

The fixed effects and standard deviations of random effects can be set to specific values or, alternatively, obtained from an object of class `lme`.

Usage

```
ccc_sim_data(
  n = 30,
  nrep = 1,
  nsim = 1,
  model = NULL,
  b = NULL,
  g = NULL,
  mu = 0,
  sa = 1,
  sab = 0,
  sag = 0,
  bg = NULL,
  se = 1,
  future_seed = TRUE,
  workers = 15,
  extra.info = TRUE,
  ...
)
```

Arguments

`n`	Integer. Number of subjects
`nrep`	Integer. Number of replicates
`nsim`	Integer. Number of data sets simulated.
`model`	Object of class `lme`.

b	Vector. Method fixed effects.
g	Vector. Time fixed effects.
mu	Integer. Overall mean.
sa	Integer. Standard deviation of subject's random effect.
sab	Integer. Standard deviation of subject-method interaction's random effect.
sag	Integer. Standard deviation of subject-time interaction's random effect.
bg	Vector. Method-time interaction's fixed effects. The vector of effects have to be ordered by method and time.
se	Integer. Standard deviation of random error effect.
future_seed	Logical/Integer. The seed to be used for parallellization. Further details in furrr_options.
workers	Integer. Number of cores to be used for parallellization. Default is 15. Capped to number of available cores minus 1.
extra.info	Logical. Should the information about CCC/ICC and variance components simulated be shown? Default is set to TRUE.
...	To pass further arguments.

Details

Random effects are simulated as normal distributions with mean 0 and the correspondign standard deviations. The simulated data is obtained as the addition of the simulated values and the fixed efffects. Parallel computation is used except if data is simulated from an object of class 'lme'. In this case. data is simulated using the simulateY function from nlmeU package.

Value

A data frame with the simulated data.

See Also

ccc_vc

Examples

```
# # Reliability data:
# 50 subjects, one method, one time, 2 replicates
# Overall mean: -0.25; Subjects standard deviation: 1.5, Random error standard deviation: 1
set.seed(101)
df <- ccc_sim_data(n=50, b = NULL, g = NULL, mu = -0.25, sa = 1.5, se = 1, nrep=2)

# Method comparison data (non-longitudinal)
# 50 subjects, two methods, 2 replicates
# Overall mean: -0.25; Subjects standard deviation: 1.5, Random error standard deviation: 1
# Difference of means between methods 2 and 1: 1
# Three data sets simulated

set.seed(202)
```

```
df <- ccc_sim_data(n=50, nsim=3,b = c(0,1), mu = -0.25, sa = 1.5, se = 1, nrep=2)

# Method comparison data (longitudinal)
# 50 subjects, two methods, 3 times, 1 replicate,
# Overall mean: -0.25; Subjects standard deviation: 1.5, Random error standard deviation: 1
# Difference of means between methods 2 and 1: 1
# Difference of means between times 3,2 and 1 respectively: 0.5 and 0.25.
# Subject-methods interaction standard deviation: 0.25
# Subject-times interaction standard deviation: 0.5
# Same difference of means at each time

set.seed(202)
df <- ccc_sim_data(n=50, b = c(0,1), g=c(0,0.25,0.5), mu = -0.25, sa = 1.5,
sab=0.25,sag=0.5,se = 1, nrep=2)

# Simulate data using the estimates of a linear mixed model
set.seed(2024)
df3 <- ccc_sim_data(n=50, b = c(0,1), g=c(0,0.25,0.5), mu = -0.25, sa = 1.5,
                    sab=0.25,sag=0.5,bg=c(0,0.5,0.75,0,1,1),se = 1, nrep=2)
mod3 <- lme_model(df3,"y","id","times","met",control.lme=nlme::lmeControl(opt = 'optim'))
ccc_sim_data(nsim=10,model=mod3)
```

`ccc_vc` *Concordance Correlation Coefficient estimation by variance components.*

Description

Estimation of the concordance correlation coefficient for either non-repeated, non-longitudinal, or longitudinal repeated measurements using the variance components from a linear mixed model. The appropriate intraclass correlation coefficient is used as estimator of the concordance correlation coefficient.

Usage

```
ccc_vc(
  dataset,
  ry,
  rind,
  rmet = NULL,
  rtime = NULL,
  vecD = NULL,
  covar = NULL,
  int = F,
  rho = 0,
  cl = 0.95,
  control.lme = list(),
```

```
    transf = "F2",
    boot = FALSE,
    boot_param = FALSE,
    boot_ci = "BCa",
    nboot = 300,
    parallel = FALSE,
    future_seed = TRUE,
    workers = 15,
    sd_est = TRUE,
    apVar = TRUE,
    ...
)
```

Arguments

- `dataset` an object of class `data.frame`.
- `ry` Character string. Name of the outcome in the data set.
- `rind` Character string. Name of the subject variable in the data set.
- `rmet` Character string. Name of the method variable in the data set.
- `rtime` Character string. Name of the time variable in the data set.
- `vecD` Vector of weights. The length of the vector must be the same as the number of repeated measures.
- `covar` Character vector. Name of covariates to include in the linear mixed model as fixed effects.
- `int` Binary indicating if the subject-method interaction has to be included in the model when analyzing the non-longitudinal setting (defaults to FALSE).
- `rho` Within subject correlation structure. A value of 0 (default option) stands for compound symmetry and 1 is used for autorregressive of order 1 structure.
- `cl` Confidence level.
- `control.lme` A list of control values for the estimation algorithm used in `lme` function. For further details see `lme` help.
- `transf` Character string. Whether to apply a transformation of the coefficient for inference. Valid options are: "F" for Fisher's Z-transformation; "F2" For Fisher's Z-transformation setting m=2 (default); "KG" Konishi-Gupta transformation; "None", no transformation is applied. See *Details* for further information.
- `boot` Logical. Whether to compute the CCC confidence interval by bootstrapping or asymptotic methods (defaults to FALSE).
- `boot_param` Logical. Whether to compute a parametric bootstrap or a non-parametric bootstrap (defaults to FALSE).
- `boot_ci` Character. Type of bootstrap confidence interval. Either "BCa" (which is the default) or "empirical".
- `nboot` Integer. Number of bootstrap resamples. Default is 300.
- `parallel` Logical. Whether the code is parallellized. The parallellization method is `multisession`.

`future_seed`	Logical/Integer. The seed to be used for parallellization. Further details in `furrr_options`.
`workers`	Integer. Number of cores to be used for parallellization. Default is 15. Capped to number of available cores minus 1.
`sd_est`	Logical. Whether to estimate the asymptotic standard deviation (defaults to TRUE) or to only report the `ccc/icc` value.
`apVar`	Logical. Should the asymptotic variance-covariance matrix of the variance components be estimated in the linear mixed model? (Defaults to TRUE).
`...`	To pass further arguments.

Details

The concordance correlation coefficient is estimated using the appropriate intraclass correlation coefficient (see Carrasco and Jover, 2003; Carrasco et al., 2009; Carrasco et al, 2013).

The scenarios considered are: a) reliability assessment (several measurements taken with one method); b) methods comparison data with non-repeated measurements (only one measurement by subject and method); c) Methods comparison data with non-longitudinal repeated measurements, i.e. replicates (multiple measurements by subject and method); and d) Methods comparison data with longitudinal repeated measurements (multiple longitudinal measurements by subject and method).

The variance components estimates are obtained from a linear mixed model (LMM) estimated by restricted maximum likelihood. The function *lme* from package *nlme* (Pinheiro et al., 2021) is used to estimate the LMM.

The standard error of CCC and its confidence interval can be obtained: a) asymptotically, using Taylor's series expansion of 1st order (Ver Hoef, 2012); b) using balanced randomized cluster bootstrap approach (Davison and Hinkley, 1997; Field and Welsh, 2007); c) using parametric bootstrap (Davison and Hinkley, 1997).

When estimating asymptotically the standard error, the confidence intervals are built using the point estimate of the CCC/ICC, its standard error, and the appropriate quantile of the standard Normal distribution. However, the approximation to the asymptotic Normal distribution is improved if the CCC/ICC is transformed using the Fisher's Z-transformation (Fisher, 1925), or the Konishi-Gupta transformation (Konishi and Gupta, 1989). In case the number of replicates is equal to 2, both transformations give the same result.

Value

A `ccc` class object. Generic function `summary` show a summary of the results. The output is a list with the following components:

- `ccc`. CCC/ICC estimate
- `model`. nlme object with the fitted linear mixed model.
- `vc`. Variance components estimates.
- `sigma`. Variance components asymptotic covariance matrix.

References

Carrasco, JL; Jover, L. (2003). Estimating the generalized concordance correlation coefficient through variance components. Biometrics, 59, 849:858.

Carrasco, JL; King, TS; Chinchilli, VM. (2009). The concordance correlation coefficient for repeated measures estimated by variance components. Journal of Biopharmaceutical Statistics, 19, 90:105.

Davison A.C., Hinkley D.V. (1997). Bootstrap Methods and Their Application. Cambridge: Cambridge University Press.

Field, C.A., Welsh, A.H. (2007). Bootstrapping Clustered Data. Journal of the Royal Statistical Society. Series B (Statistical Methodology). 69(3), 369-390.

Fisher, R. A. (1925) Statistical Methods for Research Workers. Edinburgh: Oliver

Konishi, S. and Gupta, A. K. (1989) Testing the equality of several intraclass correlation coefficients. J Statist. Planng Inf., 21, 93-105.

Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team (2021). nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-152, https://CRAN.R-project.org/package=nlme.

Ver Hoef, J.M. (2012) Who Invented the Delta Method?, The American Statistician, 66:2, 124-127.

Examples

```
# Scenario 1. Reliability
newdat <- bpres |> dplyr::filter(METODE==1)
icc_rel<-ccc_vc(newdat,"DIA","ID")
icc_rel
summary(icc_rel)

# Confidence interval using non-parametric bootstrap

icc_rel_bt<-ccc_vc(newdat,"DIA","ID",boot=TRUE,sd_est=FALSE,
nboot=500,parallel=TRUE)
icc_rel_bt
summary(icc_rel_bt)


# Scenario 2. Non-longitudinal methods comparison.
# Only 1 measure by subject and method.
# No subjects-method interaction included in the model.

newdat <- bpres |> dplyr::filter(NM==1)
ccc_mc<-ccc_vc(newdat,"DIA","ID","METODE")
ccc_mc
summary(ccc_mc)

# Confidence interval using parametric bootstrap

ccc_mc_bt<-ccc_vc(newdat,"DIA","ID",boot=TRUE,boot_param=TRUE,
sd_est=FALSE,nboot=500,parallel=TRUE)
ccc_mc_bt
summary(ccc_mc_bt)
```

```
# Scenario 3. Non-longitudinal methods comparison.
# Two measures by subject and method.
# No subject-method interaction included in the model.

ccc_mc_int=ccc_vc(bpres,"DIA","ID","METODE",int=TRUE,control.lme=nlme::lmeControl(opt = 'optim'))
ccc_mc_int
summary(ccc_mc_int)

# Scenario 4. Methods comparison in longitudinal repeated measures setting.
ccc_mc_lon<-ccc_vc(bdaw,"AUC","SUBJ","MET","VNUM")
ccc_mc_lon
summary(ccc_mc_lon)

# Scenario 5. Methods comparison in longitudinal repeated measures setting.
# More weight given to readings from first time.

ccc_mc_lonw<-ccc_vc(bfat,"BF","SUBJECT","MET","VISITNO",vecD=c(2,1,1))
ccc_mc_lonw
summary(ccc_mc_lonw)
```

`lme_model` *Fits a Linear Mixed Effects Model*

Description

Fits a Linear Mixed Effects Model

Usage

```
lme_model(
  dataset,
  ry,
  rind,
  rtime = NULL,
  rmet = NULL,
  vecD = NULL,
  covar = NULL,
  rho = 0,
  int = FALSE,
  cl = 0.95,
  control.lme = list(),
  apVar = TRUE,
  ...
)
```

Arguments

`dataset`	an object of class `data.frame`.
`ry`	Character string. Name of the outcome in the data set.
`rind`	Character string. Name of the subject variable in the data set.
`rtime`	Character string. Name of the time variable in the data set.
`rmet`	Character string. Name of the method variable in the data set.
`vecD`	Vector of weights. The length of the vector must be the same as the number of repeated measures.
`covar`	Character vector. Name of covariates to include in the linear mixed model as fixed effects.
`rho`	Within subject correlation structure. A value of 0 (default option) stands for compound symmetry and 1 is used for autoregressive of order 1 structure.
`int`	Boolean indicating if the subject-method interaction has to be included in the model.
`cl`	Confidence level.
`control.lme`	A list of control values for the estimation algorithm used in `lme` function. For further details see `lme` help.
`apVar`	Logical. Should the asymptotic variance-covariance matrix of the variance components be estimated in the linear mixed model? (Defaults to TRUE).
`...`	To pass further arguments.

Value

an object of class `lme`.

Examples

```
# Reliability ICC
set.seed(2024)
df <- ccc_sim_data(b = NULL, g = NULL, mu = -0.25, sa = 1.5, se = 1)
mod1 <- lme_model(df,"y","id")
mod1

#Non-longitudinal Methods comparison data
set.seed(2024)
df2 <- ccc_sim_data(n=50,b = c(0,1), mu = -0.25, sa = 1.5, se = 1, nrep=2)
mod2 <- lme_model(df2,"y","id",rmet="met")
mod2

# Longitudinal Methods comparison data
set.seed(2024)
df3 <- ccc_sim_data(n=50, b = c(0,1), g=c(0,0.25,0.5), mu = -0.25, sa = 1.5,
                    sab=0.25,sag=0.5,bg=c(0,0.5,0.75,0,1,1),se = 1, nrep=2)

mod3 <- lme_model(df3,"y","id","times","met",control.lme=nlme::lmeControl(opt = 'optim'))
mod3
```

sim_power_ccc *Power and confidence interval range*

Description

Power and confidence interval range obtained by simulation

Usage

```
sim_power_ccc(
  n = 30,
  nrep = 2,
  nsim = 300,
  r0 = 0,
  alpha = 0.05,
  model = NULL,
  b = NULL,
  g = NULL,
  mu = 0,
  sa = 1,
  sab = 0,
  sag = 0,
  bg = NULL,
  se = 1,
  extra.info = TRUE,
  vecD = NULL,
  covar = NULL,
  int = FALSE,
  rho = 0,
  cl = 0.95,
  control.lme = list(),
  transf = "F2",
  future_seed = TRUE,
  workers = 15
)
```

Arguments

n	Integer. Number of subjects
nrep	Integer. Number of replicates
nsim	Integer. Number of data sets simulated.
r0	Integer. Null hypothesis value.
alpha	Type-I error rate.
model	object of class `lme`.
b	Vector. Method fixed effects.

g	Vector. Time fixed effects.
mu	Integer. Overall mean.
sa	Integer. Standard deviation of subject's random effect.
sab	Integer. Standard deviation of subject-method interaction's random effect.
sag	Integer. Standard deviation of subject-time interaction's random effect.
bg	Vector. Method-time interaction's fixed effects
se	Integer. Standard deviation of random error effect.
extra.info	Logical. Should the information about CCC and variance components simulated be shown? Default is set to TRUE.
vecD	Vector of weights. The length of the vector must be the same as the number of repeated measures.
covar	Character vector. Name of covariates to include in the linear mixed model as fixed effects.
int	Binary indicating if the subject-method interaction has to be included in the model when analyzing the non-longitudinal setting (defaults to FALSE).
rho	Within subject correlation structure. A value of 0 (default option) stands for compound symmetry and 1 is used for autorregressive of order 1 structure.
cl	Confidence level.
control.lme	A list of control values for the estimation algorithm used in `lme` function. For further details see `lme` help.
transf	Character string. Whether to apply a transformation of the coefficient for inference. Valid options are: "F" for Fisher's Z-transformation; "F2" For Fisher's Z-transformation setting m=2 (default); "KG" Konishi-Gupta transformasion; "None", no transformation is applied. See *Details* for further information.
future_seed	Logical/Integer. The seed to be used for parallellization. Further details in `furrr_options`.
workers	Integer. Number of cores to be used for parallellization. Default is 15. Capped to number of available cores minus 1.

Details

The power and the range of the confidence interval are computed using the approach suggested in Choudhary and Nagaraja (2018). Data sets are simulated by setting the fixed effects values and the standard deviation of the random effects. The CCC and its standard error are estimated in each data set, along with its 95% confidence interval and the Wald test `Ztest`.

Value

A data frame with the following components:

- `n` Number of subjects
- `reps` Number of replicates
- `CCC`. Median of the CCC estimates.

- `Power`. Empirical power computed as proportion of times the null hypothesis is rejected using a type-I error rate of `alpha`.
- `SEICC`. Average of CCC standard errors.
- `SEZ`. Average of transformed CCC standard errors.
- `Range IC95`. Average of CCC confidence interval widths.

References

Choudhary, P.K. and Nagaraja, H.N. (2018). Measuring Agreement-Models, Methods, and Applications. John Wiley & Sons

Examples

```
# Power to test the CCC is above 0.8 with 35 subjects and 4 replicates.
# Two methods, three times. Simulated CCC=0.87.
sim_pw<-sim_power_ccc(n = 35, nrep=4, nsim=500, r0=0.8, b = c(-0.5,0.5),
g=c(-0.25,0,0.25), mu = -0.25, sa = 4,sab=0.5,sag=1,
bg=c(-0.5,-0.25,0.25,-0.5,0.25,0.75),se = 1)
```

`Ztest` *Wald's test on the Concordance Correlation Coefficient*

Description

Estimation of the concordance correlation coefficient for either non-repeated, non-longitudinal, or longitudinal repeated measurements using the variance components from a linear mixed model. The appropriate intraclass correlation coefficient is used as estimator of the concordance correlation coefficient.

Usage

```
Ztest(cccfit, r0 = 0, tr = TRUE)
```

Arguments

`cccfit`	An object of class `ccc`.
`r0`	Integer. Null hypothesis value.
`tr`	Logical. Should the transformed CCC be used? Only applies if a transformation was applied in the `ccc` object. Default value is TRUE.

Details

A one sided test to the null hypothesis value

$$\rho_0$$

.

$$z = \frac{\hat{\theta} - \rho_0}{SE\left(\hat{\theta}\right)}$$

where

$$\hat{\theta}$$

stands for the CCC estimate and

$$SE\left(\hat{\theta}\right)$$

its standard error. The p-value is computed as

$$P\left(X > z\right)$$

where X follows a standard Normal distribution.

Value

A data frame with two columns: Z, the statistical test value; and the P-value associated.

Examples

```
# Testing the CCC is above 0.8
ccc_mc=ccc_vc(bpres,"DIA","ID","METODE")
ccc_mc
Ztest(ccc_mc,r0=0.8)
```

Index