Package 'CrossCarry'

September 25, 2024

Type Package

Title Analysis of Data from a Crossover Design with GEE

Version 0.4.0

Description Analyze data from a crossover design using generalized estimation equations (GEE), including carryover effects and various correlation structures based on the Kronecker product. It contains functions for semiparametric estimates of carry-over effects in repeated measures and allows estimation of complex carry-over effects. Related work includes: a) Cruz N.A., Melo O.O., Martinez C.A. (2023). ``CrossCarry: An R package for the analysis of data from a crossover design with GEE''. <doi:10.48550/arXiv.2304.02440>. b) Cruz N.A., Melo O.O., Martinez C.A. (2023). ``A correlation structure for the analysis of Gaussian and non-Gaussian responses in crossover experimental designs with repeated measures''. <doi:10.1007/s00362-022-01391-z> and c) Cruz N.A., Melo O.O., Martinez C.A. (2023). ``Semiparametric generalized estimating equations for repeated measurements in cross-over designs''. <doi:10.1177/09622802231158736>.

License GPL (>= 3)

Encoding UTF-8

LazyData true

Depends R (>= 4.0)

Imports dplyr, gee, ggplot2, splines, stats, MASS

Suggests testthat (>= 3.0.0)

Config/testthat/edition 3

RoxygenNote 7.3.2

NeedsCompilation no

Author Nelson Alirio Cruz Gutierrez [aut, cre, cph],
Oscar Orlando Melo [aut],
Carlos Alberto Martinez [aut]

Maintainer Nelson Alirio Cruz Gutierrez <nelson-alirio.cruz@uib.es>

Repository CRAN

Date/Publication 2024-09-24 22:10:02 UTC

Contents

Arterial   *Arterial pressure data of crossover design*

Description

Data of a crossover experiment with three treatments to control arterial pressure: treatment A is a placebo, treatments B and C are 20 and 40 mg doses of a test drug. Thus, there were six three-period sequences: ABC, ACB, BCA, BAC, CAB and CBA, each one of them was applied to two individuals lasting six weeks each. Due to ethical reasons, there was no washout period between the treatments. In each period, 10 consecutive measurements of diastolic arterial pressure were taken: 30 and 15 minutes before, and 15, 30, 45, 60, 75, 90, 120 and 240 minutes after the administration of the treatment

Usage

```
Arterial
```

Format

A data frame with 360 rows and 5 columns:

Subject The unique identifier of each of the patients

Period The period of application of each treatment

Treatment A is a placebo, treatments B and C are 20 and 40 mg doses of a test drug

Pressure diastolic arterial pressure

Time Measurement time

Source

Jones, B. and Kenward, M. G. (2015). Design and Analysis of Cross-Over Trials Third Edition. Chapman & Hall/CRC, Boca Raton

---

| computeqic | *Quasi Information Criterion* |

Description

Function for calculating the quasi-likelihood under the independence model information criterion (QIC), quasi-likelihood, correlation information criterion (CIC), and corrected QIC for one or several fitted geeglm model object from the geepack package.

Usage

```
computeqic(object)
```

Arguments

object   a fitted GEE model from the gee package.

Details

QIC is used to select a correlation structure. The QICu is used to compare models that have the same working correlation matrix and the same quasi-likelihood form but different mean specifications. CIC has been suggested as a more robust alternative to QIC when the model for the mean may not fit the data very well and when models with different correlation structures are compared.

Models with smaller values of QIC, CIC, QICu, or QICC are preferred.

Value

A vector or matrix with the QIC, QICu, quasi likelihood, CIC, the number of mean effect parameters, and the corrected QIC for each GEE object

Author(s)

Alirio Cruz <nelson-alirio.cruz@uib.es>, Claus Ekstrom <claus@rprimer.dk>, Brian McLoone <bmcloone@pdx.edu> and Steven Orzack <orzack@freshpond.org>

References

Pan, W. (2001). Akaike's information criterion in generalized estimating equations. Biometrics, 57, 120-125. Hardin, J.W. and Hilbe, J.M. (2012). Generalized Estimating Equations, 2nd Edition, Chapman and Hall/CRC: New York.

Hin, L.-Y. and Wang, Y-G. (2009). Working-correlation-structure identification in generalized estimating equations, Statistics in Medicine 28: generalized estimating equations, Statistics in Medicine 28: 642-658. Thall, P.F. and Vail, S.C. (1990). Covariance Models for Longitudinal Count Data with Overdispersion. Biometrics, 46, 657-671.

Examples

```
library(gee)
data(Arterial)
fit <- gee(Pressure ~ Time + Treatment, id=Subject,
        data=Arterial, family=gaussian, corstr="AR-M")
computeqic(fit)
```

---

createCarry   *Add carryover dummy variables*

Description

Create dummy variables associated with first-order carryover effect in a Crossover Design

Usage

```
createCarry(data, treatment, period, id, carrySimple = TRUE)
```

Arguments

data	A data frame with the variables of the crossover experimental design
treatment	A character string specifying the name of vector with the treatment applied at each observation of the crossover experimental design
period	A character string specifying the name of vector with the observation period of the responses of the crossover experimental design
id	A character string specifying the name of vector which identifies the experimental units.
carrySimple	TRUE = simple carry-over, where the residual effect of a treatment affects equally each of the treatments that are preceded by it, and FALSE = carry-over complex, where the residual effect of the treatment affects each of the other treatments differently.

Value

data A data frame with all the variables of the crossover experimental design and the carryover variables

carryover The new carryover variables

Examples

```
data(Water)
carryover <- createCarry(data=Water,
                        treatment = "Treatment", id = "ID",
                        period = "Period", carrySimple = FALSE)
carryover$carryover
carryover$data
```

CrossGEE	*Run a GEE model for data from a crossover experiment*

Description

Provides a GEE model for the data of a crossover design with S sequences of T periods. There must be one observation of each experimental unit in each period.

Usage

```
CrossGEE(
  response,
  period,
  treatment,
  id,
  carry,
  covar = NULL,
  data,
  family = gaussian(),
  correlation = "independence",
  formula = NULL,
  Mv = 1
)
```

Arguments

response	A character string specifying the name of the response variable of the crossover experimental design
period	A character string specifying the name of vector with the observation period of the responses of the crossover experimental design
treatment	A character string specifying the name of vector with the treatment applied at each observation of the crossover experimental design
id	A character string specifying the name of vector which identifies the experimental units. The length of 'id' should be the same as the number of observations. Data are assumed to be sorted so that observations on each cluster appear as contiguous rows in data. If data is not sorted this way, the function will not identify the clusters correctly. If data is not sorted this way, a warning will be issued.
carry	A vector of character string specifying the name set of dummy variables that indicates the treatment applied in the previous period of each experimental unit. They must be 0 in period 1
covar	A vector of character string specifying the name of possible covariates of the crossover experimental design
data	A data frame with all the variables of the crossover experimental design
family	See corresponding documentation to glm

correlation	a character string specifying the correlation structure. The following are permitted: "independence", "fixed", "stat_M_dep", "non_stat_M_dep", "exchangeable", "AR-M" and "unstructured"
formula	A formula related the response variable with the explanatory variables. If it is NULL, formula `response~period+treatment+carry+covar` will be evaluated
Mv	When correlation is "stat_M_dep", "non_stat_M_dep", or "AR-M" then Mv must be specified.

Value

`QIC` The QIC of the models: The model are fitted by `geeglm`

`model` The model fitted by `geeglm`.

Source

https://doi.org/10.1111/stan.12295

References

Cruz, N. A., López Pérez, L. A., & Melo, O. O. (2023). Analysis of cross-over experiments with count data in the presence of carry-over effects. Statistica Neerlandica, 77(4), 516-542.

Examples

```
data(Water)
model <- CrossGEE(response="LCC", covar=c("Age"), period="Period",
                  treatment = "Treatment", id="ID", carry="Carry_Agua",
                  family=gaussian(),correlation ="AR-M", Mv=1 ,data=Water)

model$QIC
model$model

## Aproximate p-values
(pvalues <- 2 * pnorm(abs(coef(summary(model$model))[,5]), lower.tail = FALSE))

summary(model$model)
```

CrossGEEKron	*Run a GEE model for data from a crossover experiment with repeated measures*

Description

Provides a GEE model for the data of a crossover design with S sequences of T periods. There must be at least two observations of each experimental unit in each period.

Usage

```
CrossGEEKron(
  response,
  period,
  treatment,
  id,
  time,
  carry,
  covar = NULL,
  data,
  family = gaussian(),
  correlation = "independence",
  formula = NULL,
  tol = 1e-04,
  niter = 100,
  Mv
)
```

Arguments

response	A character string specifying the name of the response variable of the crossover experimental design
period	A character string specifying the name of vector with the observation period of the responses of the crossover experimental design
treatment	A character string specifying the name of vector with the treatment applied at each observation of the crossover experimental design
id	A character string specifying the name of vector which identifies the experimental units. The length of 'id' should be the same as the number of observations. Data are assumed to be sorted so that observations on each cluster appear as contiguous rows in data. If data is not sorted this way, the function will not identify the clusters correctly. If data is not sorted this way, a warning will be issued.
time	A character string specifying the name of the vector with the measurement time within each period
carry	A vector of character string specifying the name set of dummy variables that indicates the treatment applied in the previous period of each experimental unit. They must be 0 in period 1
covar	A vector of character string specifying the name of possible covariates of the crossover experimental design
data	A data frame with all the variables of the crossover experimental design
family	See corresponding documentation to `glm`.
correlation	a character string specifying the correlation structure. The following are permitted: "independence", "fixed", "stat_M_dep", "non_stat_M_dep", "exchangeable", "AR-M" and "unstructured"
formula	A formula related the response variable with the explanatory variables. If it is NULL the formula, `response~period+treatment+carry+time+covar` will be evaluated

tol	the tolerance used in the fitting algorithm.
niter	the maximum number of iterations. response~period+treatment+carry+time+covar will be evaluated
Mv	When correlation is "stat_M_dep", "non_stat_M_dep", or "AR-M" then Mv must be specified.

Value

`QIC` The QIC of the model: The model are fitted by `geeglm`

`model` The model fitted by `geeglm`.

`Within` The estimated correlation matrix within the period with the structure determined by `correlation`.

`Between` The estimated correlation matrix between periods

Source

https://doi.org/10.1007/s00362-022-01391-z

References

Cruz, N.A., Melo, O.O. & Martinez, C.A. A correlation structure for the analysis of Gaussian and non-Gaussian responses in crossover experimental designs with repeated measures. Statistical Papers 65, 263–290 (2024)

Examples

```
data(Arterial)

carrydata <- createCarry(data=Arterial, treatment = "Treatment",
 period = "Period",id="Subject")

data <- carrydata$data
carry <- carrydata$carryover
model <- CrossGEEKron(response = "Pressure", treatment = "Treatment",
period = "Period", id="Subject", time="Time",
 carry=c("Carry_B","Carry_C"),data=data, correlation = "AR-M", Mv=1)

model$QIC
model$Within
model$Between
summary(model$model)

## Aproximate p-values for model
(pvalues <- 2 * pnorm(abs(coef(summary(model$model))[,5]), lower.tail = FALSE))

model2 <- CrossGEEKron(response = "Pressure", treatment = "Treatment",
 period = "Period", id="Subject", time="Time",
 carry=c("Carry_B","Carry_C"), data=data,
 correlation = "AR-M", Mv=1,formula=Pressure ~ Treatment+
 Period+ Carry_B+Carry_C)
```

```
model2$QIC
model2$Within
model2$Between
summary(model2$model)
```

CrossGEESP	*Run a semi-parametric GEE model for data from a crossover experiment with repeated measures*

Description

Provides a GEE model for the data of a crossover design with S sequences of T periods. There must be at least two observations of each experimental unit in each period. The effect of time within period and the possible carryover effects are modeled by means of splines.

Usage

```
CrossGEESP(
  response,
  period,
  treatment,
  id,
  time,
  carry,
  covar = NULL,
  data,
  family = gaussian,
  correlation = "independence",
  formula = NULL,
  tol = 1e-04,
  niter = 100,
  nodes = NULL,
  Mv = 1
)
```

Arguments

response	A character string specifying the name of the response variable of the crossover experimental design
period	A character string specifying the name of vector with the observation period of the responses of the crossover experimental design
treatment	A character string specifying the name of vector with the treatment applied at each observation of the crossover experimental design

id	A character string specifying the name of vector which identifies the experimental units. The length of 'id' should be the same as the number of observations. Data are assumed to be sorted so that observations on each cluster appear as contiguous rows in data. If data is not sorted this way, the function will not identify the clusters correctly. If data is not sorted this way, a warning will be issued.
time	A character string specifying the name of the vector with the measurement time within each period
carry	A vector of character string specifying the name set of dummy variables that indicates the treatment applied in the previous period of each experimental unit. They must be 0 in period 1
covar	A vector of character string specifying the name of possible covariates of the crossover experimental design
data	A data frame with all the variables of the crossover experimental design
family	See corresponding documentation to `glm`
correlation	a character string specifying the correlation structure. The following are permitted: "independence", "fixed", "stat_M_dep", "non_stat_M_dep", "exchangeable", "AR-M" and "unstructured"
formula	A formula related the response variable with the explanatory variables. If it is NULL, formula `response~period+treatment+carry+time+covar` will be evaluated.
tol	the tolerance used in the fitting algorithm.
niter	the maximum number of iterations.
nodes	Number of nodes in the estimation of the splines. By default, the base 2 logarithm of the number of observations per period is used.
Mv	When correlation is "stat_M_dep", "non_stat_M_dep", or "AR-M" then Mv must be specified.

Value

`QIC` The QIC of the model: The model are fitted by `geeglm`

`model` The model fitted by `geeglm`.

`graphs` The graphs estimated by splines. In position 1 the graph of the effect of time appears and from then on, it appears one for each carryover effect declared in the `carry` option. The graphs are built with `ggplot2`, therefore they allow manipulation of axes and other graphic parameters of that library.

Source

https://doi.org/10.1177/09622802231158736

References

Cruz Gutierrez NA, Melo OO, Martinez CA. Semiparametric generalized estimating equations for repeated measurements in cross-over designs. Statistical Methods in Medical Research, 2023;32(5):1033-1050.

Examples

```
data(Arterial)

carrydata <- createCarry(data=Arterial, treatment = "Treatment",
                          period = "Period",id="Subject", carrySimple = FALSE)
data <- carrydata$data
carry <- carrydata$carryover
model1 <- CrossGEESP(response = "Pressure", treatment = "Treatment",
                      period = "Period", id="Subject", time="Time",
                      carry=carrydata$carryover,data=data,
                      correlation = "exchangeable")


model2 <- CrossGEESP(response = "Pressure", treatment = "Treatment",
                       period = "Period", id="Subject", time="Time",
                       carry=carrydata$carryover,data=data, correlation = "AR-M")


model1$QIC
model2$QIC
summary(model1$model)
summary(model2$model)

## Aproximate p-values for model 2

(pvalues <- 2 * pnorm(abs(coef(summary(model2$model))[,5]), lower.tail = FALSE))

model1$graph[[1]]
model1$graph[[2]]
plot <- model1$graph[[1]] + ggplot2::xlab("Time in minutes")+
ggplot2::ylab("Change in systolic blood pressure")
plot
```

Water   *Water student data of crossover design*

Description

A pilot study to investigate the impact of providing supplementary drinking water on the cognitive performance of pupils of two school grades (5 and 6) in water-scarce schools in rural Mali. 47 students were assigned to take the control treatment (normal conditions) on the first day and receive the treatment (controlled hydration) on the second day. 60 received the treatments in reverse (Hydration the first day and control the second day). One part of this test assesses visual attention. This test assesses visual attention. Pupils were given a grid containing target letters randomly dispersed among non-target letters and were given one minute to cross out as many target letters as possible. Scores were calculated by subtracting the number of non-target letters identified from the number of target letters identified; the maximum test score was 38.

Usage

```
Water
```

Format

A data frame with 214 rows and 10 columns:

ID The unique identifier of each of the students

Age The age in years of each of the students

LCS Letter Cancel incorrect (raw score)

LCC Letter Cancel correct (raw score)

LCI Letter Cancel score (LCC-LCI)

sex f=female, 1=male

school school indicator A or B

Treatment Condition indicator 0=Control 1=Water

Period date of visit

Carry_Agua Carry indicator 0=First period, 1=Water in the first period, -1 = Control in the first period

Source

<https://journals.plos.org/plosone/article/authors?id=10.1371/journal.pone.0210568>

Index