

# Package ‘geesmv’

October 12, 2015

**Type** Package

**Title** Modified Variance Estimators for Generalized Estimating Equations

**Version** 1.3

**Date** 2015-10-12

**Author** Ming Wang <mwang@phs.psu.edu>

**Maintainer** Zheng Li <zheng.li@outlook.com>

**Depends** R (>= 3.0.2)

**Imports** stats, nlme, gee, matrixcalc, MASS

**Description** Generalized estimating equations with the original sandwich variance estimator proposed by Liang and Zeger (1986), and eight types of more recent modified variance estimators for improving the finite small-sample performance.

**License** GPL (>= 3)

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2015-10-12 20:41:42

## R topics documented:

geesmv-package . . . . .	2
cluster.size . . . . .	4
cormax.ar1 . . . . .	5
cormax.exch . . . . .	5
cormax.ind . . . . .	5
dental . . . . .	5
GEE.var.fg . . . . .	6
GEE.var.gst . . . . .	7
GEE.var.kc . . . . .	9
GEE.var.lz . . . . .	10
GEE.var.mbn . . . . .	12
GEE.var.md . . . . .	13
GEE.var.mk . . . . .	15

GEE.var.pan . . . . .	16
GEE.var.wl . . . . .	18
mat.prod . . . . .	19
mat.sqrt . . . . .	19
mat.sqrt.inv . . . . .	19
seizure . . . . .	20
toenail . . . . .	20

<b>Index</b>	<b>22</b>
--------------	-----------

---

geesmv-package	<i>Modified Variance Estimators for Generalized Estimating Equations</i>
----------------	--

---

## Description

Generalized estimating equations with the original sandwich variance estimator proposed by Liang and Zeger (1986), and eight types of more recent modified variance estimators for improving the finite small-sample performance.

## Details

Generalized estimating equations with the original sandwich variance estimator proposed by Liang and Zeger (1986), and eight types of more recent modified variance estimators for improving the finite small-sample performance.

GEE.var.pan(), GEE.var.gst() and GEE.var.wl() are only for the balanced data, while the others can be used for both balanced and unbalanced data.

## Author(s)

Ming Wang <mwang@phs.psu.edu>

Maintainer: Zheng Li <zheng.li@outlook.com>

## References

De Backer M, De Vroey C, Lesaffre E, Scheys I, De Keyser P. Twelve weeks of continuous oral therapy for toenail onychomycosis caused by dermatophytes: a double-blind comparative trial of terbinafine 250 mg/day versus itraconazole 200 mg/day. *Journal of the American Academy of Dermatology* 1998; 38: 57-63.

Fay MP and Graubard BI. Small-sample adjustments for Wald-type tests using sandwich estimators. *Biometrics* 2001;57: 1198-1206.

Gosho M, Sato Y and Takeuchi H. Robust covariance estimator for small-sample adjustment in the generalized estimating equations: A simulation study. *Science Journal of Applied Mathematics and Statistics* 2014;2(1):20-25.

Kauermann G and Carroll RJ. A note on the efficiency of sandwich covariance matrix estimation. *Journal of the American Statistical Association* 2001;96: 1387-1398.

Morel JG, Bokossa MC, and Neerchal NK. Small sample correction for the variance of GEE estimators. *Biometrical Journal* 2003;45(4): 395-409.

- MacKinnon JG. Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics* 1985;29: 305-325.
- Mancini LA and DeRouen TA. A covariance estimator for GEE with improved small-sample properties. *Biometrics* 2001;57: 126-134.
- Pan W. On the robust variance estimator in Generalized Estimating Equations. *Biometrika* 2001;88: 901-906.
- Pottho R. F. and Roy, S. W. A generalized multivariate analysis of variance model useful especially for growth curve problems. *Biometrika* 1964;51:313-326
- Thall PF, and Vail SC. Some covariance models for longitudinal count data with overdispersion. *Biometrics* 1990; 46: 657-671.
- Wang M and Long Q. Modified robust variance estimator for generalized estimating equations with improved small-sample performance. *Statistics in Medicine* 2011;30(11): 1278-1291.
- Zeger SL and Liang KY. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* 1986; 121-130.

### See Also

GEE.var.lz()  
GEE.var.mk()  
GEE.var.kc()  
GEE.var.pan()  
GEE.var.fg()  
GEE.var.gst()  
GEE.var.md()  
GEE.var.mbn()  
GEE.var.wl()

### Examples

```
### Run the data analysis using the package using seizure dataset (Poisson);
data(seizure)
seizure$subject <- 1:length(seizure[,1])
data_alt <- reshape(seizure, direction="long", idvar="subject", timevar="Time",
                    varying=names(seizure)[1:4], v.names="response", times=1:4)
data_alt <- data_alt[order(data_alt$subject),]
data_alt <- data_alt[,c(4,1:3,5,6)]

### independence working correlation structure;
formula <- response~base+trt+Time
lz.ind <- GEE.var.lz(formula,id="subject",family=poisson,
                    data_alt,corstr="independence")

mk.ind <- GEE.var.mk(formula,id="subject",family=poisson,
                    data_alt,corstr="independence")
```

```
pan.ind <- GEE.var.pan(formula,id="subject",family=poisson,
  data_alt,corstr="independence")

gst.ind <- GEE.var.gst(formula,id="subject",family=poisson,
  data_alt,corstr="independence")

kc.ind <- GEE.var.kc(formula,id="subject",family=poisson,
  data_alt,corstr="independence")

md.ind <- GEE.var.md(formula,id="subject",family=poisson,
  data_alt,corstr="independence")

fg.ind <- GEE.var.fg(formula,id="subject",family=poisson,data_alt,
  corstr="independence",b=0.75)
mbn.ind <- GEE.var.mbn(formula,id="subject",family=poisson,data_alt,
  corstr="independence",d=2,r=1)
wl.ind <- GEE.var.wl(formula,id="subject",family=poisson,
  data_alt,corstr="independence")
```

---

cluster.size

*Get Summary Information of The Data.*

---

### **Description**

Get necessary information (i.e., the number of clusters, cluster sizes) of the data set.

### **Usage**

```
cluster.size(id)
```

### **Arguments**

id                    The variable name of the cluster id in the data set.

### **Details**

This function will get the number of clusters and cluster sizes of the data set.

### **Value**

n                    a vector of numbers denote the cluster sizes for different id.  
m                    the total number of clusters in the data set.

### **Author(s)**

Ming Wang <mwang@phs.psu.edu> Maintainer: Zheng Li <zheng.li@outlook.com>

---

cormax.ar1	<i>Calculate variance estimator</i>
------------	-------------------------------------

---

**Description**

Being used to calculate variance estimator by the functions in this package. It cannot be directly used in other packages.

---

cormax.exch	<i>Calculate variance estimator</i>
-------------	-------------------------------------

---

**Description**

Being used to calculate variance estimator by the functions in this package. It cannot be directly used in other packages.

---

cormax.ind	<i>Calculate variance estimator</i>
------------	-------------------------------------

---

**Description**

Being used to calculate variance estimator by the functions in this package. It cannot be directly used in other packages.

---

dental	<i>A Data Set of Orthodontic Measurements on Children</i>
--------	---

---

**Description**

The dental data consists of 27 subjects from the University of North Carolina Dental School.

**Usage**

dental

**Format**

- subject:cluster id
- gender:the sex of subject
- age\_8:the measurements at age 8
- age\_10:the measurements at age 10
- age\_12:the measurements at age 12
- age\_14:the measurements at age 14

**Details**

The data was collected by investigators at the University of North Carolina Dental School. It is a study of orthodontic measurements on children including 11 girls and 16 boys measured at ages 8, 10, 12, and 14.

**Source**

Pottho R. F. and Roy, S. W. A generalized multivariate analysis of variance model useful especially for growth curve problems. *Biometrika* 1964;51:313-326

---

GEE.var.fg	<i>A Modified GEE Variance Estimator Proposed by Fay and Graubard (2001)</i>
------------	--

---

**Description**

Calculate the modified GEE variance estimator proposed by Fay and Graubard (2001).

**Usage**

```
GEE.var.fg(formula, id, family=gaussian, data, corstr="independence", b=0.75)
```

**Arguments**

formula	an object of class "formula" (or one that can be coerced to that class): a formula expression as for other regression models to be fitted, of the form response ~ predictors. The details of formula specification can be seen in <code>glm()</code> and <code>gee()</code> .
id	a vector which identifies the clusters. The length of <code>id</code> should be the same as the total number of observations. Data is assumed to be sorted so that observations on a cluster are contiguous rows for all entities in the formula.
family	a family object: a list of functions and expressions for defining link and variance functions to be used in the model. This can be a character string naming a family function including "gaussian", "binomial", and "poisson". See family for details of family functions in <code>glm()</code> and <code>gee()</code> .
data	an optional data frame, list or environment (or object coercible by <code>as.data.frame</code> to a data frame) containing the variables in the model. If not found in data, the variables are taken from <code>environment(formula)</code> , typically the environment from which <code>glm</code> is called.
corstr	a character string specifying working correlation structure: "independence", "AR-M", "exchangeable", "unstructured" are possible.
b	adjustment factor in the proposed variance estimator, which is prespecified subjectively to avoid extreme adjustments when the cluster size is close to 1. The default value is 0.75.

**Details**

Fit the model by GEE and calculate the variance estimator proposed by Fay and Graubard (2001).

**Value**

cov.beta	estimate of the variance for beta
cov.var	estimate of the variance-covariance matrix for the variance estimator

**Author(s)**

Ming Wang <mwang@phs.psu.edu>

Maintainer: Zheng Li <zheng.li@outlook.com>

**References**

Fay MP and Graubard BI. Small-sample adjustments for Wald-type tests using sandwich estimators. *Biometrics* 2001;57: 1198-1206.

**Examples**

```
data(dental)
data_alt <- reshape(dental, direction="long", timevar="Time",
  varying=names(dental)[3:6], v.names="response", times=c(8,10,12,14))
data_alt <- data_alt[order(data_alt$subject),]
data_alt$gender <- as.numeric(data_alt$gender)
data_alt$Time <- sqrt(data_alt$Time)
formula <- response~Time+gender
fg.ind <- GEE.var.fg(formula,id="subject",family=gaussian,
  data_alt,corstr="independence") ##Independence correlation structure;
fg.exch <- GEE.var.fg(formula,id="subject",family=gaussian,
  data_alt,corstr="exchangeable") ##Exchangeable correlation structure;
fg.ar1 <- GEE.var.fg(formula,id="subject",family=gaussian,
  data_alt,corstr="AR-M") ##AR-1 correlation structure;
fg.unstr <- GEE.var.fg(formula,id="subject",family=gaussian,
  data_alt,corstr="unstructured") ##Unstructured correlation structure;
```

---

GEE.var.gst

*A modified GEE Variance Estimator Proposed by Gosho et al.(2014)*

---

**Description**

Calculate the modified GEE variance estimator proposed by Gosho et al. (2014).

**Usage**

```
GEE.var.gst(formula,id,family=gaussian,data,corstr="independence")
```

**Arguments**

formula	an object of class "formula" (or one that can be coerced to that class): a formula expression as for other regression models to be fitted, of the form response ~ predictors. The details of formula specification can be seen in <code>glm()</code> and <code>gee()</code> .
id	a vector which identifies the clusters. The length of id should be the same as the total number of observations. Data is assumed to be sorted so that observations on a cluster are contiguous rows for all entities in the formula.
family	a family object: a list of functions and expressions for defining link and variance functions to be used in the model. This can be a character string naming a family function including "gaussian", "binomial", and "poisson". See family for details of family functions in <code>glm()</code> and <code>gee()</code> .
data	an optional data frame, list or environment (or object coercible by <code>as.data.frame</code> to a data frame) containing the variables in the model. If not found in data, the variables are taken from <code>environment(formula)</code> .
corstr	a character string specifying working correlation structure: "independence", "AR-M", "exchangeable", "unstructured" are possible.

**Details**

Fit the model by GEE and provide the variance estimator based on Gosho et al.(2014). Gosho et al (2014) proposed to made an additional modification on Pan's estimator.

**Value**

cov.beta	estimate of the variance for beta
cov.var	estimate of the variance-covariance matrix for the variance estimator

**Author(s)**

Ming Wang <mwang@phs.psu.edu>  
 Maintainer: Zheng Li <zheng.li@outlook.com>

**References**

Gosho, M., Sato, Y. and Takeuchi, H. Robust covariance estimator for small-sample adjustment in the generalized estimating equations: A simulation study. Science Journal of Applied Mathematics and Statistics 2014;2(1):20-25.

**Examples**

```
data(dental)
data_alt <- reshape(dental, direction="long", timevar="Time",
  varying=names(dental)[3:6], v.names="response", times=c(8,10,12,14))
data_alt <- data_alt[order(data_alt$subject),]
data_alt$gender <- as.numeric(data_alt$gender)
data_alt$Time <- sqrt(data_alt$Time)
formula <- response~Time+gender
gst.ind <- GEE.var.gst(formula,id="subject",family=gaussian,
```



```

      data_alt,corstr="independence") ##Independence correlation structure;
gst.exch <- GEE.var.gst(formula,id="subject",family=gaussian,
      data_alt,corstr="exchangeable") ##Exchangeable correlation structure;
gst.ar1 <- GEE.var.gst(formula,id="subject",family=gaussian,
      data_alt,corstr="AR-M") ##AR-1 correlation structure;
gst.unstr <- GEE.var.gst(formula,id="subject",family=gaussian,
      data_alt,corstr="unstructured") ##Unstructured correlation structure;

```

---

GEE.var.kc	<i>A Modified GEE Variance Estimator Proposed by Kauermann and Carroll (2001).</i>
------------	--

---

### Description

Calculate the modified GEE variance estimator proposed by Kauermann and Carroll (2001).

### Usage

```
GEE.var.kc(formula,id,family=gaussian,data,corstr="independence")
```

### Arguments

formula	an object of class "formula" (or one that can be coerced to that class): a formula expression as for other regression models to be fitted, of the form response ~ predictors. The details of formula specification can be seen in glm() and gee().
id	a vector which identifies the clusters. The length of id should be the same as the total number of observations. Data is assumed to be sorted so that observations on a cluster are contiguous rows for all entities in the formula.
family	a family object: a list of functions and expressions for defining link and variance functions to be used in the model. This can be a character string naming a family function including "gaussian", "binomial", and "poisson". See family for details of family functions in glm() and gee().
data	an optional data frame, list or environment (or object coercible by as.data.frame to a data frame) containing the variables in the model. If not found in data, the variables are taken from environment(formula), typically the environment from which glm is called.
corstr	a character string specifying working correlation structure: "independence", "AR-M","exchangeable", "unstructured" are possible.

### Details

Fit the model by GEE and calculate the bias-corrected sandwich variance estimator proposed by Kauermann and Carroll(2001).

### Value

cov.beta	estimate of the variance for beta
cov.var	estimate of the variance-covariance matrix for the variance estimator

**Author(s)**

Ming Wang <mwang@phs.psu.edu>

Maintainer: Zheng Li <zheng.li@outlook.com>

**References**

Kauermann G and Carroll RJ. A note on the efficiency of sandwich covariance matrix estimation. *Journal of the American Statistical Association* 2001;96: 1387-1398.

**Examples**

```
data(dental)
data_alt <- reshape(dental, direction="long", timevar="Time",
  varying=names(dental)[3:6], v.names="response", times=c(8,10,12,14))
data_alt <- data_alt[order(data_alt$subject),]
data_alt$gender <- as.numeric(data_alt$gender)
data_alt$Time <- sqrt(data_alt$Time)
formula <- response~Time+gender
kc.ind <- GEE.var.kc(formula,id="subject",family=gaussian,
  data_alt,corstr="independence") ##Independence correlation structure;
kc.exch <- GEE.var.kc(formula,id="subject",family=gaussian,
  data_alt,corstr="exchangeable") ##Exchangeable correlation structure;
kc.ar1 <- GEE.var.kc(formula,id="subject",family=gaussian,
  data_alt,corstr="AR-M") ##AR-1 correlation structure;
kc.unstr <- GEE.var.kc(formula,id="subject",family=gaussian,
  data_alt,corstr="unstructured") ##Unstructured correlation structure;
```

---

GEE.var.lz

*GEE Sandwich Variance Estimator Proposed by Liang and Zeger (1986).*

---

**Description**

Calculate the GEE original variance estimator proposed by Liang and Zeger (1986).

**Usage**

```
GEE.var.lz(formula,id="subject",family=gaussian,data,corstr="independence")
```

**Arguments**

formula	an object of class "formula" (or one that can be coerced to that class): a formula expression as for other regression models to be fitted, of the form response ~ predictors. The details of formula specification can be seen in glm() and gee().
id	a vector which identifies the clusters. The length of id should be the same as the total number of observations. Data is assumed to be sorted so that observations on a cluster are contiguous rows for all entities in the formula.

family	a family object: a list of functions and expressions for defining link and variance functions to be used in the model. This can be a character string naming a family function including "gaussian", "binomial", and "poisson". See family for details of family functions in glm() and gee().
data	an optional data frame, list or environment (or object coercible by as.data.frame to a data frame) containing the variables in the model. If not found in data, the variables are taken from environment(formula), typically the environment from which glm is called.
corstr	a character string specifying working correlation structure: "independence", "AR-M", "exchangeable", "unstructured" are possible.

### Details

Fit the model by GEE and calculate the original variance estimator proposed by Liang and Zeger (1986).

### Value

cov.beta	estimate of the variance for beta
cov.var	estimate of the variance-covariance matrix for the variance estimator

### Author(s)

Ming Wang <mwang@phs.psu.edu>

Maintainer: Zheng Li <zheng.li@outlook.com>

### References

Zeger SL and Liang KY. Longitudinal data analysis for discrete and continuous outcomes. Biometrics 1986: 121-130.

### Examples

```
data(dental)
data_alt <- reshape(dental, direction="long", timevar="Time",
  varying=names(dental)[3:6], v.names="response", times=c(8,10,12,14))
data_alt <- data_alt[order(data_alt$subject),]
data_alt$gender <- as.numeric(data_alt$gender)
data_alt$Time <- sqrt(data_alt$Time)
formula <- response~Time+gender
lz.ind <- GEE.var.lz(formula,id="subject",family=gaussian,
  data_alt,corstr="independence") ##Independence correlation structure;
lz.exch <- GEE.var.lz(formula,id="subject",family=gaussian,
  data_alt,corstr="exchangeable") ##Exchangeable correlation structure;
lz.ar1 <- GEE.var.lz(formula,id="subject",family=gaussian,
  data_alt,corstr="AR-M") ##AR-1 correlation structure;
lz.unstr <- GEE.var.lz(formula,id="subject",family=gaussian,
  data_alt,corstr="unstructured") ##Unstructured correlation structure;
```

---

GEE.var.mbn	<i>A modified GEE Variance Estimator Proposed by Morel, Bokossa and Neerchal (2003).</i>
-------------	--

---

### Description

Calculate the modified GEE variance estimator proposed by Morel, Bokossa and Neerchal (2003).

### Usage

```
GEE.var.mbn(formula, id, family=gaussian, data, corstr="independence", d=2, r=1)
```

### Arguments

formula	an object of class "formula" (or one that can be coerced to that class): a formula expression as for other regression models to be fitted, of the form response ~ predictors. The details of formula specification can be seen in <code>glm()</code> and <code>gee()</code> .
id	a vector which identifies the clusters. The length of <code>id</code> should be the same as the total number of observations. Data is assumed to be sorted so that observations on a cluster are contiguous rows for all entities in the formula.
family	a family object: a list of functions and expressions for defining link and variance functions to be used in the model. This can be a character string naming a family function including "gaussian", "binomial", and "poisson". See family for details of family functions in <code>glm()</code> and <code>gee()</code> .
data	an optional data frame, list or environment (or object coercible by <code>as.data.frame</code> to a data frame) containing the variables in the model. If not found in data, the variables are taken from <code>environment(formula)</code> , typically the environment from which <code>glm</code> is called.
corstr	a character string specifying working correlation structure: "independence", "AR-M", "exchangeable", "unstructured" are possible.
d	adjustment factor in the variance estimator. The default value is 2. Details see Morel, Bokossa and Neerchal (2003).
r	adjustment factor in the variance estimator. The default value is 1. Details see Morel, Bokossa and Neerchal (2003).

### Details

Fit the model by GEE and calculate the variance estimator proposed by Morel, Bokossa and Neerchal (2003).

### Value

cov.beta	estimate of the variance for beta
cov.var	estimate of the variance-covariance matrix for the variance estimator

**Author(s)**

Ming Wang <mwang@phs.psu.edu>

Maintainer: Zheng Li <zheng.li@outlook.com>

**References**

Morel JG, Bokossa MC, and Neerchal NK. Small sample correction for the variance of GEE estimators. *Biometrical Journal* 2003;45(4): 395-409.

**Examples**

```
data(dental)
data_alt <- reshape(dental, direction="long", timevar="Time",
  varying=names(dental)[3:6], v.names="response", times=c(8,10,12,14))
data_alt <- data_alt[order(data_alt$subject),]
data_alt$gender <- as.numeric(data_alt$gender)
data_alt$Time <- sqrt(data_alt$Time)
formula <- response~Time+gender
mbn.ind <- GEE.var.mbn(formula,id="subject",family=gaussian,
  data_alt,corstr="independence",d=2,r=1) ##Independence correlation structure;
mbn.exch <- GEE.var.mbn(formula,id="subject",family=gaussian,
  data_alt,corstr="exchangeable",d=2,r=1) ##Exchangeable correlation structure;
mbn.ar1 <- GEE.var.mbn(formula,id="subject",family=gaussian,
  data_alt,corstr="AR-M",d=2,r=1) ##AR-1 correlation structure;
mbn.unstr <- GEE.var.mbn(formula,id="subject",family=gaussian,
  data_alt,corstr="unstructured",d=2,r=1) ##Unstructured correlation structur;
```

---

GEE.var.md

*A Modified GEE Variance Estimator Proposed by Mancl and DeRouen (2001).*

---

**Description**

Calculate the modified GEE variance estimator proposed by Mancl and DeRouen (2001).

**Usage**

```
GEE.var.md(formula,id,family=gaussian,data,corstr="independence")
```

**Arguments**

formula	an object of class "formula" (or one that can be coerced to that class): a formula expression as for other regression models to be fitted, of the form response ~ predictors. The details of formula specification can be seen in glm() and gee().
id	a vector which identifies the clusters. The length of id should be the same as the total number of observations. Data is assumed to be sorted so that observations on a cluster are contiguous rows for all entities in the formula.

family	a family object: a list of functions and expressions for defining link and variance functions to be used in the model. This can be a character string naming a family function including "gaussian", "binomial", and "poisson". See family for details of family functions in glm() and gee().
data	an optional data frame, list or environment (or object coercible by as.data.frame to a data frame) containing the variables in the model. If not found in data, the variables are taken from environment(formula), typically the environment from which glm is called.
corstr	a character string specifying working correlation structure: "independence", "AR-M", "exchangeable", "unstructured" are possible.

### Details

Fit the model by GEE and calculate the variance estimator proposed by Mancl and DeRouen (2001).

### Value

cov.beta	estimate of the variance for beta
cov.var	estimate of the variance-covariance matrix for the variance estimator

### Author(s)

Ming Wang <mwang@phs.psu.edu>

Maintainer: Zheng Li <zheng.li@outlook.com>

### References

Mancl LA and DeRouen TA. A Covariance Estimator for GEE with Improved Small- Sample Properties. Biometrics 2001;57: 126-134.

### Examples

```
data(dental)
data_alt <- reshape(dental, direction="long", timevar="Time",
  varying=names(dental)[3:6], v.names="response", times=c(8,10,12,14))
data_alt <- data_alt[order(data_alt$subject),]
data_alt$gender <- as.numeric(data_alt$gender)
data_alt$Time <- sqrt(data_alt$Time)
formula <- response~Time+gender
md.ind <- GEE.var.md(formula,id="subject",family=gaussian,
  data_alt,corstr="independence") ##Independence correlation structure;
md.exch <- GEE.var.md(formula,id="subject",family=gaussian,
  data_alt,corstr="exchangeable") ##Exchangeable correlation structure;
md.ar1 <- GEE.var.md(formula,id="subject",family=gaussian,
  data_alt,corstr="AR-M") ##AR-1 correlation structure;
md.unstr <- GEE.var.md(formula,id="subject",family=gaussian,
  data_alt,corstr="unstructured") ##Unstructured correlation structure;
```

---

GEE.var.mk

*A Modified GEE Variance Estimator Proposed by Mackinnon (1985).*


---

**Description**

Calculate the variance estimators proposed by Mackinnon (1985).

**Usage**

```
GEE.var.mk(formula, id, family=gaussian, data, corstr="independence")
```

**Arguments**

formula	an object of class "formula" (or one that can be coerced to that class): a formula expression as for other regression models to be fitted, of the form response ~ predictors. The details of formula specification can be seen in <code>glm()</code> and <code>gee()</code> .
id	a vector which identifies the clusters. The length of id should be the same as the total number of observations. Data is assumed to be sorted so that observations on a cluster are contiguous rows for all entities in the formula.
family	a family object: a list of functions and expressions for defining link and variance functions to be used in the model. This can be a character string naming a family function including "gaussian", "binomial", and "poisson". See family for details of family functions in <code>glm()</code> and <code>gee()</code> .
data	an optional data frame, list or environment (or object coercible by <code>as.data.frame</code> to a data frame) containing the variables in the model. If not found in data, the variables are taken from <code>environment(formula)</code> , typically the environment from which <code>glm</code> is called.
corstr	a character string specifying working correlation structure: "independence", "AR-M", "exchangeable", "unstructured" are possible.

**Details**

Fit the model by GEE and calculate the variance estimator proposed by Mackinnon (1985). This estimator improved the sandwich variance estimator by adjusting the degree of freedom.

**Value**

cov.beta	estimate of the variance for beta
cov.var	estimate of the variance-covariance matrix for the variance estimator

**Author(s)**

Ming Wang <mwang@phs.psu.edu>

Maintainer: Zheng Li <zheng.li@outlook.com>

## References

MacKinnon JG. Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics* 1985;29: 305-325.

## Examples

```
data(dental)
data_alt <- reshape(dental, direction="long", timevar="Time",
                    varying=names(dental)[3:6], v.names="response", times=c(8,10,12,14))
data_alt <- data_alt[order(data_alt$subject),]
data_alt$gender <- as.numeric(data_alt$gender)
data_alt$Time <- sqrt(data_alt$Time)
formula <- response~Time+gender
mk.ind <- GEE.var.mk(formula,id="subject",family=gaussian,
                    data_alt,corstr="independence") ##Independence correlation structure;
mk.exch <- GEE.var.mk(formula,id="subject",family=gaussian,
                     data_alt,corstr="exchangeable") ##Exchangeable correlation structure;
mk.ar1 <- GEE.var.mk(formula,id="subject",family=gaussian,
                    data_alt,corstr="AR-M") ##AR-1 correlation structure;
mk.unstr <- GEE.var.mk(formula,id="subject",family=gaussian,
                      data_alt,corstr="unstructured") ##Unstructured correlation structure;
```

---

GEE.var.pan

*A Modified GEE Variance Estimator Proposed by Pan (2001).*

---

## Description

Calculate the modified GEE variance estimator proposed by Pan (2001).

## Usage

```
GEE.var.pan(formula,id,family=gaussian,data,corstr="independence")
```

## Arguments

formula	an object of class "formula" (or one that can be coerced to that class): a formula expression as for other regression models to be fitted, of the form response ~ predictors. The details of formula specification can be seen in <code>glm()</code> and <code>gee()</code> .
id	a vector which identifies the clusters. The length of id should be the same as the total number of observations. Data is assumed to be sorted so that observations on a cluster are contiguous rows for all entities in the formula.
family	a family object: a list of functions and expressions for defining link and variance functions to be used in the model. This can be a character string naming a family function including "gaussian", "binomial", and "poisson". See family for details of family functions in <code>glm()</code> and <code>gee()</code> .



data	an optional data frame, list or environment (or object coercible by <code>as.data.frame</code> to a data frame) containing the variables in the model. If not found in data, the variables are taken from <code>environment(formula)</code> , typically the environment from which <code>glm</code> is called.
corstr	a character string specifying working correlation structure: "independence", "AR-M", "exchangeable", "unstructured" are possible.

### Details

it the model by GEE and calculate the variance estimator proposed by Pan(2001). There are two assumptions for the model: 1) The conditional variance is correctly specified; 2) A common correlation structure exists across all subjects.

### Value

cov.beta	estimate of the variance for beta
cov.var	estimate of the variance-covariance matrix for the variance estimator

### Author(s)

Ming Wang <mwang@phs.psu.edu>

Maintainer: Zheng Li <zheng.li@outlook.com>

### References

Pan W. On the Robust Variance Estimator in Generalized Estimating Equations. *Biometrika* 2001;88: 901-906.

### Examples

```
data(dental)
data_alt <- reshape(dental, direction="long", timevar="Time",
  varying=names(dental)[3:6], v.names="response", times=c(8,10,12,14))
data_alt <- data_alt[order(data_alt$subject),]
data_alt$gender <- as.numeric(data_alt$gender)
data_alt$Time <- sqrt(data_alt$Time)
formula <- response~Time+gender
pan.ind <- GEE.var.pan(formula,id="subject",family=gaussian,
  data_alt,corstr="independence") ##Independence correlation structure;
pan.exch <- GEE.var.pan(formula,id="subject",family=gaussian,
  data_alt,corstr="exchangeable") ##Exchangeable correlation structure;
pan.ar1 <- GEE.var.pan(formula,id="subject",family=gaussian,
  data_alt,corstr="AR-M") ##AR-1 correlation structure;
pan.unstr <- GEE.var.pan(formula,id="subject",family=gaussian,
  data_alt,corstr="unstructured") ##Unstructured correlation structure;
```

---

GEE.var.wl                      *A Modified GEE Variance Estimator Proposed by Wang and Long (2011).*

---

### Description

Calculate the modified GEE variance estimator proposed by Wang and Long (2011).

### Usage

```
GEE.var.wl(formula, id, family=gaussian, data, corstr="independence")
```

### Arguments

formula	an object of class "formula" (or one that can be coerced to that class): a formula expression as for other regression models to be fitted, of the form response ~ predictors. The details of formula specification can be seen in <code>glm()</code> and <code>gee()</code> .
id	a vector which identifies the clusters. The length of id should be the same as the total number of observations. Data is assumed to be sorted so that observations on a cluster are contiguous rows for all entities in the formula.
family	a family object: a list of functions and expressions for defining link and variance functions to be used in the model. This can be a character string naming a family function including "gaussian", "binomial", and "poisson". See family for details of family functions in <code>glm()</code> and <code>gee()</code> .
data	an optional data frame, list or environment (or object coercible by <code>as.data.frame</code> to a data frame) containing the variables in the model. If not found in data, the variables are taken from <code>environment(formula)</code> , typically the environment from which <code>glm</code> is called.
corstr	a character string specifying working correlation structure: "independence", "AR-M", "exchangeable", "unstructured" are possible.

### Details

Fit the model by GEE and calculate the variance estimator proposed by Wang and Long (2011). Two assumptions need to be satisfied: 1) The conditional variance is correctly specified; 2) A common correlation structure exists across all subjects.

### Value

cov.beta	estimate of the variance for beta
cov.var	estimate of the variance-covariance matrix for the variance estimator

### Author(s)

Ming Wang <mwang@phs.psu.edu>  
 Maintainer: Zheng Li <zheng.li@outlook.com>

## References

Wang M and Long Q. Modified robust variance estimator for generalized estimating equations with improved small-sample performance. *Statistics in Medicine* 2011;30(11): 1278-1291.

## Examples

```
data(dental)
data_alt <- reshape(dental, direction="long", timevar="Time",
                    varying=names(dental)[3:6], v.names="response", times=c(8,10,12,14))
data_alt <- data_alt[order(data_alt$subject),]
data_alt$gender <- as.numeric(data_alt$gender)
data_alt$Time <- sqrt(data_alt$Time)
formula <- response~Time+gender
wl.ind <- GEE.var.wl(formula,id="subject",family=gaussian,
                    data_alt,corstr="independence") ##Independence correlation structure;
wl.exch <- GEE.var.wl(formula,id="subject",family=gaussian,
                    data_alt,corstr="exchangeable") ##Exchangeable correlation structure;
wl.ar1 <- GEE.var.wl(formula,id="subject",family=gaussian,
                    data_alt,corstr="AR-M") ##AR-1 correlation structure;
wl.unstr <- GEE.var.wl(formula,id="subject",family=gaussian,
                    data_alt,corstr="unstructured") ##Unstructured correlation structure;
```

---

mat.prod

*Calculate variance estimator*

---

## Description

Being used to calculate variance estimator by the functions in this package. It cannot be directly used in other packages.

---

mat.sqrt

*Calculate variance estimator*

---

## Description

Being used to calculate variance estimator by the functions in this package. It cannot be directly used in other packages.

---

mat.sqrt.inv

*Calculate variance estimator*

---

## Description

Being used to calculate variance estimator by the functions in this package. It cannot be directly used in other packages.

---

seizure

*Epileptic seizure counts from the Randomized Progabide Trial*

---

**Description**

The data consists of 59 subjects from the randomized progabide trial.

**Usage**

seizure

**Format**

- base:baseline epileptic seizure counts
- trt:the indicator for treatment (1=progabide; 0=control)
- age:baseline age for each subject
- y1:epileptic seizure counts in the first week
- y2:epileptic seizure counts in the second week
- y3:epileptic seizure counts in the third week
- y4:epileptic seizure counts in the forth week

**Details**

The outcome is the count number of epileptic seizures in each of four consecutive two-week intervals, and the variables recorded include baseline epileptic seizure counts in an eight-week interval prior to the treatment assignment, the indicator for treatment (1=progabide; 0=control), and baseline age.

**Source**

Thall PF, and Vail SC. Some covariance models for longitudinal count data with overdispersion. *Biometrics* 1990; 46: 657-671.

---

toenail

*Toenail infection data from a multicenter study*

---

**Description**

The data frame consists of 250 patients from a multicenter study comparing to treatments for toenail infection.

**Usage**

toenail

**Format**

- ID: patient id. The variable for cluster id
- Response: toenail infection status: 0=none or mild; 1=moderate or severe
- Treatment: 1=oral treatment A; 0=oral treatment B
- Month: the exact time point of measurement in month
- Visit: the visit index

**Details**

In a randomized, double-blinded multicenter study, treatment A and B were compared to treat toenail infection disease. Patients were evaluated for the degree of onycholysis. There are 1908 measurements in total.

**Source**

De Backer M, De Vroey C, Lesaffre E, Scheys I, De Keyser P. Twelve weeks of continuous oral therapy for toenail onychomycosis caused by dermatophytes: a double-blind comparative trial of terbinafine 250 mg/day versus itraconazole 200 mg/day. *Journal of the American Academy of Dermatology* 1998; 38: 57-63.

# Index

- \*Topic **Fay**
    - GEE.var.fg, 6
  - \*Topic **GEE**
    - cluster.size, 4
    - dental, 5
    - GEE.var.gst, 7
    - GEE.var.kc, 9
    - GEE.var.lz, 10
    - GEE.var.mbn, 12
    - GEE.var.md, 13
    - GEE.var.mk, 15
    - GEE.var.pan, 16
    - seizure, 20
    - toenail, 20
  - \*Topic **Gee**
    - GEE.var.w1, 18
  - \*Topic **Gosho**
    - GEE.var.gst, 7
  - \*Topic **Kauermann**
    - GEE.var.kc, 9
  - \*Topic **Liang**
    - GEE.var.lz, 10
  - \*Topic **Mackinnon**
    - GEE.var.mk, 15
  - \*Topic **Mancl**
    - GEE.var.md, 13
  - \*Topic **Morel**
    - GEE.var.mbn, 12
  - \*Topic **Pan**
    - GEE.var.pan, 16
  - \*Topic **cluster**
    - cluster.size, 4
    - dental, 5
  - \*Topic **efficiency**
    - GEE.var.w1, 18
  - \*Topic **efficient**
    - GEE.var.fg, 6
  - \*Topic **package**
    - geesmv-package, 2
  - \*Topic **seizure**
    - seizure, 20
  - \*Topic **toenail**
    - toenail, 20
- cluster.size, 4
  - cormax.ar1, 5
  - cormax.exch, 5
  - cormax.ind, 5
  - dental, 5
  - GEE.var.fg, 6
  - GEE.var.gst, 7
  - GEE.var.kc, 9
  - GEE.var.lz, 10
  - GEE.var.mbn, 12
  - GEE.var.md, 13
  - GEE.var.mk, 15
  - GEE.var.pan, 16
  - GEE.var.w1, 18
  - geesmv (geesmv-package), 2
  - geesmv-package, 2
  - mat.prod, 19
  - mat.sqrt, 19
  - mat.sqrt.inv, 19
  - seizure, 20
  - toenail, 20