# Package 'multipleOutcomes'

May 30, 2024

**Title** Asymptotic Covariance Matrix of Regression Models for Multiple
Outcomes

**Version** 0.4

**Description** Regression models can be fitted for multiple outcomes simultaneously. This package computes estimates of parameters across fitted models and returns the matrix of asymptotic covariance. Various applications of this package, including CUPED (Controlled Experiments Utilizing Pre-Experiment Data), multiple comparison adjustment, are illustrated.

**License** MIT + file LICENSE

**Encoding** UTF-8

**RoxygenNote** 7.3.1

**Imports** dplyr, momentfit, numDeriv, stringr, survival

**Suggests** asaur, coin, ggplot2, iBST, invGauss, JM, joint.Cox, knitr, mvtnorm, pec, randomForestSRC, rmarkdown, survminer, tidyr

**VignetteBuilder** knitr

**Depends** R (>= 2.10)

**LazyData** true

**NeedsCompilation** no

**Author** Han Zhang [aut, cre]

**Maintainer** Han Zhang <zhangh.ustc@gmail.com>

**Repository** CRAN

**Date/Publication** 2024-05-30 15:00:03 UTC

## R topics documented:

---

actg                              *ACTG 320 Clinical Trial Dataset*

---

### Description

actg dataset from Hosmer et al.

### Format

A data frame

**id** Identification Code

**time** Time to AIDS diagnosis or death (days).

**censor** Event indicator. 1 = AIDS defining diagnosis, 0 = Otherwise.

**time_d** Time to death (days)

**censor_d** Event indicator for death (only). 1 = Death, 0 = Otherwise.

**tx** Treatment indicator. 1 = Treatment includes IDV, 0 = Control group.

**txgrp** Treatment group indicator. 1 = ZDV + 3TC. 2 = ZDV + 3TC + IDV. 3 = d4T + 3TC. 4 = d4T + 3TC + IDV.

**strat2** CD4 stratum at screening. 0 = CD4 <= 50. 1 = CD4 > 50.

**sex** 0 = Male. 1 = Female.

**raceth** Race/Ethnicity. 1 = White Non-Hispanic. 2 = Black Non-Hispanic. 3 = Hispanic. 4 = Asian, Pacific Islander. 5 = American Indian, Alaskan Native. 6 = Other/unknown.

**ivdrug** IV drug use history. 1 = Never. 2 = Currently. 3 = Previously.

**hemophil** Hemophiliac. 1 = Yes. 0 = No.

**karnof** Karnofsky Performance Scale. 100 = Normal; no complaint no evidence of disease. 90 = Normal activity possible; minor signs/symptoms of disease. 80 = Normal activity with effort; some signs/symptoms of disease. 70 = Cares for self; normal activity/active work not possible.

**cd4** Baseline CD4 count (Cells/Milliliter).

**priorzdv** Months of prior ZDV use (months).

**age** Age at Enrollment (years).

### Source

ftp://ftp.wiley.com/public/sci_tech_med/survival

### References

Hosmer, D.W. and Lemeshow, S. and May, S. (2008) Applied Survival Analysis: Regression Modeling of Time to Event Data: Second Edition, John Wiley and Sons Inc., New York, NY

### Examples

data(actg)

---

coef.multipleOutcomes     *Extract Model Coefficients*

---

### Description

`coef` is a generic function.

### Usage

```
## S3 method for class 'multipleOutcomes'
coef(object, model_index = NULL, ...)
```

### Arguments

| | |
|---|---|
| `object` | an object returned by `multipleOutcomes()`. |
| `model_index` | NULL if displaying coefficients of all fitted models; otherwise, an integer indicating the fitted model. |
| `...` | for debugging only |

### Value

a vector of coefficient estimates

---

multipleOutcomes     *Fitting Regression Models for Multiple Outcomes and Returning the Matrix of Covariance*

---

### Description

`multipleOutcomes` can fit different types of models for multiple outcomes simultaneously and return model parameters and variance-covariance matrix for further analysis.

### Usage

```
multipleOutcomes(..., family, data, data_index = NULL, score_epsilon = 1e-06)
```

### Arguments

| | |
|---|---|
| `...` | formulas of models to be fitted, or moment functions for gmm. |
| `family` | a character vector of families to be used in the models. Currently only `gaussian`, `binomial`, `coxph` and `gmm` are supported. `long` for longitudinal data may be supported in the future. `family` can be of length 1 if all models are fitted in the same family; otherwise family should be specified for each of the models in `...`. |

| data | a data frame if all models are fitted on the same dataset; otherwise a list of data frames for fitting models in .... Note that a dataset can be used to fit multiple models, thus, length(data) is unnecessary to be equal to the number of models in .... The row names in a data frame are treated as sample IDs. Consequently, for any two records in different data frames that correspond to the same sample, their row names should be consistent. |
|---|---|
| data_index | NULL if data is a data frame; otherwise, a vector in integer specifying mapping a model in ... to a data frame in data (a list). |
| score_epsilon | whatever. |

## Value

It returns an object of class "multipleOutcomes", which is a list containing the following components:

| coefficients | an unnamed vector of coefficients of all fitted models. Use id_map for variable mapping. |
|---|---|
| mcov | a unnamed matrix of covariance of coefficients. Use id_map for variable mapping. |
| id_map | a list mapping the elements in coefficients and mcov to variable names. |
| n_shared_sample_sizes | a matrix of shared sample sizes between datasets being used to fit the models. |
| call | the matched call. |

## Examples

```
## More examples can be found in the vignettes.
library(mvtnorm)
genData <- function(seed = NULL){

  set.seed(seed)
  n <- 400
  sigma <- matrix(c(1, .6, .6, 1), 2)
  x <- rmvnorm(n, sigma = sigma)
  gam <- c(.1, -.2)
  z <- rbinom(n, 1, plogis(1-1/(1+exp(-.5+x%*%gam+.1*rnorm(n)))))

  bet <- c(-.2,.2)
  #y <- rbinom(n, 1, plogis(1-1/(1+exp(-.5+x%*%bet + .2*z-.3*rnorm(n)))))
  y <- -.5+x%*%bet + .2*z-.3*rnorm(n)

  data.frame(y = y, z = z, x1 = x[, 1], x2 = x[, 2])

}

dat <- genData(123456)
dat1 <- head(dat,200)
dat2 <- tail(dat,200)
```

```
## fitting four models simultaneously.
fit <-
  multipleOutcomes(
    y ~ z + x1 - 1,
    z ~ x1 + x2,
    z ~ x1 - 1,
    y ~ x2,
    ## z can be fitted with a linear or logistic regression
    family = c('gaussian', 'binomial', 'gaussian','gaussian'),
    data = list(dat1, dat2),
    ## each dataset is used to fit two models
    data_index = c(1, 1, 2, 2)
  )

  ## unnamed coefficients of all model parameters
  coef(fit)

  ## named coefficients of a specific model
  coef(fit, 2)

  ## unnamed covariance matrix of all model parameters
  vcov(fit)

  ## named covariance matrix of a specific model
  vcov(fit, 1)

  ## summary of all parameter estimates
  summary(fit)

  ## summary of parameters in a specific model
  summary(fit, 4)
```

---

pated                    *Prognostic Variables Assisted Treatment Effect Detection*

---

### Description

`pated` is a wrapper function of `multipleOutcomes` for testing treatment effect in randomized clinical trials. It assumes that prognostic variables are fully randomized. This assumption can help enhancing statistical power of conventional approaches in detecting the treatment effect. Specifically, the sensitivity of the conventional models specified in ... are improved by `pated`.

### Usage

```
pated(..., family, data)
```

**Arguments**

| | |
|---|---|
| `...` | formulas of models to be fitted, or moment functions for gmm. |
| `family` | a character vector of families to be used in the models. Currently only `gaussian`, `binomial`, `coxph` and `gmm` are supported. `long` for longitudinal data may be supported in the future. `family` can be of length 1 if all models are fitted in thesame family; otherwise family should be specified for each of the models in `...`. |
| `data` | a data frame if all models are fitted on the same dataset; otherwise a list of data frames for fitting models in `...`. Note that a dataset can be used to fit multiple models, thus, `length(data)` is unnecessary to be equal to the number of models in `...`. The row names in a data frame are treated as sample IDs. Consequently, for any two records in different data frames that correspond to the same sample, their row names should be consistent. |

**Value**

a data frame of testing results.

**Examples**

```
# see vignettes
```

---

print.summary.multipleOutcomes
*Title Summarize an Analysis of Multiple Outcomes.*

---

**Description**

Summarize an analysis of multiple outcomes.

**Usage**

```
## S3 method for class 'summary.multipleOutcomes'
print(x, ...)
```

**Arguments**

| | |
|---|---|
| `x` | an object returned by `multipleOutcomes()`. |
| `...` | for debugging only. |

**Value**

an invisible object.

**Examples**

```
## no example
```

---

summary.multipleOutcomes

*Object Summaries*

---

### Description

summary method for class `multipleOutcomes`.

### Usage

```
## S3 method for class 'multipleOutcomes'
summary(object, model_index = NULL, ...)
```

### Arguments

| | |
|---|---|
| `object` | an object returned by `multipleOutcomes()`. |
| `model_index` | NULL if displaying summary of all fitted models; otherwise, an integer indicating the fitted model. |
| `...` | for debugging only |

### Value

a list

---

vcov.multipleOutcomes   *Calculate Variance-Covariance Matrix for a Fitted Model Object*

---

### Description

Returns the variance-covariance matrix of the main parameters of fitted model objects. The "main" parameters of models correspond to those returned by `coef`.

### Usage

```
## S3 method for class 'multipleOutcomes'
vcov(object, model_index = NULL, ...)
```

### Arguments

| | |
|---|---|
| `object` | an object returned by `multipleOutcomes()`. |
| `model_index` | NULL if displaying covariance matrix of all fitted models; otherwise, an integer indicating the fitted model. |
| `...` | for debugging only |

### Value

a matrix of covariance of all estimates

# Index