Package ‘LongDecompHE’

May 7, 2026

Type Package

Title Longitudinal Decomposition of Health Expectancy by Age and Cause

Version 0.1.0

Description Provides tools to decompose differences in cohort health expectancy (HE) by age and cause using longitudinal data. The package implements a novel longitudinal attribution method based on a semiparametric additive hazards model with time-dependent covariates, specifically designed to address interval censoring and semi-competing risks via a copula framework. The resulting age-cause-specific contributions to disability prevalence and death probability can be used to quantify and decompose differences in cohort HE between groups. The package supports stepwise replacement decomposition algorithms and is applicable to cohort-based health disparity research across diverse populations. Related methods include Sun et al. (2023) <doi:10.1177/09622802221133552>.

License GPL (>= 3)

Imports copula, corpcor, stats, ggplot2, patchwork, grDevices, tidyr

Encoding UTF-8

LazyData true

RoxygenNote 7.3.2

NeedsCompilation no

Author Huiping Zheng [aut, cre],
Tao Sun [aut],
Xiaojun Wang [aut]

Maintainer Huiping Zheng <`zhenghuiping@ruc.edu.cn`>

Depends R (>= 3.5.0)

Repository CRAN

Date/Publication 2025-07-03 15:20:02 UTC

Contents

attributionA 2

attributionB . . . 3
Attribution_sullivan . . . 4
copula_additive . . . 5
Decomp_sullivan . . . 7
decom_results . . . 8
Plot.sullivan . . . 9
print.LongDecompHE . . . 10
simulated_dataA . . . 10
simulated_dataB . . . 11
summary.LongDecompHE . . . 12

Index **13**

`attributionA` *Attribution Results for Population A*

Description

Attribution results for Population A, generated by applying the `Attribution_sullivan()` function under a copula-based additive hazards model. This object contains age- and cause-specific contributions to disability prevalence and death probability, enabling decomposition of cohort health expectancy.

Usage

```
data(attributionA)
```

Format

A list with the following components:

`Relative_Contributions_1` A matrix of age-cause-specific disability prevalence relative contributions. Rows represent causes (with the first row as total), columns represent age.

`Relative_Contributions_2` A matrix of age-cause-specific death probability relative contributions. Format matches `Relative_Contributions_1`.

`Absolute_Contributions_1` A matrix of age-cause-specific disability prevalence contributions. Rows represent causes (with the first row as total), columns represent age.

`Absolute_Contributions_2` A matrix of age-cause-specific death probability contributions. Format matches `Absolute_Contributions_1`.

`var_list` A character vector of cause names used in the attribution method.

`age_seq` A vector of age range.

Source

Generated for the manuscript: Tao Sun, Huiping Zheng, and Xiaojun Wang (2025+).\ *Decomposing Differences in Cohort Health Expectancy by Cause and Age with Longitudinal Data*.

Examples

```
data(attributionA)
str(attributionA)
```

`attributionB` *Attribution Results for Population B*

Description

Attribution results for Population B, generated by applying the `Attribution_sullivan()` function under a copula-based additive hazards model. This object contains age- and cause-specific contributions to disability prevalence and death probability, enabling decomposition of cohort health expectancy.

Usage

```
data(attributionB)
```

Format

A list with the following components:

`Relative_Contributions_1` A matrix of age-cause-specific disability prevalence relative contributions. Rows represent causes (with the first row as total), columns represent age.

`Relative_Contributions_2` A matrix of age-cause-specific death probability relative contributions. Format matches `Relative_Contributions_1`.

`Absolute_Contributions_1` A matrix of age-cause-specific disability prevalence contributions. Rows represent causes (with the first row as total), columns represent age.

`Absolute_Contributions_2` A matrix of age-cause-specific death probability contributions. Format matches `Absolute_Contributions_1`.

`var_list` A character vector of cause names used in the attribution method.

`age_seq` A vector of age range.

Source

Generated for the manuscript: Tao Sun, Huiping Zheng, and Xiaojun Wang (2025+).\ *Decomposing Differences in Cohort Health Expectancy by Cause and Age with Longitudinal Data*.

Examples

```
data(attributionB)
str(attributionB)
```

`Attribution_sullivan` *Longitudinal Attribution of Disability and Death Based on Sullivan Method*

Description

Computes the relative and/or absolute contributions of causes (i.e., covariates) to disability and death across a specified age range, based on a copula regression models with semiparametric additive hazards margins object. This attribution approach adopts a Sullivan method for cohort health expectancy.

Usage

```
Attribution_sullivan(object, type.attrib = "both")
```

Arguments

`object`	A fitted `LongDecompHE` object returned by `copula_additive`.
`type.attrib`	Type of attribution output to return; can be one of: • `"rel"`: Relative contributions of causes over age (i.e., time); • `"abs"`: Absolute contributions of causes over age (i.e., time); • `"both"`: Both relative and absolute contributions (default).

Details

This function loops over all attribution-eligible ages in the fitted model (from 0 to `floor(u-1)`), Use `summary()` on the returned object to print the attribution tables.

Value

A list containing the following components:

Relative_Contributions_1 Matrix of relative contributions to disability prevalence by cause and age

Relative_Contributions_2 Matrix of relative contributions to death probability by cause and time

Absolute_Contributions_1 Matrix of absolute contributions to disability prevalence by cause and age

Absolute_Contributions_2 Matrix of absolute contributions to death probability by cause and time

var_list Vector of covariate names used in the model

time_list Vector of attribution ages

copula Character indicating attribution type ("Attribution_sullivan")

#' Additional components `copula` and `summary` are included for compatibility with generic

See Also

`copula_additive` for model fitting; `summary.LongDecompHE` for output methods.

Examples

```
# Fit a model (see copula_additive)
data(simulated_dataA)
u1 = u2 = max(simulated_dataA$visit_time)
var_list = c("Z1", "Z2", "Z3")
copula_additive_model <-  copula_additive(data = simulated_dataA,
                                          var_list = var_list,
                                          l1=0, u1 = u1, m1 = 3,
                                          l2=0, u2 = u2, m2 = 3,
                                          method = "combined", iter=1000,
                                          stepsize=1e-6,
                                          hes = TRUE,
                                          control = list(maxit = 10000))
summary(copula_additive_model)
# Attribution analysis (both relative and absolute)
attributionA <- Attribution_sullivan(object = copula_additive_model, type.attrib = "both")
summary(attributionA)

# Relative only
attributionA_rel <- Attribution_sullivan(object = copula_additive_model, type.attrib = "rel")

# Absolute only
attributionA_abs <- Attribution_sullivan(object = copula_additive_model, type.attrib = "abs")
```

`copula_additive` *Copula regression models with semiparametric additive hazards margins for bivariate interval censoring and semi-competing risks.*

Description

Fits a copula model with semiparametric additive hazards marginsfor bivariate interval censoring and semi-competing risks.

Usage

```
copula_additive(
  data,
  var_list,
  l1 = 0,
  u1,
  m1 = 3,
  l2 = 0,
  u2,
  m2 = 3,
```

```
  method = "Newton",
  iter = 1000,
  stepsize = 1e-05,
  hes = TRUE,
  control = list(),
  eta_ini = NULL
)
```

Arguments

`data`	a data frame; must have `id` (subject id), `visit` (the visit number for the subject), `visit_time` (the time for each visit in years), `status` (the disability status at visit_time, 1 for disability, 0 for non-disability), `statusD` (the death status at visit_time, 1 for dead, 0 for alive), and `weight` sampling weight and each subject received the same weight across visits.
`var_list`	the list of covariates to be fitted into the copula model.
`l1`	Left boundary of event 1 time interval.
`u1`	Right boundary of event 1 time interval.
`m1`	Degree of Bernstein polynomial for event 1.
`l2`	Left boundary of event 2 time interval.
`u2`	Right boundary of event 2 time interval.
`m2`	Degree of Bernstein polynomial for event 2.
`method`	optimization method (see `?optim`); default is `"BFGS"`; also can be `"Newton"` (see `?nlm`).
`iter`	number of iterations when `method = "Newton"`; default is 300.
`stepsize`	size of optimization step when method is `"Newton"`; default is 1e-6.
`hes`	default is `TRUE` for hessian calculation; if LRT is desired, can set `hes = FALSE` to save time.
`control`	a list of control parameters for methods other than `"Newton"`; see ?optim.
`eta_ini`	Optional initial values for copula parameters.

Details

must have `id` (subject id), `visit` (the visit number for the subject), `visit_time` (the time for each visit in years), `status` (the disability status at visit_time, 1 for disability, 0 for non-disability), `statusD` (the death status at visit_time, 1 for dead, 0 for alive), and `weight` sampling weight and each subject received the same weight across visits.

Optimization methods can be all methods (except `"Brent"`) from `optim`, such as `"Nelder-Mead"`, `"BFGS"`, `"CG"`, `"L-BFGS-B"`, `"SANN"`. Users can also use `"Newton"` (from `nlm`).

Value

a `LongDecompHE` object summarizing the model. Can be used as an input to general `S3` methods including `summary`, `print`, `plot`, `coef`, `logLik`, `AIC`.

Source

Tao Sun, Huiping Zheng, and Xiaojun Wang (2025+). Decomposing Differences in Cohort Health Expectancy by Cause and Age with Longitudinal Data. Under review.

Examples

```
# Fit a Copula2-Semiparametric additive hazards model
data(simulated_dataA)
u1 = u2 = max(simulated_dataA$visit_time)
var_list = c("Z1", "Z2", "Z3")
copula_additive_model <-  copula_additive(data = simulated_dataA,
                                        var_list = var_list,
                                        l1=0, u1 = u1, m1 = 3,
                                        l2=0, u2 = u2, m2 = 3,
                                        method = "combined", iter=1000,
                                        stepsize=1e-6,
                                        hes = TRUE,
                                        control = list(maxit = 10000))
summary(copula_additive_model)
```

`Decomp_sullivan` *Two-dimensional decomposition of the difference between the two cohort HE by age and cause*

Description

This function implements the two-dimensional stepwise replacement decomposition algorithm (Andreev et al., 2002) to quantify age-cause-specific effect of mortality and disability to the difference in cohort health expectancy between two populations, based on longitudinal attribution results.

Usage

```
Decomp_sullivan(attribution1, attribution2)
```

Arguments

`attribution1`	A list returned by `Attribution_sullivan()`, representing Population 1.
`attribution2`	A list returned by `Attribution_sullivan()`, representing Population 2.

Value

A list including matrices of mortality effects, disability effects, and total effects, as well as cohort HE and LE for each population.

Examples

```
data(attributionA)
data(attributionB)
summary(attributionA)
summary(attributionB)
decom_results <- Decomp_sullivan(attribution1=attributionA,attribution2=attributionA)
summary(decom_results)
```

`decom_results` *Decomposition Results Comparing Two Populations*

Description

Decomposition results obtained from the `Decomp_sullivan()` function. This object includes age-cause-specific mortality, disability, and total effects to differences in cohort health expectancy (HE).

Usage

```
data(decom_results)
```

Format

A list with the following components:

- `mortality_effect` A matrix with rows representing total, background, and causes (e.g., Z1, Z2, Z3), and columns representing age groups. Each cell indicates the contribution of a given cause to the mortality effect on health expectancy differences.
- `disability_effect` A matrix structured as above, representing contributions to the disability effect.
- `total_effect` A matrix summing the mortality and disability effects for each cause at each age.
- `population1_HE` Numeric value of cohort health expectancy (HE) in population 1.
- `population2_HE` Numeric value of cohort health expectancy (HE) in population 2.
- `population1_LE` Numeric value of cohort life expectancy (LE) in population 1.
- `population2_LE` Numeric value of cohort life expectancy (LE) in population 2.
- `var_list` A character vector of cause names (e.g., `c("Z1", "Z2", "Z3")`).
- `time_list` An integer vector indicating age groups.
- `copula` A string indicating the method used for decomposition, e.g., `"Decomposition_sullivan"`.

Source

Generated for the manuscript:
Tao Sun, Huiping Zheng, and Xiaojun Wang (2025+).
Decomposing Differences in Cohort Health Expectancy by Cause and Age with Longitudinal Data.

Examples

```
data(decom_results)
str(decom_results)
```

`Plot.sullivan` *Plot Attribution or Decomposition Results Using Sullivan Method*

Description

This function generates visualization plots for results from either longitudinal attribution models (e.g., ‘Attribution_sullivan‘) or decomposition models (e.g., ‘Decomp_sullivan‘) applied to cohort health expectancy.

Usage

```
Plot.sullivan(result, var_list = NULL, colors = NULL)
```

Arguments

`result`	A list object returned by either an attribution function (must contain ‘Absolute_Contributions_1‘ and ‘Absolute_Contributions_2‘) or a decomposition function (must contain ‘total_effect‘, ‘mortality_effect‘, and ‘disability_effect‘).
`var_list`	A character vector specifying the names of variables (e.g., diseases or risk factors) to include in the plots.
`colors`	Optional. A named character vector of colors for each category (including ‘"Background"‘). If ‘NULL‘, a color palette will be generated automatically.

Details

This function supports two types of Sullivan-based outputs:

1. **Attribution results:** The input should contain components named ‘Absolute_Contributions_1‘ (for disability) and ‘Absolute_Contributions_2‘ (for mortality), as returned by the ‘Attribution_sullivan()‘ function.
2. **Decomposition results:** The input should contain ‘total_effect‘, ‘mortality_effect‘, and ‘disability_effect‘ matrices, typically produced by the ‘Decomp_sullivan()‘ function.

For attribution results, the function internally computes relative contributions as a share of total years lost or gained.

Value

A ‘patchwork‘ object combining multiple ‘ggplot2‘ barplots. The layout depends on the result type:

- For attribution results, four plots are returned: absolute and relative contributions to disability and death.
- For decomposition results, three plots are returned: total, mortality, and disability effects.

Examples

```
# For attribution result
data(attributionA)
Plot.sullivan(result=attributionA, var_list = c("Z1", "Z2", "Z3"))

# For decomposition result
data(decom_results)
Plot.sullivan(result=decom_results, var_list = c("Z1", "Z2", "Z3"))
```

`print.LongDecompHE` *Printing outputs of a LongDecompHE object*

Description

Printing outputs of a LongDecompHE object

Usage

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

Arguments

`x`	a LongDecompHE object
`...`	further arguments

Value

No return value, called for prints model details to the console.

`simulated_dataA` *Simulated Bivariate Semi-competing Risks Data (Set A)*

Description

A simulated dataset for evaluating copula-based semi-competing risks models with semiparametric additive hazards. Each subject may experience disability followed by death, with interval-censored disability times and right-censored death times.

Usage

```
data(simulated_dataA)
```

Format

A data frame with 2000 subjects and the following variables:

`id`	subject identifier
`visit`	visit index per subject
`visit_time`	observation time at each visit (in years)
`status`	disability status at visit_time (1 = disabled, 0 = non-disabled)
`statusD`	death status at visit_time (1 = dead, 0 = alive)
`Z1`	first covariate (binary)
`Z2`	second covariate (continuous)
`Z3`	third covariate (categorical: 0/1/2)
`weight`	sampling weight; constant across visits within a subject

Source

Simulated for the manuscript: Tao Sun, Huiping Zheng, and Xiaojun Wang (2025+). *Decomposing Differences in Cohort Health Expectancy by Cause and Age with Longitudinal Data*.

Examples

```
data(simulated_dataA)
head(simulated_dataA)
```

`simulated_dataB` *Simulated Bivariate Semi-competing Risks Data (Set B)*

Description

A simulated dataset for evaluating copula-based semi-competing risks models with semiparametric additive hazards. Each subject may experience disability followed by death, with interval-censored disability times and right-censored death times.

Usage

`data(simulated_dataB)`

Format

A data frame with 2000 subjects and the following variables:

`id`	subject identifier
`visit`	visit index per subject
`visit_time`	observation time at each visit (in years)
`status`	disability status at visit_time (1 = disabled, 0 = non-disabled)

`statusD` death status at visit_time (1 = dead, 0 = alive)

`Z1` first covariate (binary)

`Z2` second covariate (continuous)

`Z3` third covariate (categorical: 0/1/2)

`weight` sampling weight; constant across visits within a subject

Source

Simulated for the manuscript: Tao Sun, Huiping Zheng, and Xiaojun Wang (2025+). *Decomposing Differences in Cohort Health Expectancy by Cause and Age with Longitudinal Data*.

Examples

```
data(simulated_dataB)
head(simulated_dataB)
```

`summary.LongDecompHE` *Summarizing outputs of a LongDecompHE object*

Description

#' This function prints or returns a summary of results from an attribution or decomposition analysis using the `LongDecompHE` framework.

Usage

```
## S3 method for class 'LongDecompHE'
summary(object, ...)
```

Arguments

`object`	a LongDecompHE object
`...`	further arguments

Value

If the input is an attribution or decomposition object (e.g., from `Attribution_sullivan` or `Decomp_sullivan`), the function prints summary tables of contributions to disability and death and returns the object invisibly. Otherwise, it returns a list with model diagnostics including:

- `copula`: the copula family used.
- `m.dist`: marginal distribution type.
- `summary`: coefficient estimates and standard errors.
- `llk`: log-likelihood.
- `AIC`: Akaike Information Criterion.
- `code`: convergence code.

The returned object has class `"summary.LongDecompHE"`.

Index

∗ **datasets**
    attributionA, 2
    attributionB, 3
    decom_results, 8
    simulated_dataA, 10
    simulated_dataB, 11

Attribution_sullivan, 4
attributionA, 2
attributionB, 3

copula_additive, *5*, 5

decom_results, 8
Decomp_sullivan, 7

Plot.sullivan, 9
print.LongDecompHE, 10

simulated_dataA, 10
simulated_dataB, 11
summary.LongDecompHE, *5*, 12