# Package 'lillies'

August 7, 2023
Title Estimation of Life Years Lost
<b>Version</b> 0.2.12
Description Estimation of life expectancy and Life Years Lost (LYL, or lillies for short) for a given population, for example those with a given disease or condition. In addition, the package can be used to compare estimates from different populations, or to estimate confidence intervals. Technical details of the method are available in Plana-Ripoll et al. (2020) <doi:10.1371 journal.pone.0228073=""> and Andersen (2017) <doi:10.1002 sim.7357="">.</doi:10.1002></doi:10.1371>
<b>Depends</b> R ( $>= 3.5.0$ )
License MIT + file LICENSE
Encoding UTF-8
LazyData true
RoxygenNote 7.2.3
<b>Imports</b> dplyr, knitr, pracma, progress, rlang, survival, tidyr, utils, methods
Suggests ggplot2
NeedsCompilation no
Author Oleguer Plana-Ripoll [aut, cre] ( <a href="https://orcid.org/0000-0002-6470-7465">https://orcid.org/0000-0002-6470-7465</a> )
Maintainer Oleguer Plana-Ripoll <opr@clin.au.dk></opr@clin.au.dk>
Repository CRAN
<b>Date/Publication</b> 2023-08-07 10:40:02 UTC
R topics documented:
aggreg_data lyl . lyl_2plot . lyl_aggregated . lyl aggregated range

2 aggreg\_data

	lyl_checkplot	1
	lyl_ci	2
	lyl_compare_plot	3
	lyl_diff	5
	lyl_diff_ref	7
	yl_range	
	plot.lyl	
	plot.lyl_aggregated	
	plot.lyl_ci	
	plot.lyl_range	
	pop_ref	
	simu_data	
	summary.lyl	
	summary.lyl_aggregated	9
	summary.lyl_ci	0
	summary.lyl_range	
Index	3	3
		_
aggre	data Simulated aggregated data for Life Years Lost estimation	

## Description

A dataset containing age-specific number of new cases, number of deaths and mortality rates for a simulated disease with onset after age 40 years. Data is available for ages 40-90 years.

## Usage

aggreg\_data

## **Format**

A data frame with 50 rows and 4 variables:

age age

new\_cases number of new cases diagnosed at that specific age

deaths number of deaths among the diagnosed at that specific age

rate age-specific mortality rates among the diagnosed

## Source

Simulated data

1y1

Life Years Lost at one specific age.

## **Description**

1yl estimates remaining life expectancy and Life Years Lost for a given population after a specific age age\_speficic and restricted to a maximum theoretical age  $\tau$ .

3

## Usage

```
ly1(
  data,
  t0 = NULL
  t,
  status,
  age_specific,
  censoring_label = "Alive",
  death_labels = "Dead",
  tau = 100
)
```

## **Arguments**

data	A dataframe, where each raw represents a person. The dataframe will have a
	4' 4 4 4

time-to-event format with at least two variables: age at end of follow-up (t) and status indicator with death/censoring (status). Note that this package is not

developed to be used with tibbles.

t0 Age at start of the follow-up time. Default is NULL, which means all subjects

are followed from birth. For delayed entry, t0 indicates age at beginning of

follow-up.

Age at the end of the follow-up time (death or censoring). t

status Status indicator, normally 0=alive, 1=dead. Other choices are TRUE/FALSE

> (TRUE = death) or 1/2 (2=death). For multiple causes of death (competing risks analysis), the status variable will be a factor, whose first level is treated as censoring; or a numeric variable, whose lowest level is treated as censoring. In the

latter case, the label for censoring is censoring\_label ("Alive" by default).

Specific age at which the Life Years Lost have to be estimated. age\_specific

censoring\_label

Label for censoring status. If status is not a factor, "Alive" by default. If status is a factor, the first level will be treated as censoring label.

death\_labels Label for event status. For only one cause of death, "Dead" is the default. For

multiple causes, the default are the values given in variable status.

tau Remaining life expectancy and Life Years Lost are estimated restrictied to a

maximum theoretical age  $\tau$  ( $\tau$ =100 years by default).

lyl

4 lyl

#### Value

A list with class "lyl" containing the following components:

- data: Data frame with 3 variables and as many observations as the original data provided to estimate Life Years Lost: t0, t, and status
- LYL: Data frame with 1 observation and at least 3 variables: age which corresponds to age\_spefific; life\_exp which is the estimated remaining life expectancy at age age\_specific years and before age tau years; and one variable corresponding to the estimated Life Years Lost for each specific cause of death. If only one cause of death is considered (no competing risks), this variable is Dead and includes the total overall Life Years Lost
- tau: Maximum theoretical age au
- age\_specific: Specific age at which the Life Years Lost have been estimated
- data\_plot: A data frame in long format with 3 variables time, cause, and cip used to create a Figure of Life Years Lost with function plot.
- censoring\_label: Label for censoring status
- death\_labels: Label(s) for death status
- competing\_risks: Logical value (TRUE = more than one cause of death (competing risks))
- type: Whether the estimation is at "age\_specific" or "age\_range".

#### References

- Andersen PK. Life years lost among patients with a given disease. *Statistics in Medicine*. 2017;36(22):3573-3582.
- Andersen PK. Decomposition of number of life years lost according to causes of death. *Statistics in Medicine*. 2013;32(30):5278-5285.
- Plana-Ripoll et al. lillies An R package for the estimation of excess Life Years Lost among patients with a given disease or condition. *PLoS ONE*. 2020;15(3):e0228073.

#### See Also

- lyl\_range for estimation of Life Years Lost for a range of different ages.
- lyl\_ci to estimate bootstrapped confidence intervals.
- lyl\_diff to compare Life Years Lost for two populations.
- summary.lyl to summarize objects obtained with function lyl.
- plot.lyl to plot objects obtained with function lyl.

lyl\_2plot 5

lyl\_2plot

Plot Life Years Lost at one specific age for two different populations

## Description

lyl\_2plot was used to create a figure of Life Years Lost at one specific age for two different populations. Please use lyl\_compare\_plot instead.

## Usage

```
lyl_2plot(
    x,
    y,
    color_alive = NA,
    colors = NA,
    labels = c("Population of interest", "Reference population"),
    ...
)
```

## **Arguments**

X	An object of class 1y1 (obtained with function 1y1).
У	An object of class lyl (obtained with function lyl).
color_alive	Color to be used for the censoring category. Default is NA, and default color is "white".
colors	Vector with one color for each cause of death. Default is NA, and default colors are used.
labels	Vector with labels for the two populations (default are "Population of interest" for $x$ , and "Reference population" for $y$ )
	Additional arguments affecting the plot produced.

6 lyl\_aggregated

#### Value

A plot with survival function and stacked cause-specific cumulative incidences for two populations side by side.

#### References

• Plana-Ripoll et al. lillies – An R package for the estimation of excess Life Years Lost among patients with a given disease or condition. *PLoS ONE*. 2020;15(3):e0228073.

#### See Also

- 1yl for estimation of Life Years Lost at one specific age.
- lyl\_diff to compare Life Years Lost for two populations.

#### **Examples**

lyl\_aggregated

Life Years Lost at one specific age using aggregated data.

## Description

lyl estimates differences in remaining life expectancy and Life Years Lost for two given life tables data and data0 after a specific age age\_speficic and restricted to a maximum theoretical age  $\tau$ .

## Usage

```
lyl_aggregated(
  data,
  age,
  rates,
```

lyl\_aggregated 7

```
surv,
data0,
age0,
rates0,
surv0,
age_specific,
censoring_label = "Alive",
death_labels = "Dead",
tau = 100
)
```

## **Arguments**

data	A dataframe, where each raw represents an age, for the population of interest. The dataframe will contain information on age-specific mortality rates or survivial probability (if both parameters are provided, rates will be used).	
age	Variable in data containing information on age.	
rates	Variable in data containing information on age-specific mortality rates.	
surv	Variable in data containing information on age-specific survival probability.	
data0	A dataframe, where each raw represents an age, for the population of reference The dataframe will contain information on age-specific mortality rates or survivial probability (if both parameters are provided, rates will be used).	
age0	Variable in data0 containing information on age.	
rates0	Variable in data0 containing information on age-specific mortality rates.	
surv0	Variable in data0 containing information on age-specific survival probability.	
age_specific	Specific age at which the Life Years Lost have to be estimated.	
censoring_label		
	Label for censoring status ("Alive" by default).	
death_labels	Label for event status ("Dead" by default).	
tau	Remaining life expectancy and Life Years Lost are estimated restrictied to a maximum theoretical age $\tau$ ( $\tau$ =100 years by default).	

## Value

A list with class "lyl\_aggregated" containing the following components:

- data: Name of the dataset preovided in parameter data
- data0: Name of the dataset preovided in parameter data0
- LYL: Data frame with 1 observation and 3 variables: age which corresponds to age\_spefific; and life\_exp and life\_exp0 which are the estimated remaining life expectancies at age age\_specific years and before age tau years for the population provided in data and data0, respectively
- tau: Maximum theoretical age au
- age\_specific: Specific age at which the Life Years Lost have been estimated

lyl\_aggregated\_range

• data\_plot: A data frame in long format with 3 variables time, cause, and cip used to create a Figure of Life Years Lost with function plot.

- censoring\_label: Label for censoring status
- death\_labels: Label(s) for death status
- type: Whether the estimation is at "age\_specific" or "age\_range"

#### References

8

- Andersen PK. Life years lost among patients with a given disease. *Statistics in Medicine*. 2017;36(22):3573-3582.
- Andersen PK. Decomposition of number of life years lost according to causes of death. *Statistics in Medicine*. 2013;32(30):5278-5285.
- Plana-Ripoll et al. lillies An R package for the estimation of excess Life Years Lost among patients with a given disease or condition. *PLoS ONE*. 2020;15(3):e0228073.#'

#### See Also

- lyl\_aggregated\_range for estimation of Life Years Lost for a range of different ages.
- summary.lyl\_aggregated to summarize objects obtained with function lyl\_aggregated.
- plot.lyl\_aggregated to plot objects obtained with function lyl\_aggregated.

## **Examples**

## **Description**

lyl estimates differences in remaining life expectancy and Life Years Lost for two given life tables data and data0 after a range of specific ages (age\_begin to age\_end) and restrictied to a maximum theoretical age  $\tau$ .

lyl\_aggregated\_range 9

## Usage

```
lyl_aggregated_range(
  data,
  age,
  rates,
  surv,
  weights,
  data0,
  age0,
  rates0,
  surv0,
  age_begin,
  age_end,
  censoring_label = "Alive",
  death_labels = "Dead",
  tau = 100
)
```

## Arguments

data	A dataframe, where each raw represents an age, for the population of interest. The dataframe will contain information on age-specific mortality rates or survivial probability (if both parameters are provided, rates will be used).
age	Variable in data containing information on age.
rates	Variable in data containing information on age-specific mortality rates.
surv	Variable in data containing information on age-specific survival probability.
weights	Variable in data containing information on number of new cases per age. A weighted average is provided using these weights.
data0	A dataframe, where each raw represents an age, for the population of reference The dataframe will contain information on age-specific mortality rates or survivial probability (if both parameters are provided, rates will be used).
age0	Variable in data0 containing information on age.
rates0	Variable in data0 containing information on age-specific mortality rates.
surv0	Variable in data0 containing information on age-specific survival probability.
age_begin	Specific starting age at which the Life Years Lost have to be estimated.
age_end	Specific ending age at which the Life Years Lost have to be estimated.
censoring_labe	1
	Label for censoring status ("Alive" by default).
death_labels	Label for event status ("Dead" by default).
tau	Remaining life expectancy and Life Years Lost are estimated restrictied to a maximum theoretical age $\tau$ ( $\tau$ =100 years by default).

#### Value

A list with class "lyl\_aggregated" containing the following components:

- data: Name of the dataset preovided in parameter data
- data0: Name of the dataset preovided in parameter data0
- LYL: Data frame with 1 observation and 2 variables: life\_exp and life\_exp0 which are the estimated remaining life expectancies averaged over the age range and before age tau years for the population provided in data and data0, respectively
- tau: Maximum theoretical age au
- age\_begin: Specific starting age at which the Life Years Lost have been estimated
- age\_end: Specific ending age at which the Life Years Lost have been estimated
- data\_plot: A data frame in long format with 3 variables time, cause, and cip used to create a Figure of Life Years Lost with function plot.
- censoring\_label: Label for censoring status
- death\_labels: Label(s) for death status
- type: Whether the estimation is at "age\_specific" or "age\_range".

## References

- Andersen PK. Life years lost among patients with a given disease. *Statistics in Medicine*. 2017;36(22):3573-3582.
- Andersen PK. Decomposition of number of life years lost according to causes of death. *Statistics in Medicine*. 2013;32(30):5278-5285.
- Plana-Ripoll et al. lillies An R package for the estimation of excess Life Years Lost among patients with a given disease or condition. *PLoS ONE*. 2020;15(3):e0228073.#'

#### See Also

- lyl\_aggregated for estimation of Life Years Lost at one specific age.
- summary.lyl\_aggregated to summarize objects obtained with function lyl\_aggregated\_range.

Iyl\_checkplot

lyl\_checkplot

Plot number of persons at risk at each specific age.

## **Description**

Given a lyl\_range-class object, lyl\_checkplot draws numbers of persons at risk of dying at each specific age from age\_begin until age  $\tau$ .

## Usage

```
lyl_checkplot(x)
```

## Arguments

Х

An object of class lyl\_range obtained with the lyl\_range function.

#### Value

A plot with the number of persons at risk at each specific age.

#### References

• Plana-Ripoll et al. lillies – An R package for the estimation of excess Life Years Lost among patients with a given disease or condition. *PLoS ONE*. 2020;15(3):e0228073.

## See Also

lyl\_range for estimation of Life Years Lost for a range of different ages.

12 lyl\_ci

lyl\_ci

Confidence intervals for Life Years Lost.

## **Description**

lyl\_ci estimates confidence intervals for Life Years Lost using non-parametric bootstrap. The confidence level can be specified when summarizing the results with the function summary.lyl\_ci.

#### **Usage**

```
lyl_ci(lyl_estimation, niter = 1000)
```

## **Arguments**

```
lyl_estimation An object of class lyl or lyl_range.

niter Number of iterations for the bootstrap (default is 1,000).
```

#### Value

A list with class "lyl\_ci" containing the following components:

- LYL: Data frame with one observation per age and at least 3 variables: age; life\_exp which is the estimated remaining life expectancy at age age\_specific years and before age tau years; and one variable corresponding to the estimated Life Years Lost for each specific cause of death. If only one cause of death is considered (no competing risks), this variable is Dead and includes the total overall Life Years Lost
- LYL\_ci: Data frame with one observation per age-iteration and at least 4 variables: age; iteration, which correspond to each specific iteration; life\_exp which is the estimated remaining life expectancy at age age\_specific years and before age tau years; and one variable corresponding to the estimated Life Years Lost for each specific cause of death. If only one cause of death is considered (no competing risks), this variable is Dead and includes the total overall Life Years Lost
- tau: Maximum theoretical age au
- age\_specific: Specific age at which the Life Years Lost have been estimated
- age\_begin: Specific starting age at which the Life Years Lost have been estimated
- age\_end: Specific ending age at which the Life Years Lost have been estimated
- death\_labels: Label(s) for death status
- competing\_risks: Logical value (TRUE = more than one cause of death (competing risks))
- type: Whether the estimation is at "age\_specific" or "age\_range".
- niter: Number of iterations used to estimate the confidence intervals

#### References

• Plana-Ripoll et al. lillies – An R package for the estimation of excess Life Years Lost among patients with a given disease or condition. *PLoS ONE*. 2020;15(3):e0228073.

Iyl\_compare\_plot 13

#### See Also

- 1yl for estimation of Life Years Lost at one specific age.
- lyl\_range for estimation of Life Years Lost for a range of different ages.
- lyl\_diff to compare Life Years Lost for two populations.
- summary.lyl\_ci to summarize objects obtained with function lyl\_ci.
- plot.lyl\_ci to plot objects obtained with function lyl\_ci.

lyl\_range for estimation of Life Years Lost for a range of different ages.

## **Examples**

```
# Load simulated data as example
data(simu_data)
# Estimate remaining life expectancy and Life Years
# Lost after age 45 years and before age 95 years
lyl_estimation <- lyl(data = simu_data, t = age_death, status = death,</pre>
                      age\_specific = 45, tau = 95)
# Calculate bootstrapped confidence interval (3 iterations to test; more are necessary)
lyl_estimation_ci <- lyl_ci(lyl_estimation, niter = 3)</pre>
summary(lyl_estimation_ci)
plot(lyl_estimation_ci)
# Estimate remaining life expectancy and Life Years
# Lost after each age from 0 to 94 years and before age 95 years
lyl_estimation2 <- lyl_range(data = simu_data, t = age_death, status = death,</pre>
                             age\_begin = 0, age\_end = 94, tau = 95)
# Calculate bootstrapped confidence interval (3 iterations to test; more are necessary)
lyl_estimation_ci2 <- lyl_ci(lyl_estimation2, niter = 3)</pre>
summary(lyl_estimation_ci2, weights = simu_data$age_disease)
plot(lyl_estimation_ci2, weights = simu_data$age_disease)
```

lyl\_compare\_plot

Plot Life Years Lost at one specific age for two or more different populations

## **Description**

lyl\_compare\_plot creates a figure of Life Years Lost at one specific age for two or more different populations.

14 lyl\_compare\_plot

## Usage

```
lyl_compare_plot(
    x,
    color_alive = NA,
    colors = NA,
    nrow = NULL,
    ncol = NULL,
    dir = "h",
    reverse_legend = FALSE,
    labels = NA,
    ...
)
```

## **Arguments**

Χ	A list of objects of class lyl (obtained with function lyl).
color_alive	Color to be used for the censoring category. Default is NA, and default color is "white".
colors	Vector with one color for each cause of death. Default is NA, and default colors are used.
nrow	Number of rows to be passed to facet_wrap.
ncol	Number of columns to be passed to facet_wrap.
dir	Direction to be passed to facet_wrap: either "h" for horizontal, the default, or "v", for vertical.
reverse_legend	Reverse the order of elements in the legend. Ddefault is FALSE, indicating that first is the censoring label and then all causes of death.
labels	Vector with labels for the two populations (default are "Population of interest" for $x$ , and "Reference population" for $y$ )
	Additional arguments affecting the plot produced.

## Value

A plot with survival function and stacked cause-specific cumulative incidences for two populations side by side.

## References

• Plana-Ripoll et al. lillies – An R package for the estimation of excess Life Years Lost among patients with a given disease or condition. *PLoS ONE*. 2020;15(3):e0228073.

## See Also

- 1yl for estimation of Life Years Lost at one specific age.
- lyl\_diff to compare Life Years Lost for two populations.

lyl\_diff

## **Examples**

```
# Load simulated data as example
data(simu_data)
# Estimate remaining life expectancy and Life Years
# Lost after age 45 years and before age 95 years
lyl_estimation <- lyl(data = simu_data, t = age_death, status = cause_death,</pre>
                      age_specific = 45, tau = 95)
# Same estimate for those with a specific disease
diseased <- simu_data[!is.na(simu_data$age_disease), ]</pre>
lyl_estimation1 <- lyl(data = diseased, t0 = age_disease,</pre>
                       t = age_death, status = cause_death,
                      age_specific = 45, tau = 95)
# Plot the data
lyl_compare_plot(list(lyl_estimation1, lyl_estimation))
lyl_compare_plot(list(lyl_estimation1, lyl_estimation),
          labels = c("Population with a disease", "General population"))
# The plot can be modified with a usual ggplot2 format
lyl_compare_plot(list(lyl_estimation1, lyl_estimation)) +
 ggplot2::xlab("Age [in years]") +
 ggplot2::ggtitle("Differences in Life Years Lost at age 45 years")
```

lyl\_diff

Summarize differences in Life Years Lost.

## **Description**

 ${\tt lyl\_diff\ summarizes\ differences\ in\ estimated\ Life\ Years\ Lost\ in\ two\ different\ populations:\ lyl\_estimation\ compared\ to\ lyl\_estimation0}.$ 

## Usage

```
lyl_diff(
   lyl_population1,
   lyl_population0,
   decimals = 2,
   level = 0.95,
   weights = NA
)
```

lyl\_diff

## **Arguments**

lyl\_population1

Population of interest: An object of class lyl or lyl\_range (obtained with functions lyl or lyl\_range). Alternatively, an object of class lyl\_ci can be provided for bootstrapped confidence intervals.

lyl\_population0

decimals

Reference population: An object of class lyl or lyl\_range (obtained with functions lyl or lyl\_range). Alternatively, an object of class lyl\_ci can be provided for bootstrapped confidence intervals.

Number of decimals to be reported (default is 2).

level Confidence level if lyl\_population1 or lyl\_population0 is obtained with

the lyl\_ci function (default is 0.95 for 95% confidence intervals)

weights Vector with age distribution of disease/condition onset to be used when Life

Years Lost are estimated over a range of ages (with lyl\_range function). If weights are not provided (dafault is weights = NA), then the differences in Life Years Lost at each age is provided. If weights are provided, then a weighted

average is provided.

#### Value

A table with the summary of the differences between two populations.

#### References

• Plana-Ripoll et al. lillies – An R package for the estimation of excess Life Years Lost among patients with a given disease or condition. *PLoS ONE*. 2020;15(3):e0228073.

## See Also

- 1yl for estimation of Life Years Lost at one specific age.
- lyl\_range for estimation of Life Years Lost for a range of different ages.

lyl\_diff\_ref

```
lyl_compare_plot(list(lyl_estimation1, lyl_estimation0))
# Calculate bootstrapped confidence interval (3 iterations to test; more are necessary)
lyl_estimation1_ci <- lyl_ci(lyl_estimation1, niter = 3)</pre>
lyl_estimation0_ci <- lyl_ci(lyl_estimation0, niter = 3)</pre>
lyl_diff(lyl_estimation1_ci, lyl_estimation0_ci)
# It is also possible to assume no uncertainty for one of the estimates
lyl_diff(lyl_estimation1_ci, lyl_estimation0)
# Estimate remaining life expectancy and Life Years
# Lost due to specific causes of death after each age
# from 0 to 94 years and before age 95 years
lyl_estimation2 <- lyl_range(data = simu_data, t = age_death, status = cause_death,
                             age\_begin = 0, age\_end = 94, tau = 95)
lyl_estimation3 <- lyl_range(data = diseased, t0 = age_disease,</pre>
                             t = age_death, status = cause_death,
                              age\_begin = 0, age\_end = 94, tau = 95)
lyl_diff(lyl_estimation3, lyl_estimation2)
lyl_diff(lyl_estimation3, lyl_estimation2, weights = diseased$age_disease)
# Calculate bootstrapped confidence interval (3 iterations to test; more are necessary)
lyl_estimation3_ci <- lyl_ci(lyl_estimation3, niter = 3)</pre>
lyl_diff(lyl_estimation3_ci, lyl_estimation2, weights = diseased$age_disease)
```

lyl\_diff\_ref

Summarize differences in Life Years Lost.

#### Description

lyl\_diff summarizes differences in estimated Life Years Lost in two different populations: lyl\_estimation1 compared to a life table provided in data\_ref.

## Usage

```
lyl_diff_ref(
    lyl_population1,
    data_ref,
    age,
    surv,
    rates,
    decimals = 2,
    level = 0.95,
    weights = NA,
    lyl_population0
)
```

18 lyl\_diff\_ref

#### **Arguments**

lyl\_population1

Population of interest: An object of class lyl or lyl\_range (obtained with functions lyl or lyl\_range). Alternatively, an object of class lyl\_ci can be

provided for bootstrapped confidence intervals.

data\_ref A dataframe, where each raw represents an age, for the population of reference

The dataframe will contain information on age-specific mortality rates or sur-

vivial probability (if both parameters are provided, rates will be used).

age Variable in data\_ref containing information on age.

surv Variable in data\_ref containing information on age-specific survival probabil-

ity.

rates Variable in data\_ref containing information on age-specific mortality rates.

decimals Number of decimals to be reported (default is 2).

level Confidence level if lyl\_population1 or lyl\_population0 is obtained with

the lyl\_ci function (default is 0.95 for 95% confidence intervals)

weights Vector with age distribution of disease/condition onset to be used when Life

Years Lost are estimated over a range of ages (with lyl\_range function). If weights are not provided (dafault is weights = NA), then the differences in Life Years Lost at each age is provided. If weights are provided, then a weighted

average is provided.

lyl\_population0

Parameter automatically created.

#### Value

A table with the summary of the differences between two populations.

#### References

• Plana-Ripoll et al. lillies – An R package for the estimation of excess Life Years Lost among patients with a given disease or condition. *PLoS ONE*. 2020;15(3):e0228073.

#### See Also

- 1yl for estimation of Life Years Lost at one specific age.
- lyl\_range for estimation of Life Years Lost for a range of different ages.

```
# Load simulated data as example
data(simu_data)
data(pop_ref)

# Estimate remaining life expectancy and Life Years
# Lost due to specific causes of death after age 45
# years and before age 95 years for those with a disease
diseased <- simu_data[!is.na(simu_data$age_disease), ]</pre>
```

lyl\_range 19

```
lyl_estimation1 <- lyl(data = diseased, t0 = age_disease,
                       t = age_death, status = cause_death,
                       age_specific = 45, tau = 95)
lyl_diff_ref(lyl_estimation1, pop_ref, age = age, surv = survival)
lyl_diff_ref(lyl_estimation1, pop_ref, age = age, rates = mortality_rates)
# Calculate bootstrapped confidence interval (3 iterations to test; more are necessary)
lyl_estimation1_ci <- lyl_ci(lyl_estimation1, niter = 3)</pre>
lyl_diff_ref(lyl_estimation1_ci, pop_ref, age = age, surv = survival)
# Estimate remaining life expectancy and Life Years
# Lost due to specific causes of death after each age
# from 0 to 94 years and before age 95 years
lyl_estimation2 <- lyl_range(data = diseased, t0 = age_disease,</pre>
                             t = age_death, status = cause_death,
                             age\_begin = 0, age\_end = 94, tau = 95)
lyl_diff_ref(lyl_estimation2, pop_ref, age = age, surv = survival)
lyl_diff_ref(lyl_estimation2, pop_ref, age = age, surv = survival, weights = diseased$age_disease)
```

lyl\_range

Life Years Lost at a range of different ages.

## Description

lyl estimates remaining life expectancy and Life Years Lost for a given population after a range of specific ages (age\_begin to age\_end) and restricted to a maximum theoretical age  $\tau$ .

#### Usage

```
lyl_range(
  data,
  t0 = NULL,
  t,
  status,
  age_begin,
  age_end,
  censoring_label = "Alive",
  death_labels = "Dead",
  tau = 100
)
```

## **Arguments**

data

A dataframe, where each raw represents a person. The dataframe will have a time-to-event format with at least two variables: age at end of follow-up (t) and status indicator with death/censoring (status). Note that this package is not developed to be used with tibbles.

20 lyl\_range

Age at start of the follow-up time. Default is NULL, which means all subjects are followed from birth. For delayed entry, t0 indicates beginning of follow-up.

t Age at the end of the follow-up time (death or censoring).

status Status indicator, normally 0=alive, 1=dead. Other choices are TRUE/FALSE

(TRUE = death) or 1/2 (2=death). For multiple causes of death (competing risks analysis), the status variable will be a factor, whose first level is treated as censoring; or a numeric variable, whose lowest level is treated as censoring. In the latter case, the label for censoring is censoring\_label ("Alive" by default).

age\_begin Specific starting age at which the Life Years Lost have to be estimated.

age\_end Specific ending age at which the Life Years Lost have to be estimated.

censoring\_label

Label for censoring status. If status is not a factor, "Alive" by default. If

status is a factor, the first level will be treated as censoring label.

death\_labels Label for event status. For only one cause of death, "Dead" is the default. For

multiple causes, the default are the values given in variable status.

tau Remaining life expectancy and Life Years Lost are estimated restricted to a

maximum theoretical age  $\tau$  ( $\tau$ =100 years by default).

#### Value

A list with class "lyl\_range" containing the following components:

- data: Data frame with 3 variables and as many observations as the original data provided to estimate Life Years Lost: t0, t, and status
- LYL: Data frame with (age\_end age\_begin + 1) observations and at least 3 variables: age which corresponds to each specific age from age\_begin to age\_end; life\_exp which is the estimated remaining life expectancy at age specific age and before age tau years; and one variable corresponding to the estimated Life Years Lost for each specific cause of death. If only one cause of death is considered (no competing risks), this variable is Dead and includes the total overall Life Years Lost
- tau: Maximum theoretical age au
- age\_begin: Specific starting age at which the Life Years Lost have been estimated
- age\_end: Specific ending age at which the Life Years Lost have been estimated
- censoring\_label: Label for censoring status
- death\_labels: Label(s) for death status
- competing\_risks: Logical value (TRUE = more than one cause of death (competing risks))
- numbers\_at\_risk: Data frame with (tau age\_begin + 1) observations and 2 variables: age which corresponds to each specific age from age\_begin to tau; and number which is the number of persons at risk of dying at each specific age
- type: Whether the estimation is at "age\_specific" or "age\_range".

lyl\_range 21

#### References

• Andersen PK. Life years lost among patients with a given disease. *Statistics in Medicine*. 2017;36(22):3573-3582.

- Andersen PK. Decomposition of number of life years lost according to causes of death. *Statistics in Medicine*. 2013;32(30):5278-5285.
- Plana-Ripoll et al. lillies An R package for the estimation of excess Life Years Lost among patients with a given disease or condition. *PLoS ONE*. 2020;15(3):e0228073.

#### See Also

- 1yl for estimation of Life Years Lost at one specific age.
- lyl\_diff to compare average Life Years Lost for two populations.
- lyl\_checkplot to check whether small numbers could compromise the estimation.
- lyl\_ci to estimate bootstrapped confidence intervals.
- summary.lyl\_range to summarize objects obtained with function lyl\_range.
- plot.lyl\_range to plot objects obtained with function lyl\_range.

```
# Load simulated data as example
data(simu_data)
# Estimate remaining life expectancy and Life Years
# Lost after each age from 0 to 94 years and before age 95 years
lyl_estimation \leftarrow lyl_range(data = simu_data, t = age_death, status = death,
                            age\_begin = 0, age\_end = 94, tau = 95)
# Visualize data at each different specific age
summary(lyl_estimation)
plot(lyl_estimation)
# Summarize data over an age distribution
summary(lyl_estimation, weights = simu_data$age_disease)
# Estimate remaining life expectancy and Life Years
# Lost due to specific causes of death after each age
# from 0 to 94 years and before age 95 years
lyl_estimation2 <- lyl_range(data = simu_data, t = age_death, status = cause_death,</pre>
                             age\_begin = 0, age\_end = 94, tau = 95)
# Visualize data at each different specific age
summary(lyl_estimation2)
plot(lyl_estimation2)
# Summarize data over an age distribution
summary(lyl_estimation2, weights = simu_data$age_disease)
```

22 plot.lyl

plot.lyl

Plot Life Years Lost at one specific age

## **Description**

plot for objects of class lyl creates a figure of Life Years Lost at one specific age.

## Usage

```
## S3 method for class 'lyl'
plot(x, color_alive = NA, colors = NA, reverse_legend = FALSE, ...)
```

## **Arguments**

X	An object of class lyl (obtained with function lyl).
color_alive	Color to be used for the censoring category. Default is NA, and default color is "white".
colors	Vector with one color for each cause of death. Default is NA, and default colors are used.
reverse_legend	Reverse the order of elements in the legend. Ddefault is FALSE, indicating that first is the censoring label and then all causes of death.
	Additional arguments affecting the plot produced.

#### Value

A plot with survival function and stacked cause-specific cumulative incidences.

## References

• Plana-Ripoll et al. lillies – An R package for the estimation of excess Life Years Lost among patients with a given disease or condition. *PLoS ONE*. 2020;15(3):e0228073.

## See Also

• 1yl for estimation of Life Years Lost at one specific age.

plot.lyl\_aggregated 23

```
plot(lyl_estimation)
plot(lyl_estimation, colors = c("chocolate", "cornflowerblue"))
# The plot can be modified with a usual ggplot2 format
plot(lyl_estimation) +
    ggplot2::xlab("Age [in years]") +
    ggplot2::ggtitle("Life Years Lost at age 45 years")
```

plot.lyl\_aggregated

Plot Life Years Lost at one specific age for two different populations obtained from aggregated data

## Description

plot for objects of class lyl\_aggregated creates a figure of Life Years Lost at one specific age for two different populations.

## Usage

```
## $3 method for class 'lyl_aggregated'
plot(
    X,
    color_alive = NA,
    colors = NA,
    labels = c("Population of interest", "Reference population"),
    ...
)
```

## **Arguments**

X	An object of class $lyl_aggregated$ (obtained with function $lyl_aggregated$ ).
color_alive	Color to be used for the censoring category. Default is NA, and default color is "white".
colors	Vector with one color for each cause of death. Default is NA, and default colors are used.
labels	Vector with labels for the two populations (default are "Population of interest" for data, and "Reference population" for data0; which are provided to function lyl_aggregated.)
	Additional arguments affecting the plot produced.

## Value

A plot with survival function and stacked cause-specific cumulative incidences for two populations side by side.

plot.lyl\_ci

## References

• Plana-Ripoll et al. lillies – An R package for the estimation of excess Life Years Lost among patients with a given disease or condition. *PLoS ONE*. 2020;15(3):e0228073.

## See Also

• lyl\_aggregated for estimation of Life Years Lost at one specific age.

## **Examples**

plot.lyl\_ci

Plot evolution of bootstrapped parameters for Life Years Lost

## Description

plot for objects of class lyl\_ci creates a figure of the bootstrapped Life Years Lost to examine if the number of iterations is enough.

#### Usage

```
## S3 method for class 'lyl_ci'
plot(x, level = 0.95, weights, ...)
```

## **Arguments**

X	An object of class lyl_ci (obtained with function lyl_ci).
level	Confidence level (default is 0.95 for 95% confidence intervals)
weights	Vector with age distribution of disease/condition onset to be used when Life Years Lost are estimated over a range of ages (with lyl_range function).
	Additional arguments affecting the plot produced.

## Value

A plot with the evolution of bootstrapped parameters.

plot.lyl\_range 25

## References

• Plana-Ripoll et al. lillies – An R package for the estimation of excess Life Years Lost among patients with a given disease or condition. *PLoS ONE*. 2020;15(3):e0228073.

#### See Also

- lyl\_range for estimation of Life Years Lost for a range of different ages.
- 1yl for estimation of Life Years Lost at one specific age.
- lyl\_ci to estimate bootstrapped confidence intervals.

## **Examples**

```
# Load simulated data as example
data(simu_data)
# Estimate remaining life expectancy and Life Years
# Lost after age 45 years and before age 95 years
lyl_estimation <- lyl(data = simu_data, t = age_death, status = death,</pre>
                      age_specific = 45, tau = 95)
# Calculate bootstrapped confidence interval (10 iterations to test)
lyl_estimation_ci <- lyl_ci(lyl_estimation, niter = 10)</pre>
plot(lyl_estimation_ci)
# Estimate remaining life expectancy and Life Years
# Lost after each age from 0 to 94 years and before age 95 years
lyl_estimation2 <- lyl_range(data = simu_data, t = age_death, status = death,
                             age\_begin = 0, age\_end = 94, tau = 95)
# Calculate bootstrapped confidence interval
lyl_estimation_ci2 <- lyl_ci(lyl_estimation2)</pre>
plot(lyl_estimation_ci2, weights = simu_data$age_disease)
```

plot.lyl\_range

Plot Life Years Lost at a range of different ages

## **Description**

plot for objects of class lyl\_range creates a figure of Life Years Lost at a range of different ages.

## Usage

```
## S3 method for class 'lyl_range'
plot(x, colors = NA, ...)
```

pop\_ref

## **Arguments**

X	An object of class lyl_range (obtained with function lyl_range).
colors	Vector with one color for each cause of death. Default is NA, and default colors are used.
	Additional arguments affecting the plot produced.

#### Value

A plot with age-specific life expectancy and life years lost.

## References

• Plana-Ripoll et al. lillies – An R package for the estimation of excess Life Years Lost among patients with a given disease or condition. *PLoS ONE*. 2020;15(3):e0228073.

#### See Also

• lyl\_range for estimation of Life Years Lost for a range of different ages.

## **Examples**

pop\_ref

Aggregated data for Life Years Lost estimation.

## **Description**

A dataset containing age-specific survival probability and mortality rates for Danish women in years 2017-2018.

simu\_data 27

## Usage

```
pop_ref
```

#### **Format**

A data frame with 100 rows and 3 variables:

```
age age
survival survival probability at that specific age
mortality_rates age-specific mortality rates
```

#### **Source**

Statistics Danmark (https://www.dst.dk/en/Statistik/emner/befolkning-og-valg/)

simu\_data

Simulated population for Life Years Lost estimation.

## Description

A dataset containing age and cause of death, as well as age at disease diagnosis (or start of a condition) for 100,000 simulated persons.

## Usage

simu\_data

#### **Format**

A data frame with 100000 rows and 6 variables:

```
id unique identifier of each person
```

age\_start age at start of follow-up (0 for all individuals)

age\_death age at end of follow-up (death or censoring)

**death** logical variable (TRUE = death / FALSE = censoring)

cause\_death factor variable with 3 levels: "Alive" (for those censored) and "Natural" and
"Unnatural" (for those dying of natural and unnatural causes of death, respectively)

**age\_disease** age at developing a specific disease or condition for those 32,391 individuals that develop the disease (missing for the remaining 67,609)

#### Source

Simulated data

28 summary.lyl

summary.lyl

Summarize Life Years Lost at one specific age

## **Description**

summary for objects of class 1y1 summarizes Life Years Lost at one specific age.

## Usage

```
## S3 method for class 'lyl'
summary(object, decimals = 2, difference = FALSE, ...)
```

## **Arguments**

object An object of class 1y1 (obtained with function 1y1).

decimals Number of decimals to be reported (default is 2).

difference Parameter automatically created by the package.

... Additional arguments affecting the summary produced.

#### Value

A table with the summary of the results.

#### References

• Plana-Ripoll et al. lillies – An R package for the estimation of excess Life Years Lost among patients with a given disease or condition. *PLoS ONE*. 2020;15(3):e0228073.

#### See Also

• 1yl for estimation of Life Years Lost at one specific age.

```
summary.lyl_aggregated
```

Summarize Life Years Lost at one specific age

## Description

summary for objects of class lyl\_aggregated summarizes Life Years Lost.

## Usage

```
## S3 method for class 'lyl_aggregated'
summary(object, decimals = 2, ...)
```

## **Arguments**

object An object of class lyl\_aggregated (obtained with function lyl\_aggregated or lyl\_aggregated\_range).

decimals Number of decimals to be reported (default is 2).

Additional arguments affecting the summary produced.

#### Value

A table with the summary of the results.

#### References

• Plana-Ripoll et al. lillies – An R package for the estimation of excess Life Years Lost among patients with a given disease or condition. *PLoS ONE*. 2020;15(3):e0228073.

## See Also

- lyl\_aggregated for estimation of Life Years Lost at one specific age.
- lyl\_aggregated\_range for estimation of Life Years Lost for a range of different ages.

30 summary.lyl\_ci

summary.lyl\_ci

Summarize Life Years Lost with confidence intervals

## Description

summary for objects of class lyl\_ci summarizes Life Years Lost at one specific age or over a range of different ages, including bootstrapped confidence intervals

## Usage

```
## $3 method for class 'lyl_ci'
summary(
  object,
  decimals = 2,
  level = 0.95,
  weights = NA,
  difference = FALSE,
   ...
)
```

## Arguments

object An object of class lyl\_ci (obtained with function lyl\_ci).

decimals Number of decimals to be reported (default is 2).

level Confidence level (default is 0.95 for 95% confidence intervals)

weights Vector with age distribution of disease/condition onset to be used when Life

Years Lost are estimated over a range of ages (with lyl\_range function). If weights are not provided (dafault is weights = NA), then the differences in Life Years Lost at each age is provided. If weights are provided, then a weighted

average is provided.

difference Parameter automatically created by the package.

... Additional arguments affecting the summary produced.

## Value

A table with the summary of the results.

#### References

• Plana-Ripoll et al. lillies – An R package for the estimation of excess Life Years Lost among patients with a given disease or condition. *PLoS ONE*. 2020;15(3):e0228073.

#### See Also

- 1y1 for estimation of Life Years Lost at one specific age.
- lyl\_range for estimation of Life Years Lost for a range of different ages.
- lyl\_ci to estimate bootstrapped confidence intervals.

summary.lyl\_range 31

## **Examples**

```
# Load simulated data as example
data(simu_data)
# Estimate remaining life expectancy and Life Years
# Lost after age 45 years and before age 95 years
lyl_estimation <- lyl(data = simu_data, t = age_death, status = death,</pre>
                       age_specific = 45, tau = 95)
# Calculate bootstrapped confidence interval (3 iterations to test; more are necessary)
lyl_estimation_ci <- lyl_ci(lyl_estimation, niter = 3)</pre>
summary(lyl_estimation_ci)
# Estimate remaining life expectancy and Life Years
# Lost after each age from 0 to 94 years and before age 95 years
lyl_estimation2 <- lyl_range(data = simu_data, t = age_death, status = death,</pre>
                              age\_begin = 0, age\_end = 94, tau = 95)
# Calculate bootstrapped confidence interval (3 iterations to test; more are necessary)
lyl_estimation_ci2 <- lyl_ci(lyl_estimation2, niter = 3)</pre>
summary(lyl_estimation_ci2, weights = simu_data$age_disease)
```

summary.lyl\_range

Summarize Life Years Lost over a range of differents ages

## **Description**

summary for objects of class lyl\_range summarizes Life Years Lost over a range of different ages.

## Usage

```
## S3 method for class 'lyl_range'
summary(object, decimals = 2, weights = NA, difference = FALSE, ...)
```

## Arguments

object	An object of class lyl_range (obtained with function lyl_range).
decimals	Number of decimals to be reported (default is 2).
weights	Vector with age distribution of disease/condition onset. If weights are not provided (dafault is weights = NA), then the differences in Life Years Lost at each age are summarized. If weights are provided, then a weighted average is provided.
difference	Parameter automatically created by the package.
	Additional arguments affecting the summary produced.

32 summary.lyl\_range

## Value

A table with the summary of the results.

#### References

• Plana-Ripoll et al. lillies – An R package for the estimation of excess Life Years Lost among patients with a given disease or condition. *PLoS ONE*. 2020;15(3):e0228073.

#### See Also

• lyl\_range for estimation of Life Years Lost for a range of different ages.

## **Index**

```
* datasets
    aggreg_data, 2
    pop_ref, 26
    simu_data, 27
aggreg_data, 2
lyl, 3, 6, 13, 14, 16, 18, 21, 22, 25, 28, 30
lyl_2plot, 5
lyl_aggregated, 6, 10, 24, 29
lyl_aggregated_range, 8, 8, 29
lyl_checkplot, 11, 21
lyl_ci, 4, 12, 21, 25, 30
lyl\_compare\_plot, 13
lyl_diff, 4, 6, 13, 14, 15, 21
lyl_diff_ref, 17
lyl_range, 4, 11, 13, 16, 18, 19, 25, 26, 30, 32
plot, 4, 8, 10
plot.lyl, 4, 22
plot.lyl_aggregated, 8, 23
plot.lyl_ci, 13, 24
plot.lyl_range, 21, 25
pop_ref, 26
simu_data, 27
summary.lyl, 4, 28
summary.lyl_aggregated, 8, 10, 29
summary.lyl_ci, 12, 13, 30
summary.lyl_range, 21, 31
```