Package 'OneSampleLogRankTest'

January 20, 2025

Title One-Sample Log-Rank Test

Version 0.9.2

Imports magrittr, dplyr, ggplot2, survival, survminer, rlang

Description The log-rank test is performed to assess the survival outcomes between two group.
When there is no proper control group or obtaining such data is cumbersome, one sample
log-rank test can be applied. This package performs one sample log-
rank test as described in Finkelstein et al. (2003)<doi:10.1093/jnci/djt227> and variation of the
test for small sample sizes which is detailed in FD Liddell (1984)<doi:10.1136/jech.38.1.85> pa-
per. Visualization function in the package
generates Kaplan-Meier Curve comparing survival curve of the general popula-
tion against that of the population of interest.

License GPL (>= 3)

Encoding UTF-8

LazyData true

RoxygenNote 7.1.2

Suggests knitr, rmarkdown

VignetteBuilder knitr

NeedsCompilation no

Author Divy Kangeyan [aut, cre],
Jin Xie [aut]

Maintainer Divy Kangeyan <dkangeyan@kitepharma.com>

Depends R (>= 3.5.0)

Repository CRAN

Date/Publication 2024-02-03 12:30:15 UTC

Contents

Index   **8**

---

dataPop_1999_2020   *Population data (1999-2020)*

Description

This data set contains mortality rate for various demographic groups by age. all column has overall mortality rate for all groups. Race groups are indicated by the following notation: w - white, b - black, ai - american indian and as - asian. Female and Males are indicated by an additional suffix _f and _m.

Usage

    dataPop_1999_2020

Format

A dataframe with 16 columns and 151 rows.

Source

CDC Wonder Database. Data Colleceted from 1999 - 2020

References

https://wonder.cdc.gov/

---

dataPop_2018_2021   *Population data (2018_2021)*

Description

This data set contains mortality rate for various demographic groups by age. all column has overall mortality rate for all groups. Race groups are indicated by the following notation: w - white, b - black, ai - american indian and as - asian, nh - native hawaiian. Female and Males are indicated by an additional suffix _f and _m.

Usage

    dataPop_2018_2021

Format

A dataframe with 19 columns and 151 rows.

Source

CDC Wonder Database. Data Colleceted from 2018_2021

References

https://wonder.cdc.gov/

---

dataPop_2018_2021_race_sex_eth

Population data (2018_2021) by race, sex, ethnicity

Description

This data set contains mortality rate for various demographic groups by age. all column has overall mortality rate for all groups. Race groups are indicated by the following notation: w - white, b - black, ai - american indian and as - asian, nh - native hawaiian. Female and Males are indicated by an additional suffix _f and _m.

Usage

dataPop_2018_2021

Format

A dataframe with 16 columns and 151 rows.

Source

CDC Wonder Database. Data Colleceted from 2018_2021

References

https://wonder.cdc.gov/

---

dataSurv   *Survival data*

Description

This data set is obtained from Finkelstein et al. paper that contains the following five columns: age, time, event status, sex and race.

Usage

    dataSurv

Format

A dataframe with 5 columns and 33 rows.

Source

Finkelstein et al. (2003)

References

Finkelstein, D. M., Muzikansky, A., & Schoenfeld, D. A. (2003). Comparing survival of a sample to that of a standard population. Journal of the National Cancer Institute, 95(19), 1434-1439.

---

dataSurv_small   *Survival data*

Description

This data set is subset of data obtained from Finkelstein et al. paper that contains the following five columns: age, time, event status, sex and race. In order to apply the exact test 12 patients were randomly selected out of 33 patients.

Usage

    dataSurv

Format

A dataframe with 5 columns and 12 rows.

Source

Finkelstein et al. (2003)

References

Finkelstein, D. M., Muzikansky, A., & Schoenfeld, D. A. (2003). Comparing survival of a sample to that of a standard population. Journal of the National Cancer Institute, 95(19), 1434-1439.

---

findMatchedCumuSurvProb

Find Matched Cumulative Survival Probability

Description

Find Matched Cumulative Survival Probability

Usage

```
findMatchedCumuSurvProb(time, ageDiag, sex, race, dataPop, maxFollowUp = NULL)
```

Arguments

time	follow up length
ageDiag	age at diagnosis
sex	sex
race	race
dataPop	Population level mortality data
maxFollowUp	maximum follow-up, if max follow-up not provided then the time would be considered until death or censoring

Value

matched survival probability

Examples

```
# load data
data(dataSurv_small)
data(dataPop_2018_2021)

# Extract info for the first subject
time_vec <- dataSurv_small$time[1]
age_vec <- dataSurv_small$age[1]
sex_vec <- dataSurv_small$sex[1]
race_vec <- dataSurv_small$race[1]

# Generate cumulative survival probability
findMatchedCumuSurvProb(time = time_vec, ageDiag = age_vec, sex = sex_vec,
race = race_vec, dataPop = dataPop_2018_2021)

#If maximum followup is determined to be 20 years
```

```
findMatchedCumuSurvProb(time = time_vec, ageDiag = age_vec, sex = sex_vec,
race = race_vec, dataPop = dataPop_2018_2021, maxFollowUp = 20)
```

---

oneSampleLogRankTest   *Calculate One-Sample Log-Rank Test*

Description

Calculate One-Sample Log-Rank Test

Usage

```
oneSampleLogRankTest(dataSurv, dataPop, type = c("exact", "approximate"))
```

Arguments

dataSurv	Survival data
dataPop	Population data
type	Type of test

Value

p-value for one-sample log-rank test

Examples

```
# load data
data(dataSurv_small)
data(dataPop_2018_2021)

# Since the dataset is small run an exact test
oneSampleLogRankTest(dataSurv_small, dataPop_2018_2021, type = "exact")
```

---

plotKM   *Plot Kaplan-Meier Curve against Population*

Description

Plot Kaplan-Meier Curve against Population

Usage

```
plotKM(dataSurv, dataPop, type = c("exact", "approximate"))
```

Arguments

`dataSurv`	Survival data
`dataPop`	Population data
`type`	Type of test to conduct in order to display p-value

Value

ggplot object

Examples

```
# load data
data(dataSurv_small)
data(dataPop_2018_2021)

plotKM(dataSurv_small, dataPop_2018_2021, type = "exact")
```

```
simulated_clinical_data
```
Simultaed Clinical data

Description

This data is simulated data from clinical trial data that contains five columns: race, sex, age, event status and time in years.

Usage

```
simulated_clinical_data
```

Format

A dataframe with 5 columns and 500 rows.

Source

Simulated

References

None

Index