

Package ‘dsr’

August 23, 2019

Type Package

Title Compute Directly Standardized Rates, Ratios and Differences

Version 0.2.2

Author Matthew Kumar <mattkumar@gmail.com>

Maintainer Matthew Kumar <mattkumar@gmail.com>

URL <https://github.com/mattkumar/dsr>

Description A set of functions to compute and compare directly standardized rates, rate differences and ratios. A variety of user defined options for analysis (e.g confidence intervals) and formatting are included.

Imports stats, dplyr (>= 0.7.8) , rlang, utils, frailtypack

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 6.0.1

Suggests knitr, rmarkdown

VignetteBuilder knitr

NeedsCompilation no

Repository CRAN

Date/Publication 2019-08-23 15:10:02 UTC

R topics documented:

dsr	2
dsrr	3
dsrrec	5

Index	7
--------------	----------

dsr

*Compute Directly Standardized Rates***Description**

Computes crude and directly standardized rates by subgroup with confidence intervals.

Usage

```
dsr(data, event, fu, subgroup, ..., refdata, mp, method = "normal",
    sig = 0.95, decimals)
```

Arguments

data	A data frame with counts and unit-times summarized by the standardization variables.
event	A variable within the input data that corresponds to the event counts.
fu	A variable within the input data that corresponds to the unit-time.
subgroup	A variable within the input data frame for which rates are calculated by.
...	Variables(s) within the input data that for which rates are to be standardized by. The input data and ref data should both be summarized by these.
refdata	A data frame with population unit-times summarized by the standardization variables. The unit-time variable name must named pop.
mp	A constant to multiply rates by (e.g. mp=1000 for rates per 1000).
method	Choose between normal, lognormal and gamma confidence intervals for crude and standardized rates. The default method is normal.
sig	The desired level of confidence in computing confidence intervals. The default is 0.95 for 95 percent CIs.
decimals	Round estimates to a desired decimal place.

References

- Fay, M.P., & Feuer, E.J. (1997). Confidence intervals for directly standardized rates: a method based on the gamma distribution. *Statistics in Medicine*,16, 791-801.
- Elandt-Johnson, R. C., and Johnson, N. L. (1980). *Survival Models and Data Analysis*. New York: John Wiley & Sons.
- Chiang C. Standard error of the age-adjusted death rate. *US Department of Health, Education and Welfare: Vital Statistics Special Reports* 1961;47:271-285.
- Schoenbach, V., and Rosamond W. (2000) *Understanding the fundamentals of epidemiology: An evolving text*.

Examples

```

#An example of calculating directly standardized rates
#Data from Table 1, Page 132 of Schoenbach (2000)

#State specific death counts and fu
df_study <- data.frame(state=rep(c('Miami','Alaska'), c(5,5)),
                      age=rep(c('00-14','15-24','25-44','45-64','65+'),2),
                      deaths=c(136,57,208,1016,3605,59,18,37,90,81),
                      fu=c(114350,80259,133440,142670,92168,37164,20036,32693,14947,2077))

#US standard population
df_ref <- data.frame(age=c('00-14','15-24','25-44','45-64','65+'),
                    pop=c(23961000,15420000,21353000,19601000,10685000))

#Directly Standardized Rates (per 1000) - 95% CI's using the gamma method
my_results <- dsr(data=df_study,
                 event=deaths,
                 fu=fu,
                 subgroup=state,
                 age,
                 refdata=df_ref,
                 method="gamma",
                 sig=0.95,
                 mp=1000,
                 decimals=4)

#View results
my_results

```

dsrr

*Compare Directly Standardized Rates by Ratios or Differences.***Description**

Compare directly standardized rates by ratios or differences.

Usage

```
dsrr(data, event, fu, subgroup, ..., refdata, estimate, refgroup, mp,
     sig = 0.95, decimals)
```

Arguments

data	A data frame with counts and unit-times summarized by the standardization variables.
event	A variable within the input data that corresponds to the event counts.
fu	A variable within the input data that corresponds to the unit-time.
subgroup	A variable within the input data frame for which rates are calculated by.

...	Variables(s) within the input data that for which rates are to be standardized by. The input data and ref data should both be summarized by these.
refdata	A data frame with population unit-times summarized by the standardization variables. The unit-time variable name must named pop.
estimate	Choose between difference or ratio in comparing directly standardized rates.
refgroup	A level of the subgroup variable taken to be the reference in computing rate ratios or differences.
mp	A constant to multiply rates by (e.g. mp=1000 for rates per 1000).
sig	The desired level of confidence in computing confidence intervals. The default is 0.95 for 95 percent CIs.
decimals	Round estimates to a desired decimal place.

References

Fay, M.P., & Feuer, E.J. (1997). Confidence intervals for directly standardized rates: a method based on the gamma distribution. *Statistics in Medicine*, 16, 791-801.

Elandt-Johnson, R. C., and Johnson, N. L. (1980). *Survival Models and Data Analysis*. New York: John Wiley & Sons.

Chiang C. Standard error of the age-adjusted death rate. *US Department of Health, Education and Welfare: Vital Statistics Special Reports 1961*;47:271-285.

Schoenbach, V., and Rosamond W. (2000) *Understanding the fundamentals of epidemiology: An evolving text*.

Examples

```
#An example of comparing directly standardized rates
#Data from Table 1, Page 132 of Schoenbach (2000)

#State specific death counts and fu
df_study <- data.frame(state=rep(c('Miami','Alaska'), c(5,5)),
                      age=rep(c('00-14','15-24','25-44','45-64','65+'),2),
                      deaths=c(136,57,208,1016,3605,59,18,37,90,81),
                      fu=c(114350,80259,133440,142670,92168,37164,20036,32693,14947,2077))

#US standard population
df_ref <- data.frame(age=c('00-14','15-24','25-44','45-64','65+'),
                    pop=c(23961000,15420000,21353000,19601000,10685000))

#Directly Standardized Rate Ratio (per 1000) - 95% log-normal CI's, Alaska as the reference
my_results2 <- dsrr(data=df_study,
                   event=deaths,
                   fu=fu,
                   subgroup=state,
                   age,
                   refdata=df_ref,
                   refgroup="Alaska",
                   estimate="ratio",
                   sig=0.95,
```

```

                                mp=1000,
                                decimals=4)
#View results
my_results2

```

dsrrec

Compute Directly Standardized Rates for Recurrent Events

Description

Computes directly standardized rates for recurrent events by subgroup with confidence intervals.

Usage

```
dsrrec(data, event, fu, subgroup, ..., refdata, sig = 0.95, mp, decimals)
```

Arguments

data	A data frame with counts and unit-times summarized by the standardization variables.
event	A variable within the input data that corresponds to the event counts.
fu	A variable within the input data that corresponds to the unit-time.
subgroup	A variable within the input data frame for which rates are calculated by.
...	Variables(s) within the input data that for which rates are to be standardized by. The input data and ref data should both be summarized by these.
refdata	A data frame with population unit-times summarized by the standardization variables. The unit-time variable name must named pop.
sig	The desired level of confidence in computing confidence intervals. The default is 0.95 for 95 percent CIs.
mp	A constant to multiply rates by (e.g. mp=1000 for rates per 1000).
decimals	Round estimates to a desired decimal place.

References

Stukel, T. A., Glynn, R. J., Fisher, E. S., Sharp, S. M., Lu-Yao, G and Wennberg, J. E. (1994). Standardized rates of recurrent outcomes. *Statistics in Medicine*, 13, 1781-1791.

Fay, M.P., & Feuer, E.J. (1997). Confidence intervals for directly standardized rates: a method based on the gamma distribution. *Statistics in Medicine*, 16, 791-801.

Examples

```
#An example of directly standardized rates for recurrent events

library(frailtypack)
library(dplyr)
library(dsr)
data(readmission)

#Make an individual level dataset with total event counts and total observation times
treadm <- as.data.frame(readmission %>%
  group_by(id) %>%
  filter(max(enum)==enum ) %>%
  mutate(events=enum-1, time=t.stop) %>%
  select(id, events, time, sex, dukes))

#Make the standard pop
tref <- as.data.frame(treadm %>%
  group_by(sex) %>%
  mutate(pop=sum(time)) %>%
  select(sex, pop) %>%
  distinct(sex, pop))

#Get directly standardized rates (age-adjusted) for readmissions by Dukes' tumor grade.
analysis <- dsrrec(data=treadm,
  event=events,
  fu=time,
  refdata=tref,
  subgroup=dukes,
  sex,
  mp=1000,
  decimals=3)
```

Index

dsr, 2
dsrr, 3
dsrrec, 5