Package 'outcomerate'

October 14, 2022

Version 1.0.1

Title AAPOR Survey Outcome Rates

Description Standardized survey outcome rate functions, including the response rate, contact rate, cooperation rate, and refusal rate. These outcome rates allow survey researchers to measure the quality of survey data using definitions published by the American Association of Public Opinion Research (AAPOR). For details on these standards, see AAPOR (2016) <https://www.aapor.org/Standards-Ethics/Standard-Definitions-(1).aspx>.

License CC0

Encoding UTF-8

LazyData true

ByteCompile true

RoxygenNote 6.1.0

Depends R (>= 2.10)

Suggests dplyr, forcats, ggplot2, knitr, testthat, tidyr, covr, rmarkdown, spelling, stringr

Imports Rdpack (>= 0.7), stats

RdMacros Rdpack

Language en-US

URL https://github.com/ropensci/outcomerate

BugReports https://github.com/ropensci/outcomerate/issues

VignetteBuilder knitr

NeedsCompilation no

Author Rafael Pilliard Hellwig [aut, cre]
(<https://orcid.org/0000-0002-3092-3493>),
Carl Ganz [rev],
Neal Richardson [rev]

Maintainer Rafael Pilliard Hellwig <rafael.taph@gmail.com>

Repository CRAN

Date/Publication 2018-10-06 21:20:03 UTC

R topics documented:

---

eligibility_rate   *Survey Eligibility Rate*

Description

Provides an estimate for the proportion of cases of unknown eligibility that are eligible, as described by (Valliant et al. 2013). The rate is typically (but not necessarily) calculated on the screener data or other sources depending on the type of survey, and approaches to calculating 'e' may therefore differ from one survey to the next.

Usage

```
eligibility_rate(x, weight = NULL)
```

Arguments

x	a character vector of disposition outcomes (I, P, R, NC, O, UH, UO, U, or NE). Alternatively, a named vector/table of (weighted) disposition counts.
weight	an optional numeric vector that specifies the weight of each element in 'x' if x is a character vector. If none is provided (the default), an unweighted estimate is returned.

Details

The present implementation follows the default used in the Excel-based *AAPOR Outcome Rate Calculator (Version 4.0, May, 2016)* on the basis of known ineligibles being coded as "NE".

The eligibility rate (ELR) is defined as

- ELR = (I + P + R + NC + O) / (I + P + R + NC + O + NE)

References

The American Association for Public Opinion Research (2016). "Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys." https://www.aapor.org/Standards-Ethics/Standard-Definitions-(1).aspx. Valliant R, Dever J, Kreuter F (2013). *Practical Tools for Designing and Weighting Survey Samples*, series Statistics for Social and Behavioral Sciences. Springer New York.

See Also

outcomerate

Examples

```
# load the outcomerate package
library(outcomerate)

# Create a vector of survey dispositions
#
# I  = Complete interview
# P  = Partial interview
# R  = Refusal and break-off
# NC = Non-contact
# O  = Other
# UH = Unknown if household/occupied housing unit
# UO = Unknown, other
# NE = Not eligible
x <- c("I", "P", "I", "NE", "NC", "UH", "I", "R", "UO", "I", "O", "P", "I")

# calculate all rates, assume 80% of unknown cases are elligble
eligibility_rate(x)

# calculate weighted rates
w <- runif(13, 0, 5)
eligibility_rate(x, weight = w)

# alternatively, provide input as counts
freq <- c(I = 6, P = 2, NC = 3, NE = 1)
eligibility_rate(freq)
```

---

fmat   *outcomerate Formula Matrix (Internal Data)*

Description

The fmat object is the internal dataset used by the outcomerate package. It holds all definitions for the outcome rates. With the exception of location rates, these are taken from the AAPOR Standard Definitions (2016).

Details

The data is a 3-dimensional binary array consisting of:

- outcome: codes I, P, R, NC, O, UH, UO, eUH, eUO, NE
- rate: the shorthand name for the rate (e.g. RR1)
- side: numerator (NUM) and denominator (DEN)

Given these three dimensions, each outcome rate can be defined as a rational number (i.e. a fraction) consisting of a summation of frequencies of outcome codes (where the matrix entries are nonzero).

The input parameters given by the user are I, P, R, NC, O, UH, UO and the parameter 'e'. The parameter e is multiplied by UH, UO internally so as to produce eUH, eUO.

The reason for this implementation is:

a) It conforms to a DRY (don't repeat yourself) philosophy by holding all definitions in one place. These definitions can be used as upstream inputs to functions/test suites requiring them.

b) It makes it easier to use intermediate steps in the formula calculations. For instance, it may be of use to a researchers to want to obtain the numerator/denominators of calculations, instead of only the output.

c) it makes it easy to compare the output

d) It is easier to maintain

References

<https://www.aapor.org/Standards-Ethics/Standard-Definitions-(1).aspx>

Examples

```
fmat <- outcomerate:::fmat

# Print the dimensions
dimnames(fmat)

# Say we want to know the defintion of Response Rate 2, RR2. We see
# below that the numerator (NUM) column is defined by the entries with a 1,
# or (I + P). Likewise, the denominator (DEN) is defined as
# (I + P + R + NC + O + UH + UO)
fmat[, "RR2", ]


# To use linear algebra, we define a zero-one numerator matrix 'N'
# and a zero-one denominator matrix 'D'. Our count of disposition codes
# is given here manually as 'x' (in the same order as N and D).
N = fmat[ , , 1]
D = fmat[ , , 2]
x <- c(I = 5, P = 2, R = 1, NC = 7, O = 3,
       UH = 4, UO = 8,  NE = 1, eUH = 3, eUO = 6)

# Return all rates
(x %*% N) / (x %*% D)


# The same thing can be achieved with the apply family of functions
numden <- apply(x * fmat, 2:3, sum)
numden[, 1] / numden[, 2]
```

---

middleearth	*middleearth Dataset*

Description

`middlearth` is a toy dataset consisting of 1691 fake survey interviews conducted in J.R.R. Tolkien's fictional world of Middle Earth.

Details

Variables contained in the data:

- **code:** one of the outcome codes I, P, R, NC, O, UH, UO, UH, UO, NE
- **outcome:** A human-interpretable label for the code variable
- **researcher**: An identifier for the researcher conducting the interview
- **region**: The region of the respondent (one of five)
- **Q1**: A hypothetical binary research question posed to respondents
- **Q2**: A hypothetical continuous scale question posed to respondents
- **day**: The day the interview took place (1 being the first day of fieldwork)
- **race**: The race of the respondent in middle earth (Dwarf, Elf, Hobbit, Man, or Wizard)
- **svywt**: The survey weight (inverse probability of selection)

---

outcomerate	*AAPOR Survey Outcome Rates*

Description

Provides standardized outcome rates for surveys, primarily as defined by the American Association for Public Opinion Research (AAPOR). Details can be found in the Standard Definitions manual (The American Association for Public Opinion Research 2016).

Usage

```
outcomerate(x, e = NULL, rate = NULL, weight = NULL,
  return_nd = FALSE)
```

Arguments

x	a character vector of disposition outcomes (I, P, R, NC, O, UH, or UO). Alternatively, a named vector/table of (weighted) disposition counts.
e	a scalar number that specifies the eligibility rate (the estimated proportion of unknown cases which are eligible). A default method of calculating 'e' is provided by `eligibility_rate()`.
rate	an optional character vector specifying the rates to be calculated. If set to NA (the default), all rates are returned.
weight	an optional numeric vector that specifies the weight of each element in 'x' if x is a character vector or factor. If none is provided (the default), an unweighted estimate is returned.
return_nd	a logical to switch to having the function return the numerator and denominator instead of the rate. Defaults to FALSE.

Details

Survey and public opinion research often categorizes interview attempts of of a survey according to a set of outcome codes as follows:

- I = Complete interview
- P = Partial interview
- R = Refusal and break-off
- NC = Non-contact
- O = Other
- UH = Unknown if household/occupied housing unit
- UO = Unknown, other
- NE = Known ineligible

These high-level classes are used to calculate outcome rates that provide some measure of quality over the fieldwork. These outcome rates are defined here as follows:

AAPOR Response Rate

The proportion of your intended sample that participate in the survey.

- RR1 = I / ((I + P) + (R + NC + O) + (UH + UO))
- RR2 = (I + P) / ((I + P) + (R + NC + O) + (UH + UO))
- RR3 = I / ((I + P) + (R + NC + O) + e(UH + UO))
- RR4 = (I + P) / ((I + P) + (R + NC + O) + e(UH + UO))
- RR5 = I / ((I + P) + (R + NC + O))
- RR6 = (I + P) / ((I + P) + (R + NC + O))

AAPOR Cooperation Rates

The proportion of contacted respondents who participate in the survey.

- COOP1 = I / ((I + P) + R + O)

- COOP2 = (I + P) / ((I + P) + R + O)
- COOP3 = I / ((I + P) + R)
- COOP4 = (I + P) / ((I + P) + R)

AAPOR Refusal Rates

The proportion of the sample that refuses to participate in the survey.

- REF1 = R / ((I + P) + (R + NC + O) + (UH + UO))
- REF2 = R / ((I + P) + (R + NC + O) + e(UH + UO))
- REF3 = R / ((I + P) + (R + NC + O))

AAPOR Contact Rates

The proportion of the sample that is successfully contacted for an interview (whether they chose to participate or not).

- CON1 = ((I + P) + (R + O)) / ((I + P) + (R + NC + O) + (UH+ UO))
- CON2 = ((I + P) + (R + O)) / ((I + P) + (R + NC + O) + e(UH + UO))
- CON3 = ((I + P) + (R + O)) / ((I + P) + (R + NC + O))

Location Rate

The proportion of cases that could be located for an interview.

The location rate is not defined in AAPOR's Standards, but can be found in (Valliant et al. 2013). Note: depending on how the located cases are encoded, this may or may not be the correct formula.

- LOC1 = ((I + P) + (R + O + NC)) / ((I + P) + (R + NC + O) + (UH + UO))
- LOC2 = ((I + P) + (R + O + NC)) / ((I + P) + (R + NC + O) + e(UH + UO))

References

The American Association for Public Opinion Research (2016). "Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys." https://www.aapor.org/Standards-Ethics/Standard-Definitions-(1).aspx.

Valliant R, Dever J, Kreuter F (2013). *Practical Tools for Designing and Weighting Survey Samples*, series Statistics for Social and Behavioral Sciences. Springer New York. The American Association for Public Opinion Research (2016). "Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys." https://www.aapor.org/Standards-Ethics/Standard-Definitions-(1).aspx.

Examples

```
# load the outcomerate package
library(outcomerate)

# Create a vector of survey dispositions
#
# I  = Complete interview
# P  = Partial interview
# R  = Refusal and break-off
```

```
# NC = Non-contact
# O  = Other
# UH = Unknown if household/occupied housing unit
# UO = Unknown, other
# NE = Known ineligible
x <- c("I", "P", "I", "NC", "UH", "I", "R", "NE",
       "UO", "I", "O", "P", "I")

# calculate all rates
elr <- eligibility_rate(x)
outcomerate(x, e = elr)

# return only one rate
outcomerate(x, rate = "COOP1")

# calculate weighted rates
w <- runif(length(x), 0, 5)
outcomerate(x, e = elr, weight = w)

# alternatively, provide input as counts
freq <- c(I = 6, P = 2, NC = 3, R = 1)
outcomerate(freq, e = elr)
```

Index