Package: healthequal (via r-universe)

September 11, 2024

Title Compute Summary Measures of Health Inequality

Version 1.0.1

Description Compute 21 summary measures of health inequality and its
corresponding confidence intervals for ordered and non-ordered
dimensions using disaggregated data. Measures for ordered
dimensions (e.g., Slope Index of Inequality, Absolute
Concentration Index) also accept individual and survey data.

License AGPL (>= 3)

URL https://github.com/WHOequity/healthequal,
https://whoequity.github.io/healthequal/

BugReports https://github.com/WHOequity/healthequal/issues

Depends R (>= 3.5.0)

Imports dplyr, emmeans, marginaleffects, rlang, srvyr, survey

Suggests bookdown, knitr, rmarkdown, sandwich, spelling, testthat (>=
3.0.0)

VignetteBuilder knitr

Config/testthat/edition 3

Encoding UTF-8

Language en-US

LazyData true

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.1

Repository https://whoequity.r-universe.dev

RemoteUrl https://github.com/whoequity/healthequal

RemoteRef HEAD

RemoteSha 0414ffe5dce6bf7cd56b50d93f3310db40855490

Contents

---

aci   *Absolute concentration index (ACI)*

Description

The absolute concentration index (ACI) is an absolute measure of inequality that indicates the extent to which an indicator is concentrated among disadvantaged or advantaged subgroups, on an absolute scale.

Usage

```
aci(
  est,
  subgroup_order,
  pop = NULL,
  weight = NULL,
  psu = NULL,
```

```
      strata = NULL,
      fpc = NULL,
      lmin = NULL,
      lmax = NULL,
      conf.level = 0.95,
      force = FALSE,
      ...
  )
```

Arguments

est	The indicator estimate. Estimates must be available for all subgroups/individuals (unless force=TRUE).
subgroup_order	The order of subgroups/individuals in an increasing sequence.
pop	The number of people within each subgroup (for disaggregated data). Population size must be available for all subgroups.
weight	Individual sampling weight (required if data come from a survey).
psu	Primary sampling unit (required if data come from a survey)
strata	Strata (required if data come from a survey)
fpc	Finite population correction (if data come from a survey and sample size is large relative to population size).
lmin	Minimum limit for bounded indicators (i.e., variables that have a finite upper and/or lower limit).
lmax	Maximum limit for bounded indicators (i.e., variables that have a finite upper and/or lower limit).
conf.level	Confidence level of the interval. Default is 0.95 (95%).
force	TRUE/FALSE statement to force calculation with missing indicator estimate values.
...	Further arguments passed to or from other methods.

Details

ACI can be calculated using disaggregated data and individual-level data. Subgroups in disaggregated data are weighted according to their population share, while individuals are weighted by sample weight in the case of data from surveys.

The calculation of ACI is based on a ranking of the whole population from the most-disadvantaged subgroup (at rank 0) to the most-advantaged subgroup (at rank 1), which is inferred from the ranking and size of the subgroups. ACI can be calculated as twice the covariance between the health indicator and the relative rank. Given the relationship between covariance and ordinary least squares regression, ACI can be obtained from a regression of a transformation of the health variable of interest on the relative rank. For more information on this inequality measure see Schlotheuber, A., & Hosseinpoor, A. R. (2022) below.

Interpretation: The larger the absolute value of ACI, the higher the level of inequality. For favourable indicators, positive values indicate a concentration of the indicator among the advantaged, while negative values indicate a concentration of the indicator among the disadvantaged.

For adverse indicators, it is the reverse: positive values indicate a concentration of the indicator among the disadvantaged, while negative values indicate a concentration of the indicator among the advantaged. ACI is zero if there is no inequality.

Type of summary measure: Complex; absolute; weighted

Applicability: Ordered; more than two subgroups

Warning: The confidence intervals are approximate and might be biased.

Value

The estimated ACI value, corresponding estimated standard error, and confidence interval as a `data.frame`.

References

Schlotheuber, A., & Hosseinpoor, A. R. (2022). Summary measures of health inequality: A review of existing measures and their application. International Journal of Environmental Research and Public Health, 19 (6), 3697.

Examples

```
# example code
data(IndividualSample)
head(IndividualSample)
with(IndividualSample,
     aci(est = sba,
         subgroup_order = subgroup_order,
         weight = weight,
         psu = psu,
         strata = strata
         )
     )
```

---

bgsd   *Between-group standard deviation (BGSD)*

Description

Between-Group Standard Deviation (BGSD) is an absolute measure of inequality that considers all population subgroups. Subgroups are weighted according to their population share.

Usage

```
bgsd(pop, est, se = NULL, scaleval, sim = NULL, seed = 123456, ...)
```

Arguments

`pop`	The number of people within each subgroup. Population size must be available for all subgroups.
`est`	The subgroup estimate. Estimates must be available for at least 85% of subgroups.
`se`	The standard error of the subgroup estimate. If this is missing, 95% confidence intervals of BGSD cannot be calculated.
`scaleval`	The scale of the indicator. For example, the scale of an indicator measured as a percentage is 100. The scale of an indicator measured as a rate per 1000 population is 1000.
`sim`	The number of simulations to estimate confidence intervals. Default is 100.
`seed`	The random number generator (RNG) state for the confidence interval simulation. Default is 123456.
`...`	Further arguments passed to or from other methods.

Details

BGSD is calculated as the square root of the weighted average of squared differences between the subgroup estimates and the setting average. Squared differences are weighted by each subgroup's population share. For more information on this inequality measure see Schlotheuber, A., & Hosseinpoor, A. R. (2022) below.

95% confidence intervals are calculated using a methodology of simulated estimates. The dataset is simulated a large number of times (e.g., 100) and BGSD is calculated for each of the simulated samples. The 95% confidence intervals are based on the 2.5th and 97.5th percentiles of the BGSD results.

Interpretation: BGSD has only positive values, with larger values indicating higher levels of inequality. BGSD is zero if there is no inequality. It has the same unit as the health indicator.

Type of summary measure: Complex; absolute; weighted

Applicability: Non-ordered; more than two subgroups

Value

The estimated BGSD value, corresponding estimated standard error, and confidence interval as a `data.frame`.

References

Schlotheuber, A., & Hosseinpoor, A. R. (2022). Summary measures of health inequality: A review of existing measures and their application. International Journal of Environmental Research and Public Health, 19 (6), 3697.

Examples

```
# example code
data(NonorderedSample)
head(NonorderedSample)
```

```
with(NonorderedSample,
     bgsd(pop = population,
          est = estimate,
          se = se,
          scaleval = indicator_scale
         )
    )
```

---

bgv   *Between-group variance (BGV)*

Description

Between-Group Variance (BGV) is an absolute measure of inequality that considers all population subgroups. Subgroups are weighted according to their population share.

Usage

```
bgv(pop, est, se = NULL, conf.level = 0.95, ...)
```

Arguments

pop	The number of people within each subgroup. Population size must be available for all subgroups.
est	The subgroup estimate. Estimates must be available at least 85% of subgroups.
se	The standard error of the subgroup estimate. If this is missing, confidence intervals of BGV cannot be calculated.
conf.level	Confidence level of the interval. Default is 0.95 (95%).
...	Further arguments passed to or from other methods.

Details

BGV is calculated as the weighted average of squared differences between the subgroup estimates and the setting average. Squared differences are weighted by each subgroup's population share. For more information on this inequality measure see Schlotheuber, A., & Hosseinpoor, A. R. (2022) below.

Interpretation: BGV has only positive values, with larger values indicating higher levels of inequality. BGV is zero if there is no inequality. BGV is more sensitive to outlier estimates as it gives more weight to the estimates that are further from the setting average. It is reported as the squared unit of the health indicator.

Type of summary measure: Complex; absolute; weighted

Applicability: Non-ordered; more than two subgroups

Warning: The confidence intervals are approximate and might be biased. See Ahn J. et al. (2018) below for further information on the standard error formula.

Value

The estimated BGV value, corresponding estimated standard error, and confidence interval as a `data.frame`.

References

Schlotheuber, A., & Hosseinpoor, A. R. (2022). Summary measures of health inequality: A review of existing measures and their application. International Journal of Environmental Research and Public Health, 19 (6), 3697.

Ahn J, Harper S, Yu M, Feuer EJ, Liu B, Luta G. (2018). Variance Estimation and Confidence Intervals for 11 Commonly Used Health Disparity Measures. JCO Clin Cancer Inform. Dec;2:1–19.

Examples

```
# example code
data(NonorderedSample)
head(NonorderedSample)
with(NonorderedSample,
     bgv(pop = population,
         est = estimate,
          se = se
         )
     )
```

---

covar   *Coefficient of variation (COV)*

Description

The Coefficient of Variation (COV) is a relative measure of inequality that considers all population subgroups. Subgroups are weighted according to their population share.

Usage

```
covar(pop, est, se = NULL, scaleval, sim = NULL, seed = 123456, ...)
```

Arguments

pop	The number of people within each subgroup. Population size must be available for all subgroups.
est	The subgroup estimate. Estimates must be available for at least 85% of subgroups.
se	The standard error of the subgroup estimate. If this is missing, 95% confidence intervals of COV cannot be calculated.

scaleval	The scale of the indicator. For example, the scale of an indicator measured as a percentage is 100. The scale of an indicator measured as a rate per 1000 population is 1000.
sim	The number of simulations to estimate confidence intervals. Default is 100.
seed	The random number generator (RNG) state for the confidence interval simulation. Default is 123456.
...	Further arguments passed to or from other methods.

Details

COV is calculated by dividing the between-group standard deviation (BGSD) by the setting average and multiplying the fraction by 100. BGSD is calculated as the square root of the weighted average of squared differences between the subgroup estimates and the setting average. Squared differences are weighted by each subgroup's population share. For more information on this inequality measure see Schlotheuber, A., & Hosseinpoor, A. R. (2022) below.

95% confidence intervals are calculated using a methodology of simulated estimates. The dataset is simulated a large number of times (e.g., 100) and COV is calculated for each of the simulated samples. The 95% confidence intervals are based on the 2.5th and 97.5th percentiles of the COV results.

Interpretation: COV only has positive values, with larger values indicating higher levels of inequality. COV is zero if there is no inequality.

Type of summary measure: Complex; relative; weighted

Applicability: Non-ordered; more than two subgroups

Value

The estimated COV value, corresponding estimated standard error, and confidence interval as a `data.frame`.

References

Schlotheuber, A., & Hosseinpoor, A. R. (2022). Summary measures of health inequality: A review of existing measures and their application. International Journal of Environmental Research and Public Health, 19 (6), 3697.

Examples

```
# example code
data(NonorderedSample)
head(NonorderedSample)
with(NonorderedSample,
     covar(pop = population,
           est = estimate,
           se = se,
           scaleval = indicator_scale
          )
     )
```

---

d *Difference (D)*

Description

The difference (D) is an absolute measure of inequality that shows the difference in a health indicator between two population subgroups. For more information on this inequality measure see Schlotheuber, A., & Hosseinpoor, A. R. (2022) below.

Usage

```
d(
  est,
  se = NULL,
  favourable_indicator,
  ordered_dimension = NULL,
  subgroup_order = NULL,
  reference_subgroup = NULL,
  conf.level = 0.95,
  ...
)
```

Arguments

est	The subgroup estimate. Estimates must be available for the two subgroups being compared.
se	The standard error of the subgroup estimate. If this is missing, confidence intervals of D cannot be calculated.
favourable_indicator	
	Records whether the indicator is favourable (1) or non-favourable (0). Favourable indicators measure desirable health events where the ultimate goal is to achieve a maximum level (such as skilled birth attendance). Non-favourable indicators measure undesirable health events where the ultimate goal is to achieve a minimum level (such as under-five mortality rate).
ordered_dimension	
	Records whether the dimension is ordered (1) or not (0).
subgroup_order	The order of subgroups in an increasing sequence. Required if the dimension is ordered (ordered_dimension=1).
reference_subgroup	
	Identifies a reference subgroup with the value of 1.
conf.level	Confidence level of the interval. Default is 0.95 (95%).
...	Further arguments passed to or from other methods.

Details

D is calculated as: `D = y1 - y2`, where `y1` and `y2` indicate the estimates for subgroups 1 and 2. The selection of the two subgroups depends on the characteristics of the inequality dimension and the purpose of the analysis. In addition, the direction of the calculation may depend on the indicator type (favourable or#' adverse).

Ordered dimension: Favourable indicator: Most-advantaged subgroup - Least-advantaged subgroup Adverse indicator: Least-advantaged subgroup - Most-advantaged subgroup

Non-ordered dimension: No reference group & favourable indicator: Highest estimate - Lowest estimate No reference group & adverse indicator: Lowest estimate - Highest estimate Reference group & favourable indicator: Reference estimate - Lowest estimate Reference group & adverse indicator: Lowest estimate - Reference estimate

Interpretation: Greater absolute values indicate higher levels of inequality. D is zero if there is no inequality.

Type of summary measure: Simple; relative; unweighted.

Applicability: Any dimension of inequality.

Warning: The confidence intervals are approximate and might be biased.

Value

The estimated D value, corresponding estimated standard error, and confidence interval as a `data.frame`.

References

Schlotheuber, A., & Hosseinpoor, A. R. (2022). Summary measures of health inequality: A review of existing measures and their application. International Journal of Environmental Research and Public Health, 19 (6), 3697.

Examples

```
# example code
data(NonorderedSample)
head(NonorderedSample)
with(NonorderedSample,
      d(est = estimate,
        se = se,
        favourable_indicator = favourable_indicator,
        ordered_dimension = ordered_dimension,
        reference_subgroup = reference_subgroup
       )
    )
```

---

idisu   *Index of Disparity (unweighted) (IDISU)*

Description

The Index of Disparity (IDIS) is a relative measure of inequality that shows the average difference between each subgroup and the setting average, in relative terms. In the unweighted version (IDISU), all subgroups are weighted equally.

Usage

```
idisu(
  pop = NULL,
  est,
  se = NULL,
  scaleval,
  setting_average = NULL,
  sim = NULL,
  seed = 123456,
  ...
)
```

Arguments

pop	The number of people within each subgroup. Population size must be available for all subgroups.
est	The subgroup estimate. Estimates must be available for at least 85% of subgroups.
se	The standard error of the subgroup estimate. If this is missing, 95% confidence intervals of IDISU cannot be calculated.
scaleval	The scale of the indicator. For example, the scale of an indicator measured as a percentage is 100. The scale of an indicator measured as a rate per 1000 population is 1000.
setting_average	
	The reported setting average. Setting average must be unique for each setting, year, indicator combination. If population is not specified for all subgroups, the setting average is used.
sim	The number of simulations to estimate 95% confidence intervals. Default is 100.
seed	The random number generator (RNG) state for the 95% confidence interval simulation. Default is 123456.
...	Further arguments passed to or from other methods.

Details

IDISU is calculated as the average of absolute differences between the subgroup estimates and the setting average, divided by the number of subgroups and the setting average, and multiplied by 100. For more information on this inequality measure see Schlotheuber, A., & Hosseinpoor, A. R. (2022) below.

95% confidence intervals are calculated using a methodology of simulated estimates. The dataset is simulated a large number of times (e.g., 100) and IDISU is calculated for each of the simulated samples. The 95% confidence intervals are based on the 2.5th and 97.5th percentiles of the IDISU results.

Interpretation: IDISU has only positive values, with larger values indicating higher levels of inequality. IDISU is zero if there is no inequality.

Type of summary measure: Complex; relative; non-weighted

Applicability: Non-ordered; more than two subgroups

Value

The estimated IDISU value, corresponding estimated standard error, and confidence interval as a `data.frame`.

References

Schlotheuber, A., & Hosseinpoor, A. R. (2022). Summary measures of health inequality: A review of existing measures and their application. International Journal of Environmental Research and Public Health, 19 (6), 3697.

Examples

```
# example code
data(NonorderedSample)
head(NonorderedSample)
with(NonorderedSample,
     idisu(pop = population,
           est = estimate,
           se = se,
           scaleval = indicator_scale
          )
     )
```

---

idisw   *Index of Disparity (weighted) (IDISW)*

Description

The Index of Disparity (IDIS) is a relative measure of inequality that shows the average difference between each subgroup and the setting average, in relative terms. In the weighted version (IDISW), subgroups are weighted according to their population share.

Usage

```
idisw(pop, est, se = NULL, scaleval, sim = NULL, seed = 123456, ...)
```

Arguments

pop	The number of people within each subgroup. Population size must be available for all subgroups.
est	The subgroup estimate. Estimates must be available for at least 85% of subgroups.
se	The standard error of the subgroup estimate. If this is missing, 95% confidence intervals of IDISW cannot be calculated.
scaleval	The scale of the indicator. For example, the scale of an indicator measured as a percentage is 100. The scale of an indicator measured as a rate per 1000 population is 1000.
sim	The number of simulations to estimate 95% confidence intervals. Default is 100.
seed	The random number generator (RNG) state for the 95% confidence interval simulation. Default is 123456.
...	Further arguments passed to or from other methods.

Details

IDISW is calculated as the weighted average of absolute differences between the subgroup estimates and the setting average, divided by the setting average, and multiplied by 100. Absolute differences are weighted by each subgroup's population share. For more information on this inequality measure see Schlotheuber, A., & Hosseinpoor, A. R. (2022) below.

95% confidence intervals are calculated using a methodology of simulated estimates. The dataset is simulated a large number of times (e.g., 100) and IDISW is calculated for each of the simulated samples. The 95% confidence intervals are based on the 2.5th and 97.5th centiles of the IDISW results.

Interpretation: IDISW has only positive values, with larger values indicating higher levels of inequality. IDISW is zero if there is no inequality.

Type of summary measure: Complex; relative; weighted

Applicability: Non-ordered; more than two subgroups

Value

The estimated IDISW value, corresponding estimated standard error, and confidence interval as a `data.frame`.

References

Schlotheuber, A., & Hosseinpoor, A. R. (2022). Summary measures of health inequality: A review of existing measures and their application. International Journal of Environmental Research and Public Health, 19 (6), 3697.

Examples

```
# example code
data(NonorderedSample)
head(NonorderedSample)
with(NonorderedSample,
     idisw(pop = population,
           est = estimate,
           se,
           scaleval = indicator_scale
          )
     )
```

IndividualSample   *World Health Organization (WHO)*

Description

This dataset contains sample data for computing non-ordered summary measures of health inequality. It contains data from a household survey for two indicators, births attended by skilled health personnel (sba) and Diphtheria tetanus toxoid and pertussis (DTP3) immunization coverage, disaggregated by economic status. Both indicators are binary, (1) for those who had sba or dpt3 or (0) if the had not.

Usage

```
IndividualSample
```

Format

IndividualSample:
A data frame with 17,848 rows and 10 columns:

id individual identifier

psu Primary Sample Unit (PSU)

strata sampling strata

weight sampling weight

subgroup subgroup name

subgroup_order subgroup order

sba indicator estimate

dtp3 indicator estimate

favourable_indicator favourable (1) or non-favourable (0) indicator

indicator_scale scale of the indicator

Details

Births attended by skilled health personnel is defined as a birth attended by skilled health personnel divided by the total number of live births to women aged 15-49 years occurring in the period prior to the survey. Skilled health personnel include doctors, nurses, midwives and other medically trained personnel, as defined according to each country. DPT3 is measured among one-year-olds and indicate those who have received three doses of the combined diphtheria, tetanus toxoid and pertussis containing vaccine in a given year.This is in line with the definition used by the Countdown to 2030 Collaboration, Demographic and Health Surveys (DHS), Multiple Indicator Cluster Surveys (MICS) and Reproductive Health Surveys (RHS).

Economic status is determined using a wealth index, which is based on owning selected assets and having access to certain services. The wealth index is divided into five equal subgroups (quintiles) that each account for 20% of the population. Economic status is an ordered dimension (meaning that the subgroups have an inherent ordering).

This dataset can be used to calculate ordered summary measures of health inequality, including: absolute concentration index (ACI), relative concentration index (RCI), slope index of inequality (SII) and relative index of inequality (RII).

Source

WHO Health Inequality Data Repository https://www.who.int/data/inequality-monitor/data

Examples

```
head(IndividualSample)
summary(IndividualSample)
```

---

mdbu	*Mean difference from the best-performing subgroup (unweighted) (MDBU)*

Description

The Mean Difference from the Best-Performing Subgroup (MDB) is an absolute measure of inequality that shows the mean difference between each population subgroup and the subgroup with the best estimate. For the unweighted version (MDBU), all subgroups are weighted equally.

Usage

```
mdbu(
  est,
  se = NULL,
  favourable_indicator,
  scaleval,
  sim = NULL,
  seed = 123456,
  ...
)
```

Arguments

est	The subgroup estimate. Estimates must be available for at least 85% of subgroups.
se	The standard error of the subgroup estimate. If this is missing, 95% confidence intervals of MDBU cannot be calculated.
favourable_indicator	
	Records whether the indicator is favourable (1) or non-favourable (0). Favourable indicators measure desirable health events where the ultimate goal is to achieve a maximum level (such as skilled birth attendance). Non-favourable indicators measure undesirable health events where the ultimate goal is to achieve a minimum level (such as under-five mortality rate).
scaleval	The scale of the indicator. For example, the scale of an indicator measured as a percentage is 100. The scale of an indicator measured as a rate per 1000 population is 1000.
sim	The number of simulations to estimate 95% confidence intervals. Default is 100.
seed	The random number generator (RNG) state for the 95% confidence interval simulation. Default is 123456.
...	Further arguments passed to or from other methods.

Details

The unweighted version (MDBU) is calculated as the average of absolute differences between the subgroup estimates and the estimate for the best-performing subgroup, divided by the number of subgroups. For more information on this inequality measure see Schlotheuber, A., & Hosseinpoor, A. R. (2022) below.

95% confidence intervals are calculated using a methodology of simulated estimates. The dataset is simulated a large number of times (e.g., 100) and MDBU is calculated for each of the simulated samples. The 95% confidence intervals are based on the 2.5th and 97.5th percentiles of the MDBU results.

Interpretation: MDBU only has positive values, with larger values indicating higher levels of inequality. MDBU is zero if there is no inequality.

Type of summary measure: Complex; absolute; non-weighted

Applicability: Non-ordered; more than two subgroups

Value

The estimated MDBU value, corresponding estimated standard error, and confidence interval as a `data.frame`.

References

Schlotheuber, A., & Hosseinpoor, A. R. (2022). Summary measures of health inequality: A review of existing measures and their application. International Journal of Environmental Research and Public Health, 19 (6), 3697.

Examples

```
# example code
data(NonorderedSample)
head(NonorderedSample)
with(NonorderedSample,
    mdbu(est = estimate,
         se,
         favourable_indicator,
         scaleval = indicator_scale
       )
    )
```

---

mdbw   *Mean difference from the best-performing subgroup (weighted)*
                                (MDBW)

Description

The Mean Difference from the Best-Performing Subgroup (MDB) is an absolute measure of inequality that shows the mean difference between each population subgroup and the subgroup with the best estimate. For the weighted version (MDBW), subgroups are weighted according to their population share.

Usage

```
mdbw(
  pop,
  est,
  se = NULL,
  favourable_indicator,
  scaleval,
  sim = NULL,
  seed = 123456,
  ...
)
```

Arguments

pop	The number of people within each subgroup. Population size must be available for all subgroups.
est	The subgroup estimate. Estimates must be available for at least 85% of subgroups.
se	The standard error of the subgroup estimate. If this is missing, 95% confidence intervals of MDBW cannot be calculated.

favourable_indicator

> Records whether the indicator is favourable (1) or non-favourable (0). Favourable indicators measure desirable health events where the ultimate goal is to achieve a maximum level (such as skilled birth attendance). Non-favourable indicators measure undesirable health events where the ultimate goal is to achieve a minimum level (such as under-five mortality rate).

scaleval   The scale of the indicator. For example, the scale of an indicator measured as a percentage is 100. The scale of an indicator measured as a rate per 1000 population is 1000.

sim   The number of simulations to estimate 95% confidence intervals. Default is 100.

seed   The random number generator (RNG) state for the 95% confidence interval simulation. Default is 123456.

...   Further arguments passed to or from other methods.

Details

The weighted version (MDBW) is calculated as the weighted average of absolute differences between the subgroup estimates and the estimate for the best-performing subgroup, divided by the number of subgroups. For more information on this inequality measure see Schlotheuber, A., & Hosseinpoor, A. R. (2022) below.

95% confidence intervals are calculated using a methodology of simulated estimates. The dataset is simulated a large number of times (e.g., 100) and MDBW is calculated for each of the simulated samples. The 95% confidence intervals are based on the 2.5th and 97.5th percentiles of the MDBW results.

Interpretation: MDBW only has positive values, with larger values indicating higher levels of inequality. MDBW is zero if there is no inequality.

Type of summary measure: Complex; absolute; weighted

Applicability: Non-ordered; more than two subgroups

Value

The estimated MDBW value, corresponding estimated standard error, and confidence interval as a `data.frame`.

References

Schlotheuber, A., & Hosseinpoor, A. R. (2022). Summary measures of health inequality: A review of existing measures and their application. International Journal of Environmental Research and Public Health, 19 (6), 3697.

Examples

```
# example code
data(NonorderedSample)
head(NonorderedSample)
with(NonorderedSample,
     mdbw(pop = population,
```

```
        est = estimate,
        se,
        scaleval = indicator_scale,
        favourable_indicator
      )
  )
```

---

mdmu   *Mean difference from mean (unweighted) (MDMU)*

Description

The Mean Difference from Mean (MDM) is an absolute measure of inequality that shows the mean difference between each subgroup and the setting average. For the unweighted version (MDMU), all subgroups are weighted equally.

Usage

```
mdmu(
  pop = NULL,
  est,
  se = NULL,
  scaleval,
  setting_average = NULL,
  sim = NULL,
  seed = 123456,
  ...
)
```

Arguments

pop	The number of people within each subgroup. Population size must be available for all subgroups.
est	The subgroup estimate. Estimates must be available for at least 85% of subgroups.
se	The standard error of the subgroup estimate. If this is missing, 95% confidence intervals of MDMU cannot be calculated.
scaleval	The scale of the indicator. For example, the scale of an indicator measured as a percentage is 100. The scale of an indicator measured as a rate per 1000 population is 1000.
setting_average	
	The reported setting average. Setting average must be unique for each setting, year, indicator combination. If population is not specified for all subgroups, the setting average is used.
sim	The number of simulations to estimate 95% confidence intervals. Default is 100.
seed	The random number generator (RNG) state for the 95% confidence interval simulation. Default is 123456.
...	Further arguments passed to or from other methods.

Details

The unweighted version (MDMU) is calculated as the sum of the absolute differences between the subgroup estimates and the setting average, divided by the number of subgroups. For more information on this inequality measure see Schlotheuber, A., & Hosseinpoor, A. R. (2022) below.

95% confidence intervals are calculated using a methodology of simulated estimates. The dataset is simulated a large number of times (e.g., 100) and MDMU is calculated for each of the simulated samples. The 95% confidence intervals are based on the 2.5th and 97.5th percentiles of the MDMU results.

Interpretation: MDMU only has positive values, with larger values indicating higher levels of inequality. MDMU is zero if there is no inequality.

Type of summary measure: Complex; absolute; non-weighted

Applicability: Non-ordered; more than two subgroups

Value

The estimated MDMU value, corresponding estimated standard error, and confidence interval as a `data.frame`.

References

Schlotheuber, A., & Hosseinpoor, A. R. (2022). Summary measures of health inequality: A review of existing measures and their application. International Journal of Environmental Research and Public Health, 19 (6), 3697.

Examples

```
# example code
data(NonorderedSample)
head(NonorderedSample)
with(NonorderedSample,
     mdmu(pop = population,
         est = estimate,
         se = se,
         scaleval = indicator_scale,
         setting_average
       )
    )
```

---

mdmw   *Mean difference from mean (weighted) (MDMW)*

Description

The Mean Difference from Mean (MDM) is an absolute measure of inequality that shows the mean difference between each subgroup and the setting average. For the weighted version (MDMW), subgroups are weighted according to their population share.

Usage

```
mdmw(pop, est, se = NULL, scaleval, sim = NULL, seed = 123456, ...)
```

Arguments

pop	The number of people within each subgroup. Population size must be available for all subgroups.
est	The subgroup estimate. Estimates must be available for at least 85% of subgroups.
se	The standard error of the subgroup estimate. If this is missing, 95% confidence intervals of MDMW cannot be calculated.
scaleval	The scale of the indicator. For example, the scale of an indicator measured as a percentage is 100. The scale of an indicator measured as a rate per 1000 population is 1000.
sim	The number of simulations to estimate 95% confidence intervals. Default is 100.
seed	The random number generator (RNG) state for the 95% confidence interval simulation. Default is 123456.
...	Further arguments passed to or from other methods.

Details

The weighted version (MDMW) is calculated as the weighted average of absolute differences between the subgroup estimates and the setting average. Absolute differences are weighted by each subgroup's population share. For more information on this inequality measure see Schlotheuber, A., & Hosseinpoor, A. R. (2022) below.

95% confidence intervals are calculated using a methodology of simulated estimates. The dataset is simulated a large number of times (e.g., 100) and MDMW is calculated for each of the simulated samples. The 95% confidence intervals are based on the 2.5th and 97.5th percentiles of the MDMW results.

Interpretation: MDMW only has positive values, with larger values indicating higher levels of inequality. MDMW is zero if there is no inequality.

Type of summary measure: Complex; absolute; weighted

Applicability: Non-ordered; more than two subgroups

Value

The estimated MDMW value, corresponding estimated standard error, and confidence interval as a `data.frame`.

References

Schlotheuber, A., & Hosseinpoor, A. R. (2022). Summary measures of health inequality: A review of existing measures and their application. International Journal of Environmental Research and Public Health, 19 (6), 3697.

Examples

```
# example code
data(NonorderedSample)
head(NonorderedSample)
with(NonorderedSample,
     mdmw(pop = population,
          est = estimate,
          se = se,
          scaleval = indicator_scale
        )
     )
```

---

mdru   *Mean difference from a reference subgroup (unweighted) (MDRU)*

Description

The Mean Difference from a Reference Subgroup (MDR) is an absolute measure of inequality that shows the mean difference between each population subgroup and a reference subgroup. For the unweighted version (MDRU), all subgroups are weighted equally.

Usage

```
mdru(
  est,
  se = NULL,
  scaleval,
  reference_subgroup,
  sim = NULL,
  seed = 123456,
  ...
)
```

Arguments

est	The subgroup estimate. Estimates must be available for at least 85% of subgroups.
se	The standard error of the subgroup estimate. If this is missing, 95% confidence intervals of MDRU cannot be calculated.
scaleval	The scale of the indicator. For example, the scale of an indicator measured as a percentage is 100. The scale of an indicator measured as a rate per 1000 population is 1000.
reference_subgroup	
	Identifies a reference subgroup with the value of 1.
sim	The number of simulations to estimate 95% confidence intervals. Default is 100.
seed	The random number generator (RNG) state for the 95% confidence interval simulation. Default is 123456.
...	Further arguments passed to or from other methods.

Details

The unweighted version (MDRU) is calculated as the average of absolute differences between the subgroup estimates and the estimate for the reference subgroup, divided by the number of subgroups. For more information on this inequality measure see Schlotheuber, A., & Hosseinpoor, A. R. (2022) below.

95% confidence intervals are calculated using a methodology of simulated estimates. The dataset is simulated a large number of times (e.g., 100) and MDRU is calculated for each of the simulated samples. The 95% confidence intervals are based on the 2.5th and 97.5th percentiles of the MDRU results.

Interpretation: MDRU only has positive values, with larger values indicating higher levels of inequality. MDRU is zero if there is no inequality.

Type of summary measure: Complex; absolute; non-weighted

Applicability: Non-ordered; more than two subgroups

Value

The estimated MDRU value, corresponding estimated standard error, and confidence interval as a `data.frame`.

References

Schlotheuber, A., & Hosseinpoor, A. R. (2022). Summary measures of health inequality: A review of existing measures and their application. International Journal of Environmental Research and Public Health, 19 (6), 3697.

Examples

```
# example code
data(NonorderedSample)
head(NonorderedSample)
with(NonorderedSample,
    mdru(est = estimate,
        se = se,
        scaleval = indicator_scale,
        reference_subgroup
        )
    )
```

---

mdrw   *Mean difference from a reference subgroup (weighted) (MDRW)*

Description

An absolute measure of inequality that shows the mean difference between each population subgroup and a reference subgroup. In the weighted version (MDRW), subgroups are weighted according to their population share.

Usage

```
mdrw(
  pop,
  est,
  se = NULL,
  scaleval,
  reference_subgroup,
  sim = NULL,
  seed = 123456,
  ...
)
```

Arguments

pop	The number of people within each subgroup. Population size must be available for all subgroups.
est	The subgroup estimate. Estimates must be available for at least 85% of subgroups.
se	The standard error of the subgroup estimate. If this is missing, 95% confidence intervals of MDRW cannot be calculated.
scaleval	The scale of the indicator. For example, the scale of an indicator measured as a percentage is 100. The scale of an indicator measured as a rate per 1000 population is 1000.
reference_subgroup	
	Identifies a reference subgroup with the value of 1.
sim	The number of simulations to estimate 95% confidence intervals. Default is 100.
seed	The random number generator (RNG) state for the 95% confidence interval simulation. Default is 123456.
...	Further arguments passed to or from other methods.

Details

The weighted version (MDRW) is calculated as the weighted average of absolute differences between the subgroup estimates and the estimate for the reference subgroup. Absolute differences are weighted by each subgroup's population share. For more information on this inequality measure see Schlotheuber, A., & Hosseinpoor, A. R. (2022) below.

95% confidence intervals are calculated using a methodology of simulated estimates. The dataset is simulated a large number of times (e.g., 100) and MDRW is calculated for each of the simulated samples. The 95% confidence intervals are based on the 2.5th and 97.5th percentiles of the MDRW results.

Interpretation: MDRW only has positive values, with larger values indicating higher levels of inequality. MDRW is zero if there is no inequality.

Type of summary measure: Complex; absolute; weighted

Applicability: Non-ordered; more than two subgroups

Value

The estimated MDRW value, corresponding estimated standard error, and confidence interval as a `data.frame`.

References

Schlotheuber, A., & Hosseinpoor, A. R. (2022). Summary measures of health inequality: A review of existing measures and their application. International journal of environmental research and public health, 19 (6), 3697.

Examples

```
# example code
data(NonorderedSample)
head(NonorderedSample)
with(NonorderedSample,
    mdrw(pop = population,
        est = estimate,
        se = se,
        scaleval = indicator_scale,
        reference_subgroup
       )
    )
```

---

mld *Mean log deviation (MLD)*

Description

The Mean Log Deviation (MLD) is a relative measure of inequality that considers all population subgroups. Subgroups are weighted according to their population share.

Usage

```
mld(pop, est, se = NULL, conf.level = 0.95, ...)
```

Arguments

pop	The number of people within each subgroup. Population size must be available for all subgroups.
est	The subgroup estimate. Estimates must be available for at least 85% of subgroups.
se	The standard error of the subgroup estimate. If this is missing, confidence intervals of MLD cannot be calculated.
conf.level	Confidence level of the interval. Default is 0.95 (95%).
...	Further arguments passed to or from other methods.

Details

MLD is calculated as the sum of products between the negative natural logarithm of the share of the indicator of each subgroup and the population share of each subgroup. MLD may be more easily readable when multiplied by 1000. For more information on this inequality measure see Schlotheuber, A., & Hosseinpoor, A. R. (2022) below.

Interpretation: MLD is zero if there is no inequality. Greater absolute values indicate higher levels of inequality. MLD is more sensitive to differences further from the setting average (by the use of the logarithm).

Type of summary measure: Complex; relative; weighted

Applicability: Non-ordered; more than two subgroups

Warning: The confidence intervals are approximate and might be biased. See Ahn J. et al. (2018) below for further information on the standard error formula.

Value

The estimated MLD value, corresponding estimated standard error, and confidence interval as a `data.frame`.

References

Schlotheuber, A., & Hosseinpoor, A. R. (2022). Summary measures of health inequality: A review of existing measures and their application. International Journal of Environmental Research and Public Health, 19 (6), 3697.

Ahn J, Harper S, Yu M, Feuer EJ, Liu B, Luta G. (2018). Variance Estimation and Confidence Intervals for 11 Commonly Used Health Disparity Measures. JCO Clin Cancer Inform. Dec;2:1–19.

Examples

```
# example code
data(NonorderedSample)
head(NonorderedSample)
with(NonorderedSample,
    mld(pop = population,
        est = estimate,
        se = se
        )
    )
```

NonorderedSample   *World Health Organization (WHO)*

Description

This dataset contains sample data for computing non-ordered summary measures of health inequality. It contains data from a household survey for the proportion of births attended by skilled health personnel disaggregated by subnational region.

Usage

```
NonorderedSample
```

Format

```
NonorderedSample:
```
A data frame with 34 rows and 11 columns:

indicator indicator name

dimension dimension of inequality

subgroup population subgroup within a given dimension of inequality

estimate subgroup estimate

se standard error of the subgroup estimate

population number of people within each subgroup

setting_average indicator average for the setting

favourable_indicator favourable (1) or non-favourable (0) indicator

ordered_dimension ordered (1) or non-ordered (0) dimension

indicator_scale scale of the indicator

reference_subgroup reference subgroup

Details

The proportion of births attended by skilled health personnel is calculated as the number of births attended by skilled health personnel divided by the total number of live births to women aged 15-49 years occurring in the period prior to the survey.

Skilled health personnel include doctors, nurses, midwives and other medically trained personnel, as defined according to each country. This is in line with the definition used by the Countdown to 2030 Collaboration, Demographic and Health Surveys (DHS), Multiple Indicator Cluster Surveys (MICS) and Reproductive Health Surveys (RHS).

Subnational regions are defined using country-specific criteria. Subnational region is a non-ordered dimension (meaning that the subgroups do not have an inherent ordering).

This dataset can be used to calculate non-ordered summary measures of health inequality, including: between-group variance (BGV), between-group standard deviation (BGSD), coefficient of variation (COV), mean difference from mean (MDM), index of disparity (IDIS), Theil index (TI) and mean log deviation (MLD). It can also be used to calculate the impact measures population attributable risk (PAR) and population attributable fraction (PAF).

Source

WHO Health Inequality Data Repository https://www.who.int/data/inequality-monitor/data

Examples

```
head(NonorderedSample)
summary(NonorderedSample)
```

NonorderedSampleMultipleind

World Health Organization (WHO)

Description

This dataset contains sample data for computing non-ordered summary measures of health inequality. It contains data from a household survey for two indicators, the proportion of births attended by skilled health personnel and under-five mortality rate, disaggregated by subnational region.

Usage

NonorderedSampleMultipleind

Format

NonorderedSampleMultipleind:

A data frame with 71 rows and 11 columns:

indicator indicator name

dimension dimension of inequality

subgroup population subgroup within a given dimension of inequality

estimate subgroup estimate

se standard error of the subgroup estimate

population number of people within each subgroup

setting_average indicator average for the setting

favourable_indicator favourable (1) or non-favourable (0) indicator

ordered_dimension ordered (1) or non-ordered (0) dimension

indicator_scale scale of the indicator

reference_subgroup reference subgroup

Details

The proportion of births attended by skilled health personnel is calculated as the number of births attended by skilled health personnel divided by the total number of live births to women aged 15-49 years occurring in the period prior to the survey.

Skilled health personnel include doctors, nurses, midwives and other medically trained personnel, as defined according to each country. This is in line with the definition used by the Countdown to 2030 Collaboration, Demographic and Health Surveys (DHS), Multiple Indicator Cluster Surveys (MICS) and Reproductive Health Surveys (RHS).

The under-five mortality rate is the probability (expressed as a rate per 1000 live births) of a child born in a specific year or period dying before reaching the age of five years. It is calculated as the number of deaths at age 0-5 years divided by the number of surviving children at the beginning of the specified age range during the 10 years prior to the survey.

Subnational regions are defined using country-specific criteria. Subnational region is a non-ordered dimension (meaning that the subgroups do not have an inherent ordering).

This dataset can be used to calculate non-ordered summary measures of health inequality, including: between-group variance (BGV), between-group standard deviation (BGSD), coefficient of variation (COV), mean difference from mean (MDM), index of disparity (IDIS), Theil index (TI) and mean log deviation (MLD). It can also be used to calculate the impact measures population attributable risk (PAR) and population attributable fraction (PAF).

Source

WHO Health Inequality Data Repository <https://www.who.int/data/inequality-monitor/data> head(NonorderedSampleMultipleind) summary(NonorderedSampleMultipleind)

OrderedSample   *World Health Organization (WHO)*

Description

This dataset contains sample data for computing ordered summary measures of health inequality. It contains data from a household survey for the proportion of births attended by skilled health personnel disaggregated by economic status, measured by wealth quintiles.

Usage

```
OrderedSample
```

Format

OrderedSample:

A data frame with 5 rows and 11 columns.

indicator indicator name

dimension dimension of inequality

subgroup population subgroup within a given dimension of inequality

subgroup_order the order of subgroups in an increasing sequence

estimate subgroup estimate

se standard error of the subgroup estimate

population number of people within each subgroup

setting_average indicator average for the setting

favourable_indicator favourable (1) or non-favourable (0) indicator

ordered_dimension ordered (1) or non-ordered (0) dimension

indicator_scale scale of the indicator

Details

The proportion of births attended by skilled health personnel is calculated as the number of births attended by skilled health personnel divided by the total number of live births to women aged 15-49 years occurring in the period prior to the survey.

Skilled health personnel include doctors, nurses, midwives and other medically trained personnel, as defined according to each country. This is in line with the definition used by the Countdown to 2030 Collaboration, Demographic and Health Surveys (DHS), Multiple Indicator Cluster Surveys (MICS) and Reproductive Health Surveys (RHS).

Economic status is determined using a wealth index, which is based on owning selected assets and having access to certain services. The wealth index is divided into five equal subgroups (quintiles) that each account for 20% of the population. Economic status is an ordered dimension (meaning that the subgroups have an inherent ordering).

This dataset can be used to calculate ordered summary measures of health inequality, including: absolute concentration index (ACI), relative concentration index (RCI), slope index of inequality (SII) and relative index of inequality (RII). It can also be used to calculate the impact measures population attributable risk (PAR) and population attributable fraction (PAF).

Source

WHO Health Inequality Data Repository https://www.who.int/data/inequality-monitor/data

Examples

```
head(OrderedSample)
summary(OrderedSample)
```

OrderedSampleMultipleind

World Health Organization (WHO)

Description

This dataset contains sample data for computing ordered summary measures of health inequality. It contains data from a household survey for two indicators, the proportion of births attended by skilled health personnel and under-five mortality rate, disaggregated by economic status.

Usage

```
OrderedSampleMultipleind
```

Format

OrderedSampleMultipleind:
A data frame with 10 rows and 11 columns:

indicator indicator name
dimension dimension of inequality

subgroup population subgroup within a given dimension of inequality

subgroup_order the order of subgroups in an increasing sequence

estimate subgroup estimate

se standard error of the subgroup estimate

population number of people within each subgroup

setting_average indicator average for the setting

favourable_indicator favourable (1) or non-favourable (0) indicator

ordered_dimension ordered (1) or non-ordered (0) dimension

indicator_scale scale of the indicator

Details

The proportion of births attended by skilled health personnel is calculated as the number of births attended by skilled health personnel divided by the total number of live births to women aged 15-49 years occurring in the period prior to the survey.

Skilled health personnel include doctors, nurses, midwives and other medically trained personnel, as defined according to each country. This is in line with the definition used by the Countdown to 2030 Collaboration, Demographic and Health Surveys (DHS), Multiple Indicator Cluster Surveys (MICS) and Reproductive Health Surveys (RHS).

The under-five mortality rate is the probability (expressed as a rate per 1000 live births) of a child born in a specific year or period dying before reaching the age of five years. It is calculated as the number of deaths at age 0-5 years divided by the number of surviving children at the beginning of the specified age range during the 10 years prior to the survey.

Economic status is determined using a wealth index, which is based on owning selected assets and having access to certain services. The wealth index is divided into five equal subgroups (quintiles) that each account for 20% of the population. Economic status is an ordered dimension (meaning that the subgroups have an inherent ordering).

This dataset can be used to calculate ordered summary measures of health inequality, including: absolute concentration index (ACI), relative concentration index (RCI), slope index of inequality (SII) and relative index of inequality (RII). It can also be used to calculate the impact measures population attributable risk (PAR) and population attributable fraction (PAF).

Source

WHO Health Inequality Data Repository https://www.who.int/data/inequality-monitor/data

Examples

```
head(OrderedSampleMultipleind)
summary(OrderedSampleMultipleind)
```

---

paf   *Population attributable fraction (PAF)*

Description

The population attributable fraction (PAF) is a relative measure of inequality that shows the potential improvement in the average of an indicator, in relative terms, that could be achieved if all population subgroups had the same level of the indicator as a reference group.

Usage

```
paf(
  pop = NULL,
  est,
  ordered_dimension,
  subgroup_order = NULL,
  setting_average = NULL,
  favourable_indicator,
  scaleval,
  conf.level = 0.95,
  ...
)
```

Arguments

pop	The number of people within each subgroup. Population size must be available for all subgroups.
est	The subgroup estimate. Estimates must be available for all subgroups.
ordered_dimension	
	Records whether the dimension is ordered (1) or not (0).
subgroup_order	The order of subgroups in an increasing sequence, if the dimension is ordered (ordered dimension=1).
setting_average	
	The reported setting average. Setting average must be unique for each setting, year, indicator combination. If population is not specified for all subgroups, the setting average is used.
favourable_indicator	
	Records whether the indicator is favourable (1) or non-favourable (0). Favourable indicators measure desirable health events where the ultimate goal is to achieve a maximum level (such as skilled birth attendance). Non-favourable indicators measure undesirable health events where the ultimate goal is to achieve a minimum level (such as under-five mortality rate).
scaleval	The scale of the indicator. For example, the scale of an indicator measured as a percentage is 100. The scale of an indicator measured as a rate per 1000 population is 1000.
conf.level	Confidence level of the interval. Default is 0.95 (95%).
...	Further arguments passed to or from other methods.

Details

PAF is calculated as the difference between the estimate for the reference subgroup and the setting average, divided by the setting average and multiplied by 100. For more information on this inequality measure see Schlotheuber, A., & Hosseinpoor, A. R. (2022) below.

If the indicator is favourable and PAF < 0, then PAF is replaced with 0. If the indicator is adverse and PAF > 0, then PAF is replaced with 0. The selection of the reference subgroup depends on the characteristics of the inequality dimension and the indicator type. It is the most-advantaged subgroup for ordered dimensions. For non-ordered dimensions, it is the subgroup with the highest estimate for favourable indicators and is the subgroup with the lowest estimate for adverse indicators.

Interpretation: PAF assumes positive values for favourable indicators and negative values for non-favourable (adverse) indicators. The larger the absolute value of PAF, the higher the level of inequality. PAF is zero if no further improvement can be achieved (i.e., if all subgroups have reached the same level of the indicator as the reference subgroup or surpassed that level).

Type of summary measure: Complex; relative; weighted

Applicability: Any

Warning: The confidence intervals are approximate and might be biased. See Walter S.D. (1978) below for further information on the standard error formula.

Value

The estimated PAF value, corresponding estimated standard error, and confidence interval as a `data.frame`.

References

Schlotheuber, A., & Hosseinpoor, A. R. (2022). Summary measures of health inequality: A review of existing measures and their application. International Journal of Environmental Research and Public Health, 19 (6), 3697.

Walter, S. D. (1978). "Calculation of Attributable Risks from Epidemiological Data." International Journal of Epidemiology 7 (2): 175–82.

Examples

```
# example code
data(OrderedSample)
head(OrderedSample)
with(OrderedSample,
     paf(pop = population,
         est = estimate,
         ordered_dimension = ordered_dimension,
         subgroup_order = subgroup_order,
         favourable_indicator = favourable_indicator,
         scaleval = indicator_scale
         )
     )
```

---

parisk   *Population attributable risk (PAR)*

Description

Population Attributable Risk (PAR) is an absolute measure of inequality that shows the potential improvement in the average of an indicator, in absolute terms, that could be achieved if all population subgroups had the same level of the indicator as a reference group.

Usage

```
parisk(
  pop = NULL,
  est,
  ordered_dimension,
  subgroup_order = NULL,
  setting_average = NULL,
  favourable_indicator,
  scaleval,
  conf.level = 0.95,
  ...
)
```

Arguments

pop
: The number of people within each subgroup. Population size must be available for all subgroups.

est
: The subgroup estimate. Estimates must be available for all subgroups.

ordered_dimension
: Records whether the dimension is ordered (1) or not (0).

subgroup_order
: The order of subgroups in an increasing sequence, if the dimension is ordered.

setting_average
: The reported setting average. Setting average must be unique for each setting, year, indicator combination. If population is not specified for all subgroups, the setting average is used.

favourable_indicator
: Records whether the indicator is favourable (1) or non-favourable (0). Favourable indicators measure desirable health events where the ultimate goal is to achieve a maximum level (such as skilled birth attendance). Non-favourable indicators measure undesirable health events where the ultimate goal is to achieve a minimum level (such as under-five mortality rate).

scaleval
: The scale of the indicator. For example, the scale of an indicator measured as a percentage is 100. The scale of an indicator measured as a rate per 1000 population is 1000.

conf.level
: Confidence level of the interval. Default is 0.95 (95%).

...
: Further arguments passed to or from other methods.

Details

PAR is calculated as the difference between the estimate for the reference subgroup and the setting average. For more information on this inequality measure see Schlotheuber, A., & Hosseinpoor, A. R. (2022) below.

If the indicator is favourable and PAR < 0, then PAR is replaced with 0. If the indicator is adverse and PAR > 0, then PAR is replaced with 0. The selection of the reference subgroup depends on the characteristics of the inequality dimension and the indicator type. It is the most-advantaged subgroup for ordered dimensions. For non-ordered dimensions, it is the subgroup with the highest estimate for favourable indicators and is the subgroup with the lowest estimate for adverse indicators.

Interpretation: PAR assumes positive values for favourable indicators and negative values for non-favourable (adverse) indicators. The larger the absolute value of PAR, the higher the level of inequality. PAR is zero if no further improvement can be achieved (i.e., if all subgroups have reached the same level of the indicator as the reference subgroup or surpassed that level).

Type of summary measure: Complex; absolute; weighted

Applicability: Any

Warning: The confidence intervals are approximate and might be biased. See Walter S.D. (1978) below for further information on the standard error formula.

Value

The estimated PAR value, corresponding estimated standard error, and confidence interval as a `data.frame`.

References

Schlotheuber, A., & Hosseinpoor, A. R. (2022). Summary measures of health inequality: A review of existing measures and their application. International Journal of Environmental Research and Public Health, 19 (6), 3697.

Walter, S. D. (1978). "Calculation of Attributable Risks from Epidemiological Data." International Journal of Epidemiology 7 (2): 175–82.

Examples

```
# example code
data(OrderedSample)
head(OrderedSample)
with(OrderedSample,
     parisk(pop = population,
            est = estimate,
            ordered_dimension = ordered_dimension,
            subgroup_order = subgroup_order,
            favourable_indicator = favourable_indicator,
            scaleval = indicator_scale
            )
     )
```

---

r   *Ratio (R)*

Description

The ratio (R) is a relative measure of inequality that shows the ratio of a health indicator between two population subgroups. For more information on this inequality measure see Schlotheuber, A., & Hosseinpoor, A. R. (2022) below.

Usage

```
r(
  est,
  se = NULL,
  favourable_indicator,
  ordered_dimension = NULL,
  subgroup_order = NULL,
  reference_subgroup = NULL,
  conf.level = 0.95,
  ...
)
```

Arguments

est	The subgroup estimate. Estimates must be available for the two subgroups being compared.
se	The standard error of the subgroup estimate. If this is missing, confidence intervals of R cannot be calculated.
favourable_indicator	
	Records whether the indicator is favourable (1) or non-favourable (0). Favourable indicators measure desirable health events where the ultimate goal is to achieve a maximum level (such as skilled birth attendance). Non-favourable indicators measure undesirable health events where the ultimate goal is to achieve a minimum level (such as under-five mortality rate).
ordered_dimension	
	Records whether the dimension is ordered (1) or not (0).
subgroup_order	The order of subgroups in an increasing sequence. Required if the dimension is ordered (ordered_dimension=1).
reference_subgroup	
	Identifies a reference subgroup with the value of 1.
conf.level	Confidence level of the interval. Default is 0.95 (95%).
...	Further arguments passed to or from other methods.

Details

R is calculated as: `R = y1/y2` where `y1` and `y2` indicate the estimates for subgroups 1 and 2. The selection of the two subgroups depends on the characteristics of the inequality dimension and the purpose of the analysis. In addition, the direction of the calculation may depend on the indicator type (favourable or adverse).

Ordered dimension: Favourable indicator: Most-advantaged subgroup / Least-advantaged subgroup Adverse indicator: Least-advantaged subgroup / Most-advantaged subgroup

Non-ordered dimension: No reference group & favourable indicator: Highest estimate / Lowest estimate No reference group & adverse indicator: Lowest estimate / Highest estimate Reference group & favourable indicator: Reference estimate / Lowest estimate Reference group & adverse indicator: Lowest estimate / Reference estimate

Interpretation: R only assumes positive values. The further the value of R from one, the higher the level of inequality. R is 1 if there is no inequality.

Type of summary measure: Simple; relative; unweighted.

Applicability: Any dimension of inequality.

Warning: The confidence intervals are approximate and might be biased. See Ahn J. et al. (1978) below for further information on the standard error formula.

Value

The estimated D value, corresponding estimated standard error, and confidence interval as a `data.frame`.

References

Schlotheuber, A., & Hosseinpoor, A. R. (2022). Summary measures of health inequality: A review of existing measures and their application. International Journal of Environmental Research and Public Health, 19 (6), 3697.

Ahn J, Harper S, Yu M, Feuer EJ, Liu B, Luta G. Variance Estimation and Confidence Intervals for 11 Commonly Used Health Disparity Measures. JCO Clin Cancer Inform. 2018 Dec;2:1–19.

Examples

```
# example code
data(NonorderedSample)
head(NonorderedSample)
with(NonorderedSample,
    r(est = estimate,
      se = se,
      favourable_indicator,
      ordered_dimension = ordered_dimension,
      reference_subgroup = reference_subgroup
        )
    )
```

---

rci   *Relative concentration index (RCI)*

Description

The relative concentration index (RCI) is a relative measure of inequality that shows the gradient across population subgroups. It indicates the extent to which an indicator is concentrated among disadvantaged or advantaged subgroups, on a relative scale. RCI can be calculated using disaggregated data and individual-level data. Subgroups in disaggregated data are weighted according to their population share, while individuals are weighted by sample weight in the case of data from surveys.

Usage

```
rci(
  est,
  subgroup_order,
  scaleval = NULL,
  pop = NULL,
  weight = NULL,
  psu = NULL,
  strata = NULL,
  fpc = NULL,
  method = NULL,
  lmin = NULL,
  lmax = NULL,
  conf.level = 0.95,
  force = FALSE,
  ...
)
```

Arguments

est	The indicator estimate. Estimates must be available for all subgroups/individuals (unless force=TRUE).
subgroup_order	The order of subgroups/individuals in an increasing sequence.
scaleval	The scale of the indicator. For example, the scale of an indicator measured as a percentage is 100. The scale of an indicator measured as a rate per 1000 population is 1000.
pop	The number of people within each subgroup (for disaggregated data). Population size must be available for all subgroups.
weight	Individual sampling weight (required if data come from a survey)
psu	Primary sampling unit (required if data come from a survey)
strata	Strata (required if data come from a survey)

fpc	Finite population correction (if data come from a survey and sample size is large relative to population size).
method	Normalisation method for bounded indicators. Options available Wagstaff (`wagstaff`) or Erreygers (`erreygers`) correction.
lmin	Minimum limit for bounded indicators (i.e., variables that have a finite upper and/or lower limit).
lmax	Maximum limit for bounded indicators (i.e., variables that have a finite upper and/or lower limit).
conf.level	Confidence level of the interval. Default is 0.95 (95%).
force	TRUE/FALSE statement to force calculation with missing indicator estimate values.
...	Further arguments passed to or from other methods.

Details

The calculation of RCI is based on a ranking of the whole population from the most-disadvantaged subgroup (at rank 0) to the most-advantaged subgroup (at rank 1), which is inferred from the ranking and size of the subgroups. RCI can be calculated as twice the covariance between the health indicator and the relative rank, divided by the indicator mean. Given the relationship between covariance and ordinary least squares regression, RCI can be obtained from a regression of a transformation of the health variable of interest on the relative rank. For more information on this inequality measure see Schlotheuber, A., & Hosseinpoor, A. R. (2022) below.

Interpretation: RCI is bounded between -1 and +1 (or between -100 and +100, when multiplied by 100). The larger the absolute value of RCI, the higher the level of inequality. For favourable indicators, positive values indicate a concentration of the indicator among the advantaged, while negative values indicate a concentration of the indicator among the disadvantaged. For adverse indicators, it is the reverse: positive values indicate a concentration of the indicator among the disadvantaged, while negative values indicate a concentration of the indicator among the advantaged. RCI is zero if there is no inequality.

Type of summary measure: Complex; relative; weighted

Applicability: Ordered; more than two subgroups

Warning: The confidence intervals are approximate and might be biased.

Value

The estimated RCI value, corresponding estimated standard error, and confidence interval as a `data.frame`.

References

Schlotheuber, A., & Hosseinpoor, A. R. (2022). Summary measures of health inequality: A review of existing measures and their application. International Journal of Environmental Research and Public Health, 19 (6), 3697.

Wagstaff A. (2005). The bounds of the concentration index when the variable of interest is binary, with an application to immunization inequality. Health Economics, 14:429–432.

Erreygers G. (2009). Correcting the Concentration Index. Journal of Health Economics, 28:504–515.

Examples

```
# example code
data(IndividualSample)
head(IndividualSample)
with(IndividualSample,
    rci(est = sba,
        subgroup_order = subgroup_order,
        weight = weight,
        psu = psu,
        strata = strata
        )
    )
```

---

rii   *Relative index of inequality (RII)*

Description

The relative index of inequality (RII) is a relative measure of inequality that represents the ratio of estimated indicator values between the most-advantaged and most-disadvantaged, while taking into consideration the situation in all other subgroups/individuals – using an appropriate regression model. RII can be calculated using disaggregated data and individual-level data. Subgroups in disaggregated data are weighted according to their population share, while individuals are weighted by sample weight in the case of data from surveys.

Usage

```
rii(
  est,
  subgroup_order,
  pop = NULL,
  scaleval = NULL,
  weight = NULL,
  psu = NULL,
  fpc = NULL,
  strata = NULL,
  conf.level = 0.95,
  linear = FALSE,
  force = FALSE,
  ...
)
```

Arguments

est   The indicator estimate. Estimates must be available for all subgroups/individuals
                (unless force=TRUE).

subgroup_order The order of subgroups/individuals in an increasing sequence.

pop	The number of people within each subgroup (for disaggregated data). Population size must be available for all subgroups.
scaleval	The scale of the indicator. For example, the scale of an indicator measured as a percentage is 100. The scale of an indicator measured as a rate per 1000 population is 1000.
weight	Individual sampling weight (required if data come from a survey)
psu	Primary sampling unit (required if data come from a survey)
fpc	Finite population correction (if data come from a survey and sample size is large relative to population size).
strata	Strata (required if data come from a survey)
conf.level	Confidence level of the interval. Default is 0.95 (95%).
linear	TRUE/FALSE statement to specify the use of a linear regression model for RII estimation (default is logistic regression)
force	TRUE/FALSE statement to force calculation with missing indicator estimate values.
...	Further arguments passed to or from other methods.

Details

To calculate RII, a weighted sample of the whole population is ranked from the most-disadvantaged subgroup (at rank 0) to the most-advantaged subgroup (at rank 1). This ranking is weighted, accounting for the proportional distribution of the population within each subgroup. The population of each subgroup is then considered in terms of its range in the cumulative population distribution, and the midpoint of this range. The indicator of interest is then regressed against this midpoint value using an appropriate regression model (e.g., a generalized linear model with logit link), and the predicted values of the indicator are calculated for the two extremes (rank 1 and rank 0). The ratio between the estimated values at rank 1 and rank 0 (covering the entire distribution) generates the RII value. For more information on this inequality measure see Schlotheuber, A., & Hosseinpoor, A. R. (2022) below.

Interpretation: RII has the value of one if there is no inequality. RII has only positive values. Greater absolute values indicate higher levels of inequality. The further the value of RII from one, the higher the level of inequality. For favourable indicators, values larger than one indicate a concentration of the indicator among the advantaged and values smaller than one indicate a concentration of the indicator among the disadvantaged. For adverse indicators, values larger than one indicate a concentration of the indicator among the disadvantaged and values smaller than one indicate a concentration of the indicator among the advantaged. RII is a multiplicative measure and has to be displayed on a logarithmic scale (values larger than one are equivalent in magnitude to their reciprocal values smaller than one, e.g., a value of 2 is equivalent in magnitude to a value of 0.5).

Type of summary measure: Complex; relative; weighted

Applicability: Ordered; more than two subgroups

Warning: The confidence intervals are approximate and might be biased.

Value

The estimated RII value, corresponding estimated standard error, and confidence interval as a data.frame.

References

Schlotheuber, A., & Hosseinpoor, A. R. (2022). Summary measures of health inequality: A review of existing measures and their application. International journal of environmental research and public health, 19 (6), 3697.

Examples

```
# example code
data(IndividualSample)
head(IndividualSample)
with(IndividualSample,
    rii(est = sba,
        subgroup_order = subgroup_order,
        weight = weight,
        psu = psu,
        strata = strata
        )
    )
```

---

sii   *Slope index of inequality (SII)*

Description

The slope index of inequality (SII) is an absolute measure of inequality that represents the difference in estimated indicator values between the most-advantaged and most-disadvantaged, while taking into consideration the situation in all other subgroups/individuals – using an appropriate regression model. SII can be calculated using both disaggregated data and individual-level data. Subgroups in disaggregated data are weighted according to their population share, while individuals are weighted by sample weight in the case of data from surveys.

Usage

```
sii(
  est,
  subgroup_order,
  pop = NULL,
  scaleval = NULL,
  weight = NULL,
  psu = NULL,
  strata = NULL,
  fpc = NULL,
  conf.level = 0.95,
  linear = FALSE,
  force = FALSE,
  ...
)
```

Arguments

`est`	The indicator estimate. Estimates must be available for all subgroups/individuals (unless force=TRUE).
`subgroup_order`	The order of subgroups/individuals in an increasing sequence.
`pop`	The number of people within each subgroup (for disaggregated data). Population size must be available for all subgroups.
`scaleval`	The scale of the indicator. For example, the scale of an indicator measured as a percentage is 100. The scale of an indicator measured as a rate per 1000 population is 1000.
`weight`	Individual sampling weight (required if data come from a survey)
`psu`	Primary sampling unit (required if data come from a survey)
`strata`	Strata (required if data come from a survey)
`fpc`	Finite population correction (if data come from a survey and sample size is large relative to population size).
`conf.level`	Confidence level of the interval. Default is 0.95 (95%).
`linear`	TRUE/FALSE statement to specify the use of a linear regression model for SII estimation (default is logistic regression).
`force`	TRUE/FALSE statement to force calculation with missing indicator estimate values.
`...`	Further arguments passed to or from other methods.

Details

To calculate SII, a weighted sample of the whole population is ranked from the most-disadvantaged subgroup (at rank 0) to the most-advantaged subgroup (at rank 1). This ranking is weighted, accounting for the proportional distribution of the population within each subgroup. The indicator of interest is then regressed against this relative rank using an appropriate regression model (e.g., a generalized linear model with logit link), and the predicted values of the indicator are calculated for the two extremes (rank 1 and rank 0). The difference between the predicted values at rank 1 and rank 0 (covering the entire distribution) generates the SII value. For more information on this inequality measure see Schlotheuber, A., & Hosseinpoor, A. R. (2022) below.

Interpretation: SII is zero if there is no inequality. Greater absolute values indicate higher levels of inequality. For favourable indicators, positive values indicate a concentration of the indicator among the advantaged, while negative values indicate a concentration of the indicator among the disadvantaged. For adverse indicators, it is the reverse: positive values indicate a concentration of the indicator among the disadvantaged, while negative values indicate a concentration of the indicator among the advantaged.

Type of summary measure: Complex; absolute; weighted

Applicability: Ordered; more than two subgroups

Warning: The confidence intervals are approximate and might be biased.

Value

The estimated SII value, corresponding estimated standard error, and confidence interval as a `data.frame`.

References

Schlotheuber, A., & Hosseinpoor, A. R. (2022). Summary measures of health inequality: A review of existing measures and their application. International journal of environmental research and public health, 19 (6), 3697.

Examples

```
# example code
data(IndividualSample)
head(IndividualSample)
with(IndividualSample,
    sii(est = sba,
        subgroup_order = subgroup_order,
        weight = weight,
        psu = psu,
        strata = strata
        )
    )
```

---

ti   *Theil Index (TI)*

Description

The Theil Index (TI) is a relative measure of inequality that considers all population subgroups. Subgroups are weighted according to their population share.

Usage

```
ti(pop, est, se = NULL, conf.level = 0.95, ...)
```

Arguments

pop   The number of people within each subgroup. Population size must be available for all subgroups.

est   The subgroup estimate. Estimates must be available for at least 85% of subgroups.

se   The standard error of the subgroup estimate. If this is missing, confidence intervals of TI cannot be calculated.

conf.level   Confidence level of the interval. Default is 0.95 (95%).

...   Further arguments passed to or from other methods.

Details

TI is calculated as the sum of products of the natural logarithm of the share of the indicator of each subgroup, the share of the indicator of each subgroup and the population share of each subgroup. TI may be easily interpreted when multiplied by 1000. For more information on this inequality measure see Schlotheuber, A., & Hosseinpoor, A. R. (2022) below.

Interpretation: Greater absolute values indicate higher levels of inequality. TI is zero if there is no inequality. TI is more sensitive to differences further from the setting average (by the use of the logarithm).

Type of summary measure: Complex; relative; weighted

Applicability: Non-ordered; more than two subgroups

Warning: The confidence intervals are approximate and might be biased. See Ahn J. et al. (2018) below for further information on the standard error formula.

Value

The estimated TI value, corresponding estimated standard error, and confidence interval as a `data.frame`.

References

Schlotheuber, A., & Hosseinpoor, A. R. (2022). Summary measures of health inequality: A review of existing measures and their application. International Journal of Environmental Research and Public Health, 19 (6), 3697.

Ahn J, Harper S, Yu M, Feuer EJ, Liu B, Luta G. (2018). Variance Estimation and Confidence Intervals for 11 Commonly Used Health Disparity Measures. JCO Clin Cancer Inform. Dec;2:1–19.

Examples

```
# example code
data(NonorderedSample)
head(NonorderedSample)
with(NonorderedSample,
    ti(pop = population,
       est = estimate,
       se = se
       )
    )
```

Index