

# Package: **healthequal** (via **r-universe**)

November 25, 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

**RoxygenNote** 7.3.1

**NeedsCompilation** no

**Author** Daniel A. Antiporta [aut]  
(<<https://orcid.org/0000-0002-5143-3776>>), Patricia Menéndez [aut] (<<https://orcid.org/0000-0003-0701-6315>>), Katherine Kirkby [aut, cre] (<<https://orcid.org/0000-0002-6881-409X>>), Ahmad Hosseinpoor [aut] (<<https://orcid.org/0000-0001-7322-672X>>), World Health Organization [cph]

**Maintainer** Katherine Kirkby <kirkbyk@who.int>

**Repository** CRAN

**Date/Publication** 2024-11-25 11:40:08 UTC

**Config/pak/sysreqs** make libicu-dev

## Contents

aci . . . . .	2
bgsd . . . . .	4
bgv . . . . .	6
covar . . . . .	8
d . . . . .	9
idisu . . . . .	11
idisw . . . . .	13
IndividualSample . . . . .	15
mbu . . . . .	16
mbw . . . . .	18
mdmu . . . . .	20
mdmw . . . . .	22
mdru . . . . .	23
mdrw . . . . .	25
mld . . . . .	27
NonorderedSample . . . . .	28
NonorderedSampleMultipleind . . . . .	29
OrderedSample . . . . .	31
OrderedSampleMultipleind . . . . .	32
paf . . . . .	33
parisk . . . . .	35
r . . . . .	38
rci . . . . .	40
rii . . . . .	42
sii . . . . .	44
ti . . . . .	46
<b>Index</b>	<b>48</b>

---

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**

<code>est</code>	The indicator estimate. Estimates must be available for all subgroups/individuals (unless <code>force=TRUE</code> ).
<code>subgroup_order</code>	The order of subgroups/individuals in an increasing sequence.
<code>pop</code>	For disaggregated data, the number of people within each subgroup. This must be available for all subgroups.
<code>weight</code>	The individual sampling weight, for individual-level data from a survey. This must be available for all individuals.
<code>psu</code>	Primary sampling unit, for individual-level data from a survey.
<code>strata</code>	Strata, for individual-level data from a survey.
<code>fpc</code>	Finite population correction, for individual-level data from a survey where sample size is large relative to population size.
<code>lmin</code>	Minimum limit for bounded indicators (i.e., variables that have a finite upper and/or lower limit).
<code>lmax</code>	Maximum limit for bounded indicators (i.e., variables that have a finite upper and/or lower limit).
<code>conf.level</code>	Confidence level of the interval. Default is 0.95 (95%).
<code>force</code>	TRUE/FALSE statement to force calculation with missing indicator estimate values.
<code>...</code>	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 (2022) below.

**Interpretation:** ACI is 0 if there is no inequality. The larger the absolute value of ACI, the higher the level of inequality. Positive values indicate a concentration of the indicator among advantaged subgroups, and negative values indicate a concentration of the indicator among disadvantaged subgroups.

**Type of summary measure:** Complex; absolute; weighted

**Applicability:** Ordered dimension of inequality with 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, AR. Summary measures of health inequality: A review of existing measures and their application. *Int J Environ Res Public Health*. 2022;19(6):3697. doi:10.3390/ijerph19063697.

### Examples

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

## 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(
  est,
  se = NULL,
  pop,
  scaleval = NULL,
  sim = NULL,
  seed = 123456,
  force = FALSE,
  ...
)
```

## Arguments

<code>est</code>	The subgroup estimate. Estimates must be available for at least 85% of subgroups.
<code>se</code>	The standard error of the subgroup estimate. If this is missing, 95% confidence intervals cannot be calculated.
<code>pop</code>	The number of people within each subgroup. Population size must be available for all subgroups.
<code>scaleval</code>	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. If this is missing, 95% confidence intervals cannot be calculated.
<code>sim</code>	The number of simulations to estimate confidence intervals. Default is 100.
<code>seed</code>	The random number generator (RNG) state for the confidence interval simulation. Default is 123456.
<code>force</code>	TRUE/FALSE statement to force calculation when more than 85% of subgroup estimates are missing.
<code>...</code>	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 Schlottheuber (2022) below.

95% confidence intervals are calculated using a Monte Carlo simulation-based method. The dataset is simulated a large number of times (e.g. 100), with the mean and standard error of each simulated dataset being the same as the original dataset. BGSD is calculated for each of the simulated sample datasets. The 95% confidence intervals are based on the 2.5th and 97.5th percentiles of the BGSD results. See Ahn (2019) below for further information.

**Interpretation:** BGSD has only positive values, with larger values indicating higher levels of inequality. BGSD is 0 if there is no inequality. BGSD has the same unit as the health indicator.

**Type of summary measure:** Complex; absolute; weighted

**Applicability:** Non-ordered dimensions of inequality with more than two subgroups

### Value

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

### References

Schlotheuber, A, Hosseinpoor, AR. Summary measures of health inequality: A review of existing measures and their application. *Int J Environ Res Public Health*. 2022;19(6):3697. doi:10.3390/ijerph19063697.

Ahn J, Harper S, Yu M, Feuer EJ, Liu B. Improved Monte Carlo methods for estimating confidence intervals for eleven commonly used health disparity measures. *PLoS One*. 2019 Jul 1;14(7).

### Examples

```
# example code
data(NonorderedSample)
head(NonorderedSample)
with(NonorderedSample,
     bgsd(est = estimate,
          se = se,
          pop = population,
          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(est, se = NULL, pop, conf.level = 0.95, ...)
```

### Arguments

<code>est</code>	The subgroup estimate. Estimates must be available for at least 85% of subgroups.
<code>se</code>	The standard error of the subgroup estimate. If this is missing, 95% confidence intervals cannot be calculated.

pop	The number of people within each subgroup. Population size must be available for all subgroups.
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 (2022) below.

**Interpretation:** BGV has only positive values, with larger values indicating higher levels of inequality. BGV is 0 if there is no inequality. BGV is reported as the squared unit of the indicator. BGV is more sensitive to outlier estimates as it gives more weight to the estimates that are further from the setting average.

**Type of summary measure:** Complex; absolute; weighted

**Applicability:** Non-ordered dimensions of inequality with more than two subgroups

**Warning:** The confidence intervals are approximate and might be biased. See Ahn (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, AR. Summary measures of health inequality: A review of existing measures and their application. *Int J Environ Res Public Health*. 2022;19(6):3697. doi:10.3390/ijerph19063697.

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;2:1-19. doi:10.1200/CCI.18.00031.

## Examples

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

---

covar	<i>Coefficient of variation (COV)</i>
-------	---------------------------------------

---

### 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(
  est,
  se = NULL,
  pop,
  scaleval = NULL,
  sim = NULL,
  seed = 123456,
  force = FALSE,
  ...
)
```

### 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 cannot be calculated.
pop	The number of people within each subgroup. 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. If this is missing, 95% confidence intervals cannot be calculated.
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.
force	TRUE/FALSE statement to force calculation when more than 85% of subgroup estimates are missing.
...	Further arguments passed to or from other methods.

### Details

COV is calculated by dividing the between-group standard deviation (BGSD) by the setting average, multiplied 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 (2022) below.

95% confidence intervals are calculated using a Monte Carlo simulation-based method. The dataset is simulated a large number of times (e.g. 100), with the mean and standard error of each simulated dataset being the same as the original dataset. COV is calculated for each of the simulated sample datasets. The 95% confidence intervals are based on the 2.5th and 97.5th percentiles of the COV results. See Ahn (2019) below for further information.

**Interpretation:** COV only has positive values, with larger values indicating higher levels of inequality. COV is 0 if there is no inequality. COV has no unit.

**Type of summary measure:** Complex; relative; weighted

**Applicability:** Non-ordered dimensions of inequality with more than two subgroups

### Value

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

### References

Schlotheuber, A, Hosseinpoor, AR. Summary measures of health inequality: A review of existing measures and their application. *Int J Environ Res Public Health*. 2022;19(6):3697. doi:10.3390/ijerph19063697.

Ahn J, Harper S, Yu M, Feuer EJ, Liu B. Improved Monte Carlo methods for estimating confidence intervals for eleven commonly used health disparity measures. *PLoS One*. 2019 Jul 1;14(7).

### Examples

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

---

d

*Difference (D)*


---

### Description

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

**Usage**

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

**Arguments**

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

**Details**

D is calculated as  $D = y_1 - y_2$ , where  $y_1$  and  $y_2$  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). Please see specifications of how  $y_1$  and  $y_2$  are identified below.

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 0 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 et al. (2018) below for further information about the standard error formula.

## Value

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

## References

Schlottheuber, A, Hosseinpoor, AR. Summary measures of health inequality: A review of existing measures and their application. *Int J Environ Res Public Health*. 2022;19(6):3697. doi:10.3390/ijerph19063697.

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;2:1-19. doi:10.1200/CCI.18.00031.

## 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(
  est,
  se = NULL,
  pop = NULL,
  scaleval = NULL,
  setting_average = NULL,
  sim = NULL,
  seed = 123456,
  force = FALSE,
  ...
)

```

**Arguments**

<code>est</code>	The subgroup estimate. Estimates must be available for at least 85% of subgroups.
<code>se</code>	The standard error of the subgroup estimate. If this is missing, 95% confidence intervals cannot be calculated.
<code>pop</code>	The number of people within each subgroup. Population size must be available for all subgroups.
<code>scaleval</code>	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. If this is missing, 95% confidence intervals cannot be calculated.
<code>setting_average</code>	The overall indicator average for the setting of interest. Setting average must be unique for each setting, year and indicator combination. If population ( <code>pop</code> ) is not specified for all subgroups, the setting average is used for the calculation.
<code>sim</code>	The number of simulations to estimate 95% confidence intervals. Default is 100.
<code>seed</code>	The random number generator (RNG) state for the 95% confidence interval simulation. Default is 123456.
<code>force</code>	TRUE/FALSE statement to force calculation when more than 85% of subgroup estimates are missing.
<code>...</code>	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 inequity measure see Schlottheuber (2022) below.

95% confidence intervals are calculated using a Monte Carlo simulation-based method. The dataset is simulated a large number of times (e.g. 100), with the mean and standard error of each simulated dataset being the same as the original dataset. IDISU is calculated for each of the simulated sample datasets. The 95% confidence intervals are based on the 2.5th and 97.5th percentiles of the IDISU results. See Ahn (2019) below for further information.

**Interpretation:** IDISU has only positive values, with larger values indicating higher levels of inequality. IDISU is 0 if there is no inequality. IDISU has no unit.

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

**Applicability:** Non-ordered dimensions of inequality with more than two subgroups

### Value

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

### References

Schlotheuber, A, Hosseinpoor, AR. Summary measures of health inequality: A review of existing measures and their application. *Int J Environ Res Public Health*. 2022;19(6):3697. doi:10.3390/ijerph19063697.

Ahn J, Harper S, Yu M, Feuer EJ, Liu B. Improved Monte Carlo methods for estimating confidence intervals for eleven commonly used health disparity measures. *PLoS One*. 2019 Jul 1;14(7).

### Examples

```
# example code
data(NonorderedSample)
head(NonorderedSample)
with(NonorderedSample,
     idisu(est = estimate,
           se = se,
           pop = population,
           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(
  est,
  se = NULL,
  pop,
  scaleval = NULL,
  sim = NULL,
  seed = 123456,
  force = FALSE,
  ...
)
```

### 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 cannot be calculated.
pop	The number of people within each subgroup. 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. If this is missing, 95% confidence intervals cannot be calculated.
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.
force	TRUE/FALSE statement to force calculation when more than 85% of subgroup estimates are missing.
...	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 (2022) below.

95% confidence intervals are calculated using a Monte Carlo simulation-based method. The dataset is simulated a large number of times (e.g. 100), with the mean and standard error of each simulated dataset being the same as the original dataset. IDISW is calculated for each of the simulated sample datasets. The 95% confidence intervals are based on the 2.5th and 97.5th percentiles of the IDISW results. See Ahn (2019) below for further information.

**Interpretation:** IDISW has only positive values, with larger values indicating higher levels of inequality. IDISW is 0 if there is no inequality. IDISW has no unit.

**Type of summary measure:** Complex; relative; weighted

**Applicability:** Non-ordered dimensions of inequality with more than two subgroups

### Value

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

### References

- Schlotheuber, A, Hosseinpoor, AR. Summary measures of health inequality: A review of existing measures and their application. *Int J Environ Res Public Health*. 2022;19(6):3697. doi:10.3390/ijerph19063697.
- Ahn J, Harper S, Yu M, Feuer EJ, Liu B. Improved Monte Carlo methods for estimating confidence intervals for eleven commonly used health disparity measures. *PLoS One*. 2019 Jul 1;14(7).

## Examples

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

---

IndividualSample	<i>World Health Organization (WHO)</i>
------------------	--

---

## 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 *dtp3* or (0) if they 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	<i>Mean difference from best-performing subgroup (unweighted) (MDBU)</i>
------	--

---

## 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 best-performing subgroup. The best-performing subgroup is the subgroup with the highest value in the case of favourable indicators and the subgroup with the lowest value in the case of adverse indicators.

## Usage

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



```

    force = FALSE,
    ...
  )

```

### Arguments

<code>est</code>	The subgroup estimate. Estimates must be available for at least 85% of subgroups.
<code>se</code>	The standard error of the subgroup estimate. If this is missing, 95% confidence intervals cannot be calculated.
<code>favourable_indicator</code>	Records whether the indicator is favourable (1) or adverse (0). Favourable indicators measure desirable health events where the ultimate goal is to achieve a maximum level (such as skilled birth attendance). Adverse indicators measure undesirable health events where the ultimate goal is to achieve a minimum level (such as under-five mortality rate).
<code>scaleval</code>	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. If this is missing, 95% confidence intervals cannot be calculated.
<code>sim</code>	The number of simulations to estimate 95% confidence intervals. Default is 100.
<code>seed</code>	The random number generator (RNG) state for the 95% confidence interval simulation. Default is 123456.
<code>force</code>	TRUE/FALSE statement to force calculation when more than 85% of subgroup estimates are missing.
<code>...</code>	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. All subgroups are weighted equally. For more information on this inequality measure see Schlotheuber (2022) below.

95% confidence intervals are calculated using a Monte Carlo simulation-based method. The dataset is simulated a large number of times (e.g. 100), with the mean and standard error of each simulated dataset being the same as the original dataset. MDBU is calculated for each of the simulated sample datasets. The 95% confidence intervals are based on the 2.5th and 97.5th percentiles of the MDBU results. See Ahn (2019) below for further information.

**Interpretation:** MDBU only has positive values, with larger values indicating higher levels of inequality. MDBU is 0 if there is no inequality. MDBU has the same unit as the indicator.

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

**Applicability:** Non-ordered dimensions of inequality with more than two subgroups

### Value

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

## References

- Schlotheuber, A, Hosseinpoor, AR. Summary measures of health inequality: A review of existing measures and their application. *Int J Environ Res Public Health*. 2022;19(6):3697. doi:10.3390/ijerph19063697.
- Ahn J, Harper S, Yu M, Feuer EJ, Liu B. Improved Monte Carlo methods for estimating confidence intervals for eleven commonly used health disparity measures. *PLoS One*. 2019 Jul 1;14(7).

## Examples

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

---

mdbw	<i>Mean difference from best-performing subgroup (weighted) (MDBW)</i>
------	--

---

## 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. The best-performing subgroup is the subgroup with the highest value in the case of favourable indicators and the subgroup with the lowest value in the case of adverse indicators.

## Usage

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

## Arguments

- |                  |   |
|------------------|---|
| <code>est</code> | The subgroup estimate. Estimates must be available for at least 85% of subgroups.                               |
| <code>se</code>  | The standard error of the subgroup estimate. If this is missing, 95% confidence intervals cannot be calculated. |

pop	The number of people within each subgroup. Population size must be available for all subgroups.
favourable_indicator	Records whether the indicator is favourable (1) or adverse (0). Favourable indicators measure desirable health events where the ultimate goal is to achieve a maximum level (such as skilled birth attendance). Adverse 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. If this is missing, 95% confidence intervals cannot be calculated.
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.
force	TRUE/FALSE statement to force calculation when more than 85% of subgroup estimates are missing.
...	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. Subgroups are weighted according to their population share. For more information on this inequality measure see Schlotheuber (2022) below.

95% confidence intervals are calculated using a Monte Carlo simulation-based method. The dataset is simulated a large number of times (e.g. 100), with the mean and standard error of each simulated dataset being the same as the original dataset. MDBW is calculated for each of the simulated sample datasets. The 95% confidence intervals are based on the 2.5th and 97.5th percentiles of the MDBW results. See Ahn (2019) below for further information.

**Interpretation:** MDBW only has positive values, with larger values indicating higher levels of inequality. MDBW is 0 if there is no inequality. MDBW has the same unit as the indicator.

**Type of summary measure:** Complex; absolute; weighted

**Applicability:** Non-ordered dimensions of inequality with 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(est = estimate,
          se = se,
          pop = population,
          favourable_indicator,
          scaleval = indicator_scale))
```

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 mean (e.g. the national average).

**Usage**

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

**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 cannot be calculated.
pop	The number of people within each subgroup. 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. If this is missing, 95% confidence intervals cannot be calculated.

setting_average	The overall indicator average for the setting of interest. Setting average must be unique for each setting, year and indicator combination. If population (pop) is not specified for all subgroups, the setting average is used for the calculation.
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.
force	TRUE/FALSE statement to force calculation when more than 85% of subgroup estimates are missing.
...	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 mean, divided by the number of subgroups. All subgroups are weighted equally. For more information on this inequality measure see Schlotheuber (2022) below.

95% confidence intervals are calculated using a Monte Carlo simulation-based method. The dataset is simulated a large number of times (e.g. 100), with the mean and standard error of each simulated dataset being the same as the original dataset. MDMU is calculated for each of the simulated sample datasets. The 95% confidence intervals are based on the 2.5th and 97.5th percentiles of the MDMU results. See Ahn (2019) below for further information.

**Interpretation:** MDMU only has positive values, with larger values indicating higher levels of inequality. MDMU is 0 if there is no inequality. MDMU has the same unit as the indicator.

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

**Applicability:** Non-ordered dimensions of inequality with more than two subgroups

### Value

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

### References

Schlotheuber, A, Hosseinpoor, AR. Summary measures of health inequality: A review of existing measures and their application. *Int J Environ Res Public Health*. 2022;19(6):3697. doi:10.3390/ijerph19063697.

Ahn J, Harper S, Yu M, Feuer EJ, Liu B. Improved Monte Carlo methods for estimating confidence intervals for eleven commonly used health disparity measures. *PLoS One*. 2019 Jul 1;14(7).

### Examples

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

---

mdmw	<i>Mean difference from mean (weighted) (MDMW)</i>
------	--

---

### Description

The mean difference from mean (MDM) is an absolute measure of inequality that shows the mean difference between each subgroup and the mean (e.g. the national average).

### Usage

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

### Arguments

<code>est</code>	The subgroup estimate. Estimates must be available for at least 85% of subgroups.
<code>se</code>	The standard error of the subgroup estimate. If this is missing, 95% confidence intervals cannot be calculated.
<code>pop</code>	The number of people within each subgroup. Population size must be available for all subgroups.
<code>scaleval</code>	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. If this is missing, 95% confidence intervals cannot be calculated.
<code>sim</code>	The number of simulations to estimate 95% confidence intervals. Default is 100.
<code>seed</code>	The random number generator (RNG) state for the 95% confidence interval simulation. Default is 123456.
<code>force</code>	TRUE/FALSE statement to force calculation when more than 85% of subgroup estimates are missing.
<code>...</code>	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 mean. Absolute differences are weighted by each subgroup's population share. For more information on this inequality measure see Schlotheuber (2022) below.

95% confidence intervals are calculated using a Monte Carlo simulation-based method. The dataset is simulated a large number of times (e.g. 100), with the mean and standard error of each simulated dataset being the same as the original dataset. MDMW is calculated for each of the simulated sample datasets. The 95% confidence intervals are based on the 2.5th and 97.5th percentiles of the MDMW results. See Ahn (2019) below for further information.

**Interpretation:** MDMW only has positive values, with larger values indicating higher levels of inequality. MDMW is 0 if there is no inequality. MDMW has the same unit as the indicator.

**Type of summary measure:** Complex; absolute; weighted

**Applicability:** Non-ordered dimensions of inequality with more than two subgroups

### Value

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

### References

Schlotheuber, A, Hosseinpoor, AR. Summary measures of health inequality: A review of existing measures and their application. *Int J Environ Res Public Health*. 2022;19(6):3697. doi:10.3390/ijerph19063697.

Ahn J, Harper S, Yu M, Feuer EJ, Liu B. Improved Monte Carlo methods for estimating confidence intervals for eleven commonly used health disparity measures. *PLoS One*. 2019 Jul 1;14(7).

### Examples

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

---

 mdru

*Mean difference from a reference point (unweighted) (MDRU)*

---

### Description

The mean difference from a reference point (MDR) is an absolute measure of inequality that shows the mean difference between each population subgroup and a defined reference subgroup (e.g. the capital city or region for data disaggregated by subnational regions).

**Usage**

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

**Arguments**

<code>est</code>	The subgroup estimate. Estimates must be available for at least 85% of subgroups.
<code>se</code>	The standard error of the subgroup estimate. If this is missing, 95% confidence intervals cannot be calculated.
<code>scaleval</code>	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. If this is missing, 95% confidence intervals cannot be calculated.
<code>reference_subgroup</code>	Identifies a reference subgroup with the value of 1.
<code>sim</code>	The number of simulations to estimate 95% confidence intervals. Default is 100.
<code>seed</code>	The random number generator (RNG) state for the 95% confidence interval simulation. Default is 123456.
<code>force</code>	TRUE/FALSE statement to force calculation when more than 85% of subgroup estimates are missing.
<code>...</code>	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. All subgroups are weighted equally. For more information on this inequality measure see Schlotheuber (2022) below.

95% confidence intervals are calculated using a Monte Carlo simulation-based method. The dataset is simulated a large number of times (e.g. 100), with the mean and standard error of each simulated dataset being the same as the original dataset. MDRU is calculated for each of the simulated sample datasets. The 95% confidence intervals are based on the 2.5th and 97.5th percentiles of the MDRU results. See Ahn (2019) below for further information.

**Interpretation:** MDRU only has positive values, with larger values indicating higher levels of inequality. MDRU is 0 if there is no inequality. MDRU has the same unit as the indicator.

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

**Applicability:** Non-ordered dimensions of inequality with more than two subgroups



**Value**

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

**References**

Schlotheuber, A, Hosseinpoor, AR. Summary measures of health inequality: A review of existing measures and their application. *Int J Environ Res Public Health*. 2022;19(6):3697. doi:10.3390/ijerph19063697.

Ahn J, Harper S, Yu M, Feuer EJ, Liu B. Improved Monte Carlo methods for estimating confidence intervals for eleven commonly used health disparity measures. *PLoS One*. 2019 Jul 1;14(7).

**Examples**

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

---

 mdrw

*Mean difference from a reference point (weighted) (MDRW)*

---

**Description**

The mean difference from a reference point (MDR) is an absolute measure of inequality that shows the mean difference between each population subgroup and a defined reference subgroup (e.g. the capital city or region for data disaggregated by subnational regions).

**Usage**

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

## 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 cannot be calculated.
pop	The number of people within each subgroup. 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. If this is missing, 95% confidence intervals cannot be calculated.
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.
force	TRUE/FALSE statement to force calculation when more than 85% of subgroup estimates are missing.
...	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 (2022) below.

95% confidence intervals are calculated using a Monte Carlo simulation-based method. The dataset is simulated a large number of times (e.g. 100), with the mean and standard error of each simulated dataset being the same as the original dataset. MDRW is calculated for each of the simulated sample datasets. The 95% confidence intervals are based on the 2.5th and 97.5th percentiles of the MDRW results. See Ahn (2019) below for further information.

**Interpretation:** MDRW only has positive values, with larger values indicating higher levels of inequality. MDRW is 0 if there is no inequality. MDRW has the same unit as the indicator.

**Type of summary measure:** Complex; absolute; weighted

**Applicability:** Non-ordered dimensions of inequality with more than two subgroups

## Value

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

## References

- Schlotheuber, A, Hosseinpoor, AR. Summary measures of health inequality: A review of existing measures and their application. *Int J Environ Res Public Health*. 2022;19(6):3697. doi:10.3390/ijerph19063697.
- Ahn J, Harper S, Yu M, Feuer EJ, Liu B. Improved Monte Carlo methods for estimating confidence intervals for eleven commonly used health disparity measures. *PLoS One*. 2019 Jul 1;14(7).

## Examples

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

---

mld	<i>Mean log deviation (MLD)</i>
-----	---------------------------------

---

## 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(est, se = NULL, pop, conf.level = 0.95, force = FALSE, ...)
```

## 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 cannot be calculated.
pop	The number of people within each subgroup. Population size must be available for all subgroups.
conf.level	Confidence level of the interval. Default is 0.95 (95%).
force	TRUE/FALSE statement to force calculation when more than 85% of subgroup estimates are missing.
...	Further arguments passed to or from other methods.

## Details

MLD measures the extent to which the shares of the population and shares of the health indicator differ across subgroups, weighted by shares of the population. 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 (2022) below.

**Interpretation:** MLD is 0 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). MLD has no unit.

**Type of summary measure:** Complex; relative; weighted

**Applicability:** Non-ordered dimensions of inequality with more than two subgroups

**Warning:** The confidence intervals are approximate and might be biased. See Ahn (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, AR. Summary measures of health inequality: A review of existing measures and their application. *Int J Environ Res Public Health*. 2022;19(6):3697. doi:10.3390/ijerph19063697.

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;2:1-19. doi:10.1200/CCI.18.00031.

### Examples

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

---

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	<i>World Health Organization (WHO)</i>
---------------	--

---

### 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	<i>Population attributable fraction (PAF)</i>
-----	---

---

### Description

Population attributable fraction (PAR) is a relative 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 point. The reference point refers to the most advantaged subgroup for ordered dimensions and the best-performing subgroup for non-ordered dimensions (i.e. the subgroup with the highest value for favourable indicators and the subgroup with the lowest value for adverse indicators).

### Usage

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

```

    force = FALSE,
    ...
)

```

### Arguments

<code>est</code>	The subgroup estimate. Estimates must be available for the two subgroups being compared.
<code>pop</code>	The number of people within each subgroup. Population size must be available for all subgroups.
<code>favourable_indicator</code>	Records whether the indicator is favourable (1) or adverse (0). Favourable indicators measure desirable health events where the ultimate goal is to achieve a maximum level (such as skilled birth attendance). Adverse indicators measure undesirable health events where the ultimate goal is to achieve a minimum level (such as under-five mortality rate).
<code>ordered_dimension</code>	Records whether the dimension is ordered (1) or non-ordered (0). Ordered dimensions have subgroup with a natural order (such as economic status). Non-ordered or binary dimensions do not have a natural order (such as subnational region or sex).
<code>subgroup_order</code>	The order of subgroups in an increasing sequence. Required if the dimension is ordered ( <code>ordered_dimension=1</code> ).
<code>setting_average</code>	The overall indicator average for the setting of interest. Setting average must be unique for each setting, year and indicator combination. If population ( <code>pop</code> ) is not specified for all subgroups, the setting average is used for the calculation.
<code>scaleval</code>	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.
<code>conf.level</code>	Confidence level of the interval. Default is 0.95 (95%).
<code>force</code>	TRUE/FALSE statement to force calculation when subgroup estimates are missing.
<code>...</code>	Further arguments passed to or from other methods.

### Details

PAF is calculated as the difference between the estimate for the reference subgroup and the mean (e.g. the national average), divided by the mean and multiplied by 100. For more information on this inequality measure see Schlotheuber (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.

**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 0 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 dimension of inequality with more than two subgroups

**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, AR. Summary measures of health inequality: A review of existing measures and their application. *Int J Environ Res Public Health*. 2022;19(6):3697. doi:10.3390/ijerph19063697.

Walter, SD. Calculation of attributable risks from epidemiological data. *Int J Epidemiol*. 1978 Jun 1;7(2):175-82. doi:10.1093/ije/7.2.175.

## Examples

```
# example code
data(OrderedSample)
head(OrderedSample)
with(OrderedSample,
      paf(est = estimate,
          pop = population,
          favourable_indicator = favourable_indicator,
          ordered_dimension = ordered_dimension,
          subgroup_order = subgroup_order,
          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 point. The reference point refers to the most advantaged subgroup for ordered dimensions and the best-performing subgroup for non-ordered dimensions (i.e. the subgroup with the highest value for favourable indicators and the subgroup with the lowest value for adverse indicators).

**Usage**

```

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

```

**Arguments**

<code>est</code>	The subgroup estimate. Estimates must be available for the two subgroups being compared.
<code>pop</code>	The number of people within each subgroup. Population size must be available for all subgroups.
<code>favourable_indicator</code>	Records whether the indicator is favourable (1) or adverse (0). Favourable indicators measure desirable health events where the ultimate goal is to achieve a maximum level (such as skilled birth attendance). Adverse indicators measure undesirable health events where the ultimate goal is to achieve a minimum level (such as under-five mortality rate).
<code>ordered_dimension</code>	Records whether the dimension is ordered (1) or non-ordered (0). Ordered dimensions have subgroup with a natural order (such as economic status). Non-ordered or binary dimensions do not have a natural order (such as subnational region or sex).
<code>subgroup_order</code>	The order of subgroups in an increasing sequence. Required if the dimension is ordered ( <code>ordered_dimension=1</code> ).
<code>setting_average</code>	The overall indicator average for the setting of interest. Setting average must be unique for each setting, year and indicator combination. If population ( <code>pop</code> ) is not specified for all subgroups, the setting average is used for the calculation.
<code>scaleval</code>	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.
<code>conf.level</code>	Confidence level of the interval. Default is 0.95 (95%).
<code>force</code>	TRUE/FALSE statement to force calculation when subgroup estimates are missing.
<code>...</code>	Further arguments passed to or from other methods.

## Details

PAR is calculated as the difference between the estimate for the reference subgroup and the mean (e.g. the national average). For more information on this inequality measure see Schlotheuber (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.

**Interpretation:** PAR assumes positive values for favourable indicators and negative values for adverse indicators. The larger the absolute value of PAR, the higher the level of inequality. PAR is 0 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 dimension of inequality with more than two subgroups

**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, AR. Summary measures of health inequality: A review of existing measures and their application. *Int J Environ Res Public Health*. 2022;19(6):3697. doi:10.3390/ijerph19063697.

Walter, SD. Calculation of attributable risks from epidemiological data. *Int J Epidemiol*. 1978 Jun 1;7(2):175-82. doi:10.1093/ije/7.2.175.

## Examples

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

---

r	<i>Ratio (R)</i>
---	------------------

---

### Description

The ratio (R) is a relative measure of inequality that shows the ratio of an indicator between two population subgroups. For more information on this inequality measure see Schlottheuber (2022) below.

### Usage

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

### Arguments

<code>est</code>	The subgroup estimate. Estimates must be available for the two subgroups being compared.
<code>se</code>	The standard error of the subgroup estimate. If this is missing, confidence intervals of D cannot be calculated.
<code>favourable_indicator</code>	Records whether the indicator is favourable (1) or adverse (0). Favourable indicators measure desirable health events where the ultimate goal is to achieve a maximum level (such as skilled birth attendance). Adverse indicators measure undesirable health events where the ultimate goal is to achieve a minimum level (such as under-five mortality rate).
<code>ordered_dimension</code>	Records whether the dimension is ordered (1) or non-ordered (0). Ordered dimensions have subgroup with a natural order (such as economic status). Non-ordered or binary dimensions do not have a natural order (such as subnational region or sex).
<code>subgroup_order</code>	The order of subgroups in an increasing sequence. Required if the dimension is ordered ( <code>ordered_dimension=1</code> ).
<code>reference_subgroup</code>	Identifies a reference subgroup with the value of 1, if the dimension is non-ordered or binary.
<code>conf.level</code>	Confidence level of the interval. Default is 0.95 (95%).
<code>...</code>	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). Please see specifications of how  $y1$  and  $y2$  are identified below.

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 1, the higher the level of inequality. R is 1 if there is no inequality. R is a multiplicative measure and therefore results should be displayed on a logarithmic scale. Values larger than 1 are equivalent in magnitude to their reciprocal values smaller than 1 (e.g. a value of 2 is equivalent in magnitude to a value of 0.5).

**Type of summary measure:** Simple; relative; unweighted

**Applicability:** Any dimension of inequality

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

## Value

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

## References

Schlotheuber, A, Hosseinpoor, AR. Summary measures of health inequality: A review of existing measures and their application. *Int J Environ Res Public Health*. 2022;19(6):3697. doi:10.3390/ijerph19063697.

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;2:1-19. doi:10.1200/CCI.18.00031.

## Examples

```
# example code
data(NonorderedSample)
head(NonorderedSample)
with(NonorderedSample,
     r(est = estimate,
       se = se,
       favourable_indicator = 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 indicates the extent to which an indicator is concentrated among disadvantaged or advantaged subgroups, on a relative scale.

### Usage

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

### Arguments

<code>est</code>	The indicator estimate. Estimates must be available for all subgroups/individuals (unless <code>force=TRUE</code> ).
<code>subgroup_order</code>	The order of subgroups/individuals in an increasing sequence.
<code>pop</code>	For disaggregated data, the number of people within each subgroup. This must be available for all subgroups.
<code>weight</code>	The individual sampling weight, for individual-level data from a survey. This must be available for all individuals.
<code>psu</code>	Primary sampling unit, for individual-level data from a survey.
<code>strata</code>	Strata, for individual-level data from a survey.
<code>fpc</code>	Finite population correction, for individual-level data from a survey where sample size is large relative to population size.
<code>method</code>	Normalisation method for bounded indicators. Options available are Wagstaff ( <code>wagstaff</code> ) or Erreygers ( <code>erreygers</code> ) correction. Must be used in conjunction with minimum and maximum limits ( <code>lmin</code> and <code>lmax</code> ).
<code>lmin</code>	Minimum limit for bounded indicators (i.e., variables that have a finite upper and/or lower limit).



<code>lmax</code>	Maximum limit for bounded indicators (i.e., variables that have a finite upper and/or lower limit).
<code>conf.level</code>	Confidence level of the interval. Default is 0.95 (95%).
<code>force</code>	TRUE/FALSE statement to force calculation with missing indicator estimate values.
<code>...</code>	Further arguments passed to or from other methods.

## Details

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.

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 (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. Positive values indicate a concentration of the indicator among advantaged subgroups, and negative values indicate a concentration of the indicator among disadvantaged subgroups. RCI is 0 if there is no inequality.

**Type of summary measure:** Complex; relative; weighted

**Applicability:** Ordered dimension of inequality with 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

- Erreygers G. Correcting the Concentration Index. *J Health Econ.* 2009;28(2):504-515. doi:10.1016/j.jhealeco.2008.02.003.
- Schlotheuber, A, Hosseinpoor, AR. Summary measures of health inequality: A review of existing measures and their application. *Int J Environ Res Public Health.* 2022;19(6):3697. doi:10.3390/ijerph19063697.
- Wagstaff A. The bounds of the concentration index when the variable of interest is binary, with an application to immunization inequality. *Health Econ.* 2011;20(10):1155-1160. doi:10.1002/hec.1752.

## Examples

```
# example code
data(IndividualSample)
head(IndividualSample)
with(IndividualSample,
     rci(est = sba,
```

```

        subgroup_order = subgroup_order,
        weight = weight,
        psu = psu,
        strata = strata))
# example code
data(OrderedSample)
head(OrderedSample)
with(OrderedSample,
     rci(est = estimate,
         subgroup_order = subgroup_order,
         pop = population))

```

rii

*Relative index of inequality (RII)***Description**

The relative index of inequality (RII) is a relative measure of inequality that represents the ratio of predicted values of an indicator between the most advantaged and most disadvantaged subgroups, obtained by fitting a regression model.

**Usage**

```

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

```

**Arguments**

<code>est</code>	The indicator estimate. Estimates must be available for all subgroups/individuals (unless <code>force=TRUE</code> ).
<code>subgroup_order</code>	The order of subgroups/individuals in an increasing sequence.
<code>pop</code>	For disaggregated data, the number of people within each subgroup. This must be available for all subgroups.
<code>weight</code>	The individual sampling weight, for individual-level data from a survey. This must be available for all individuals.
<code>psu</code>	Primary sampling unit, for individual-level data from a survey.

strata	Strata, for individual-level data from a survey.
fpc	Finite population correction, for individual-level data from a survey where 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 (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

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.

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 indicator of interest is then regressed against this relative rank using an appropriate regression model, and the predicted values of the indicator are calculated for the two extremes (rank 1 and rank 0). RII is calculated as the ratio between the predicted values at rank 1 and rank 0 (covering the entire distribution). For more information on this inequality measure see Schlotheuber (2022) below.

The default regression model used is a generalized linear model with logit link. In logistic regression, the relationship between the indicator and the subgroup rank is not assumed to be linear and, due to the logit link, the predicted values from the regression model will be bounded between 0 and 1 (which is ideal for indicators measured as percentages). Specify `Linear=TRUE` to use a linear regression model, which may be more appropriate for indicators without a 0-1 or 0-100% scale.

**Interpretation:** RII takes only positive values. RII has the value of 1 if there is no inequality. Values larger than 1 indicate the level of the indicator is higher among advantaged subgroups, and values lower than 1 indicate the level of the indicator is higher among disadvantaged subgroups. Note that this results in different interpretations for favourable and adverse indicators. RII is a multiplicative measure and therefore results should be displayed on a logarithmic scale. Values larger than 1 are equivalent in magnitude to their reciprocal values smaller than 1 (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 dimension of inequality with 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, AR. Summary measures of health inequality: A review of existing measures and their application. *Int J Environ Res Public Health*. 2022;19(6):3697. doi:10.3390/ijerph19063697.

**Examples**

```
# example code 1
data(IndividualSample)
head(IndividualSample)
with(IndividualSample,
     rii(est = sba,
         subgroup_order = subgroup_order,
         weight = weight,
         psu = psu,
         strata = strata))
# example code 2
data(OrderedSample)
head(OrderedSample)
with(OrderedSample,
     rii(est = estimate,
         subgroup_order = subgroup_order,
         pop = population))
```

sii

*Slope index of inequality (SII)***Description**

The slope index of inequality (SII) is an absolute measure of inequality that represents the difference in predicted values of an indicator between the most advantaged and most disadvantaged subgroups, obtained by fitting a regression model.

**Usage**

```
sii(
  est,
  subgroup_order,
  pop = 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	For disaggregated data, the number of people within each subgroup. This must be available for all subgroups.
weight	The individual sampling weight, for individual-level data from a survey. This must be available for all individuals.
psu	Primary sampling unit, for individual-level data from a survey.
strata	Strata, for individual-level data from a survey.
fpc	Finite population correction, for individual-level data from a survey where 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 (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

SII 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.

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, and the predicted values of the indicator are calculated for the two extremes (rank 1 and rank 0). SII is calculated as the difference between the predicted values at rank 1 and rank 0 (covering the entire distribution). For more information on this inequality measure see Schlotheuber (2022) below.

The default regression model used is a generalized linear model with logit link. In logistic regression, the relationship between the indicator and the subgroup rank is not assumed to be linear and, due to the logit link, the predicted values from the regression model will be bounded between 0 and 1 (which is ideal for indicators measured as percentages). Specify `Linear=TRUE` to use a linear regression model, which may be more appropriate for indicators without a 0-1 or 0-100% scale.

**Interpretation:** SII is 0 if there is no inequality. Greater absolute values indicate higher levels of inequality. Positive values indicate that the level of the indicator is higher among advantaged subgroups, while negative values indicate that the level of the indicator is higher among disadvantaged subgroups. Note that this results in different interpretations for favourable and adverse indicators.

**Type of summary measure:** Complex; absolute; weighted

**Applicability:** Ordered dimension of inequality with 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, AR. Summary measures of health inequality: A review of existing measures and their application. *Int J Environ Res Public Health*. 2022;19(6):3697. doi:10.3390/ijerph19063697.

## Examples

```
# example code 1
data(IndividualSample)
head(IndividualSample)
with(IndividualSample,
     sii(est = sba,
         subgroup_order = subgroup_order,
         weight = weight,
         psu = psu,
         strata = strata))

# example code 2
data(OrderedSample)
head(OrderedSample)
with(OrderedSample,
     sii(est = estimate,
         subgroup_order = subgroup_order,
         pop = population))
```

---

ti	<i>Theil index (TI)</i>
----	-------------------------

---

## 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(est, se = NULL, pop, conf.level = 0.95, force = FALSE, ...)
```

## Arguments

<code>est</code>	The subgroup estimate. Estimates must be available for at least 85% of subgroups.
<code>se</code>	The standard error of the subgroup estimate. If this is missing, 95% confidence intervals cannot be calculated.
<code>pop</code>	The number of people within each subgroup. Population size must be available for all subgroups.
<code>conf.level</code>	Confidence level of the interval. Default is 0.95 (95%).
<code>force</code>	TRUE/FALSE statement to force calculation when more than 85% of subgroup estimates are missing.
<code>...</code>	Further arguments passed to or from other methods.

## Details

TI measures the extent to which the shares of the population and shares of the health indicator differ across subgroups, weighted by shares of the health indicator. 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 (2022) below.

**Interpretation:** TI is 0 if there is no inequality. Greater absolute values indicate higher levels of inequality. TI is more sensitive to differences further from the setting average (by the use of the logarithm). TI has no unit.

**Type of summary measure:** Complex; relative; weighted

**Applicability:** Non-ordered dimensions of inequality with more than two subgroups

**Warning:** The confidence intervals are approximate and might be biased. See Ahn (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, AR. Summary measures of health inequality: A review of existing measures and their application. *Int J Environ Res Public Health*. 2022;19(6):3697. doi:10.3390/ijerph19063697.
- 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;2:1-19. doi:10.1200/CCI.18.00031.

## Examples

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

# Index

## \* datasets

- IndividualSample, [15](#)
- NonorderedSample, [28](#)
- NonorderedSampleMultipleind, [29](#)
- OrderedSample, [31](#)
- OrderedSampleMultipleind, [32](#)

[aci](#), [2](#)

[bgzd](#), [4](#)

[bgv](#), [6](#)

[covar](#), [8](#)

[d](#), [9](#)

[idisu](#), [11](#)

[idisw](#), [13](#)

[IndividualSample](#), [15](#)

[mdbu](#), [16](#)

[mdbw](#), [18](#)

[mdmu](#), [20](#)

[mdmw](#), [22](#)

[mdru](#), [23](#)

[mdrw](#), [25](#)

[mld](#), [27](#)

[NonorderedSample](#), [28](#)

[NonorderedSampleMultipleind](#), [29](#)

[OrderedSample](#), [31](#)

[OrderedSampleMultipleind](#), [32](#)

[paf](#), [33](#)

[parisk](#), [35](#)

[r](#), [38](#)

[rci](#), [40](#)

[rii](#), [42](#)

[sii](#), [44](#)

[ti](#), [46](#)