# Package 'epibyhand'

November 3, 2026

**Type** Package

**Title** Worked Derivations for Classical Epidemiological Measures

**Version** 0.3.0

**Description** Computes classical epidemiological measures and returns the complete worked derivation alongside the result: every intermediate quantity, the formula, and the formula with the observed numbers substituted in. Intended for teaching, for checking hand calculations, and for generating worked solutions in course materials. Scope is deliberately limited to methods a student can compute by hand on paper. Methods follow Mantel and Haenszel (1959) <doi:10.1093/jnci/22.4.719>, Greenland and Robins (1985, Biometrics 41, 55-68), Robins, Breslow and Greenland (1986, Biometrics 42, 311-323), and Breslow and Day (1980, IARC Scientific Publications No. 32).

**License** MIT + file LICENSE

**Encoding** UTF-8

**Language** en-GB

**Depends** R (>= 4.0.0)

**Imports** stats

**Suggests** testthat (>= 3.0.0), knitr, rmarkdown

**Config/testthat/edition** 3

**VignetteBuilder** knitr

**URL** https://github.com/rajsubediresearch/epibyhand

**BugReports** https://github.com/rajsubediresearch/epibyhand/issues

**Config/roxygen2/version** 8.1.1

**NeedsCompilation** no

**Author** Raj Subedi [aut, cre]

**Maintainer** Raj Subedi <rajsubediresearch@gmail.com>

**Repository** CRAN

**Date/Publication** 2026-11-03 03:40:02 UTC

# Contents

attributable_fraction . . . 2
check_work . . . 3
collapse_strata . . . 4
confint.epibyhand_derivation . . . 5
derivation . . . 5
derivation_step . . . 6
epi2x2 . . . 7
epi_strata . . . 8
estimate . . . 9
homogeneity . . . 9
mh_odds_ratio . . . 10
mh_risk_ratio . . . 11
odds_ratio . . . 12
predictive_value . . . 12
print.epi2x2 . . . 14
print.epibyhand_derivation . . . 14
print.epi_strata . . . 15
risk_difference . . . 15
risk_ratio . . . 16
steps_table . . . 16
to_latex . . . 17
to_markdown . . . 18

**Index** **19**

---

`attributable_fraction` *Attributable fraction, with worked derivation*

---

## Description

The proportion of disease that would be prevented by removing the exposure, either among the exposed (`among = "exposed"`) or in the whole population (`among = "population"`).

## Usage

```
attributable_fraction(
  x,
  ...,
  among = c("exposed", "population"),
  conf_level = 0.95
)
```

### Arguments

| | |
|---|---|
| x | An epi2x2, or the first of four cell counts. |
| ... | Further cell counts if x is a bare count. |
| among | "exposed" (the default) or "population". |
| conf_level | Confidence level for the interval. |

### Details

For the population fraction the derivation computes the same quantity three ways – directly from risks, by Levin's formula from exposure prevalence, and by Miettinen's formula from the proportion of cases exposed – and shows that they agree. They are not competing estimators; they are one quantity written three ways, which is easier to believe once seen numerically.

### Value

An epibyhand_derivation.

### References

Levin, M. L. (1953). The occurrence of lung cancer in man. *Acta Unio Internationalis Contra Cancrum*, 9(3), 531-541.

Miettinen, O. S. (1974). Proportion of disease caused or prevented by a given exposure, trait or intervention. *American Journal of Epidemiology*, 99(5), 325-332.

### Examples

```
attributable_fraction(epi2x2(36, 14, 30, 25))
attributable_fraction(epi2x2(36, 14, 30, 25), among = "population")
```

---

## check_work *Check a hand calculation against a derivation*

---

### Description

Compare a value you computed by hand against the final estimate, or against a named intermediate step. When the value does not match, check_work() searches every step of the derivation for one that does – so instead of "wrong", the student is told *where* they stopped or slipped.

### Usage

```
check_work(x, value, step = NULL, tol = 0.005)
```

### Arguments

| | |
|---|---|
| x | An epibyhand_derivation. |
| value | Numeric. The value you computed by hand. |
| step | Optional. Which step to compare against: a step number, or the symbol of a step such as "SE" or "R1". Defaults to the final estimate. |
| tol | Relative tolerance for the comparison. The default is loose enough to accept a value rounded to two decimal places. |

### Value

Invisibly, TRUE if the value matched the requested target.

### Examples

```
d <- odds_ratio(epi2x2(36, 14, 30, 25))
check_work(d, 2.14)
check_work(d, 2.571)
```

---

## collapse_strata *Collapse strata into a single 2 x 2 table*

---

### Description

Adds the strata cell by cell, discarding the stratifying variable. The result is the *crude* table – the one you would have had if you had never stratified. Comparing a crude estimate with its adjusted counterpart is how confounding is made visible.

### Usage

```
collapse_strata(x)
```

### Arguments

| | |
|---|---|
| x | An epi_strata. |

### Value

An epi2x2.

### Examples

```
collapse_strata(epi_strata(c(10, 20, 5, 40), c(30, 15, 20, 25)))
```

## confint.epibyhand_derivation *Extract the confidence interval from a derivation*

### Description

Confidence intervals in this package are computed when the measure is computed, because the derivation displays the arithmetic that produced them. `level` therefore cannot change the result here; supplying one that differs from the level the derivation was built at raises a warning. Set the level with the `conf_level` argument of the measure instead.

### Usage

```
## S3 method for class 'epibyhand_derivation'
confint(object, parm, level, ...)
```

### Arguments

| | |
|---|---|
| `object` | An `epibyhand_derivation`. |
| `parm` | Unused, for S3 consistency. |
| `level` | Ignored, with a warning if it differs from the level the derivation was built at. Use `conf_level` on the measure instead. |
| `...` | Unused. |

### Value

Numeric vector of length 2.

### Examples

```
d <- odds_ratio(epi2x2(36, 14, 30, 25), conf_level = 0.99)
confint(d)
```

## derivation *Create a derivation object*

### Description

Create a derivation object

### Usage

```
derivation(
  method,
  estimate,
  symbol = NULL,
  ci = NULL,
  conf_level = 0.95,
  data = NULL,
  steps = list(),
  notes = character()
)
```

### Arguments

| | |
|---|---|
| method | Character. Name of the measure, e.g. "Odds ratio". |
| estimate | Numeric. The final point estimate. |
| symbol | Character. Symbol for the estimate, e.g. "OR". |
| ci | Numeric length-2 vector, or NULL. |
| conf_level | Confidence level used for ci. |
| data | Optional. The input data object, printed above the steps. |
| steps | List of epibyhand_step objects. |
| notes | Character vector of assumption notes shown after the result. |

### Value

An object of class epibyhand_derivation.

---

## derivation_step *Create one step of a derivation*

---

### Description

A step is the atomic unit of a worked solution: a human-readable label, the symbolic formula, the same formula with observed values substituted in, and the numeric result.

### Usage

```
derivation_step(
  label,
  formula,
  substituted,
  result,
  symbol = NULL,
  note = NULL,
  table = NULL,
  latex = NULL
)
```

### Arguments

| | |
|---|---|
| label | Short description of what this step computes. |
| formula | Character. The formula in symbols, e.g. "a / (a + b)". |
| substituted | Character. The formula with numbers filled in. |
| result | Numeric. The value this step evaluates to. |
| symbol | Optional character. The symbol this step defines, e.g. "R1". |
| note | Optional character. A caveat or explanation attached to the step. |
| table | Optional data frame. Per-unit working shown as a grid beneath the formula. Used when a step aggregates over strata and a single substituted line would be unreadable. |
| latex | Optional character. The right-hand side of formula as LaTeX, e.g. "\\frac{a}{a + b}", used by to_markdown(), to_latex() and in Quarto or R Markdown documents. Built-in measures do not need it. Supply it for measures you write yourself; without it the plain-text formula is shown as upright text. |

### Value

An object of class epibyhand_step.

---

## epi2x2 *Build a 2 x 2 table*

---

### Description

Accepts either four cell counts or a 2 x 2 matrix/table. Orientation is exposure in rows, outcome in columns, with the index level first in each – the layout used in standard epidemiology texts.

### Usage

```
epi2x2(
  a,
  b = NULL,
  c = NULL,
  d = NULL,
  exposure = c("Exposed", "Unexposed"),
  outcome = c("Case", "Non-case")
)
```

### Arguments

| | |
|---|---|
| a | Exposed cases, or a 2 x 2 matrix or table. |
| b | Exposed non-cases. |
| c | Unexposed cases. |
| d | Unexposed non-cases. |
| exposure | Length-2 character vector labelling the rows. |
| outcome | Length-2 character vector labelling the columns. |

## Value

An object of class epi2x2.

## Examples

```
epi2x2(36, 14, 30, 25)
epi2x2(matrix(c(36, 14, 30, 25), nrow = 2, byrow = TRUE))
```

---

epi_strata *Build a set of stratified 2 x 2 tables*

---

## Description

Accepts a list of epi2x2() objects, a list of four-element count vectors, or a 2 x 2 x K array. Stratum labels are taken from the names of the list, the third dimnames of the array, or generated as "Stratum 1", and so on.

## Usage

```
epi_strata(
  ...,
  labels = NULL,
  exposure = c("Exposed", "Unexposed"),
  outcome = c("Case", "Non-case")
)
```

## Arguments

| | |
|---|---|
| ... | Strata: epi2x2 objects, length-4 numeric vectors of counts, or a single list or 2 x 2 x K array containing them. |
| labels | Optional character vector of stratum names. |
| exposure | Length-2 character vector labelling the rows. |
| outcome | Length-2 character vector labelling the columns. |

## Value

An object of class epi_strata.

## Examples

```
epi_strata(c(10, 20, 5, 40), c(30, 15, 20, 25), labels = c("Men", "Women"))
```

## estimate — *Extract the point estimate from a derivation*

### Description

Extract the point estimate from a derivation

### Usage

```
estimate(x)
```

### Arguments

| | |
|---|---|
| x | An epibyhand_derivation. |

### Value

Numeric scalar.

## homogeneity — *Breslow-Day test of homogeneity, with worked derivation*

### Description

Tests whether one common odds ratio can be assumed across strata – the assumption that makes mh_odds_ratio() meaningful. The statistic is a sum of per-stratum contributions, and those contributions are shown, so a single badly behaved stratum can be seen rather than inferred.

### Usage

```
homogeneity(x, ..., tarone = TRUE)
```

### Arguments

| | |
|---|---|
| x | An epi_strata, or arguments passed to epi_strata(). |
| ... | Further strata if x is not already an epi_strata. |
| tarone | Apply Tarone's correction. Defaults to TRUE. |

### Details

Tarone's correction is applied by default; without it the statistic is slightly too large.

### Value

An epibyhand_derivation.

## References

Breslow, N. E. and Day, N. E. (1980). *Statistical Methods in Cancer Research, Volume I: The Analysis of Case-Control Studies*. IARC Scientific Publications No. 32, Lyon.

Tarone, R. E. (1985). On heterogeneity tests based on efficient scores. *Biometrika*, 72(1), 91-95.

## Examples

```
homogeneity(epi_strata(c(10, 20, 5, 40), c(30, 15, 20, 25)))
```

---

`mh_odds_ratio` *Mantel-Haenszel odds ratio with worked derivation*

---

## Description

Pools stratum-specific odds ratios into a single adjusted estimate, showing the stratum weights explicitly and comparing the result against the crude estimate so that confounding is visible as arithmetic.

## Usage

```
mh_odds_ratio(x, ..., conf_level = 0.95)
```

## Arguments

| | |
|---|---|
| `x` | An `epi_strata`, or arguments passed to `epi_strata()`. |
| `...` | Further strata if `x` is not already an `epi_strata`. |
| `conf_level` | Confidence level for the interval. |

## Details

The confidence interval uses the Robins-Breslow-Greenland variance, which is valid both for a few large strata and for many small ones.

## Value

An `epibyhand_derivation`.

## References

Mantel, N. and Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute*, 22(4), 719-748. doi:10.1093/jnci/22.4.719

Robins, J., Breslow, N. and Greenland, S. (1986). Estimators of the Mantel-Haenszel variance consistent in both sparse data and large-strata limiting models. *Biometrics*, 42(2), 311-323.

## Examples

```
s <- epi_strata(c(10, 20, 5, 40), c(30, 15, 20, 25),
                labels = c("Men", "Women"))
mh_odds_ratio(s)
```

---

`mh_risk_ratio` *Mantel-Haenszel risk ratio with worked derivation*

---

## Description

Mantel-Haenszel risk ratio with worked derivation

## Usage

```
mh_risk_ratio(x, ..., conf_level = 0.95)
```

## Arguments

| | |
|---|---|
| `x` | An `epi_strata`, or arguments passed to `epi_strata()`. |
| `...` | Further strata if `x` is not already an `epi_strata`. |
| `conf_level` | Confidence level for the interval. |

## Value

An `epibyhand_derivation`.

## References

Greenland, S. and Robins, J. M. (1985). Estimation of a common effect parameter from sparse follow-up data. *Biometrics*, 41, 55-68.

## Examples

```
mh_risk_ratio(epi_strata(c(10, 20, 5, 40), c(30, 15, 20, 25)))
```

## odds_ratio *Odds ratio with worked derivation*

### Description

Odds ratio with worked derivation

### Usage

```
odds_ratio(x, ..., conf_level = 0.95)
```

### Arguments

| | |
|---|---|
| x | An epi2x2, or the first of four cell counts. |
| ... | Further cell counts if x is a bare count. |
| conf_level | Confidence level for the interval. |

### Value

An epibyhand_derivation.

### References

Woolf, B. (1955). On estimating the relation between blood group and disease. *Annals of Human Genetics*, 19, 251-253.

### Examples

```
odds_ratio(epi2x2(36, 14, 30, 25))
```

## predictive_value *Predictive value of a diagnostic test, with worked derivation*

### Description

Computes the positive or negative predictive value, showing sensitivity, specificity and prevalence as separate steps so that the dependence on prevalence is visible rather than asserted.

### Usage

```
predictive_value(
  x,
  ...,
  which = c("positive", "negative"),
  prevalence = NULL,
  conf_level = 0.95
)
```

## Arguments

| | |
|---|---|
| x | An epi2x2, or the first of four cell counts. |
| ... | Further cell counts if x is a bare count. |
| which | "positive" (the default) or "negative". |
| prevalence | Optional. Apply the test to a population with this prevalence rather than the one observed in the table. No confidence interval is produced in that case, because the result is not a proportion estimated from these data. |
| conf_level | Confidence level for the interval. |

## Details

The table is expected in screening orientation, which is the same shape as every other table in this package once you read "exposed" as "test positive" and "case" as "diseased":

| | Diseased | Healthy |
|---|---|---|
| **Test +** | a (TP) | b (FP) |
| **Test -** | c (FN) | d (TN) |

Supplying prevalence applies the test's sensitivity and specificity to a population with that prevalence instead of the one in the table, via Bayes' theorem. This is how you answer "what would this test do in a screening programme?" using validation data from a high-prevalence clinical sample.

## Value

An epibyhand_derivation.

## References

Altman, D. G. and Bland, J. M. (1994). Diagnostic tests 2: predictive values. *BMJ*, 309, 102.

## Examples

```
# A test with 90% sensitivity and 90% specificity, 5% prevalence
scr <- epi2x2(90, 180, 10, 1720,
              exposure = c("Test positive", "Test negative"),
              outcome  = c("Diseased", "Healthy"))
predictive_value(scr)
predictive_value(scr, which = "negative")

# The same test in a population where 40% are diseased
predictive_value(scr, prevalence = 0.40)
```

## print.epi2x2 *Print a 2 x 2 table with margins*

### Description

Print a 2 x 2 table with margins

### Usage

```
## S3 method for class 'epi2x2'
print(x, indent = "", ...)
```

### Arguments

| | |
|---|---|
| x | An epi2x2. |
| indent | Character prefix for each line. |
| ... | Unused. |

### Value

x, invisibly.

## print.epibyhand_derivation *Print a worked derivation*

### Description

Print a worked derivation

### Usage

```
## S3 method for class 'epibyhand_derivation'
print(x, verbose = NULL, digits = NULL, ...)
```

### Arguments

| | |
|---|---|
| x | An epibyhand_derivation. |
| verbose | Detail level. 0 prints the result only, 1 adds the symbolic formulas, 2 (the default) shows the full worked solution with numbers substituted in. Set globally with options(epibyhand.verbose = 1). |
| digits | Decimal places. Set globally with options(epibyhand.digits = 3). |
| ... | Unused. |

### Value

x, invisibly.

## print.epi_strata *Print stratified tables*

### Description

Print stratified tables

### Usage

```
## S3 method for class 'epi_strata'
print(x, indent = "", ...)
```

### Arguments

| | |
|---|---|
| x | An epi_strata. |
| indent | Character prefix for each line. |
| ... | Unused. |

### Value

x, invisibly.

## risk_difference *Risk difference with worked derivation*

### Description

Risk difference with worked derivation

### Usage

```
risk_difference(x, ..., conf_level = 0.95)
```

### Arguments

| | |
|---|---|
| x | An epi2x2, or the first of four cell counts. |
| ... | Further cell counts if x is a bare count. |
| conf_level | Confidence level for the interval. |

### Value

An epibyhand_derivation.

### Examples

```
risk_difference(epi2x2(36, 14, 30, 25))
```

## risk_ratio *Risk ratio with worked derivation*

### Description

Risk ratio with worked derivation

### Usage

```
risk_ratio(x, ..., conf_level = 0.95)
```

### Arguments

| | |
|---|---|
| x | An epi2x2, or the first of four cell counts. |
| ... | Further cell counts if x is a bare count. |
| conf_level | Confidence level for the interval. |

### Value

An epibyhand_derivation.

### Examples

```
risk_ratio(epi2x2(36, 14, 30, 25))
```

## steps_table *Return the steps of a derivation as a data frame*

### Description

Useful for building answer keys, or for rendering the derivation in a format this package does not provide.

### Usage

```
steps_table(x)
```

### Arguments

| | |
|---|---|
| x | An epibyhand_derivation. |

### Value

A data frame with one row per step.

`to_latex` *Render a derivation as LaTeX*

## Description

Produces a LaTeX fragment of the worked derivation, for dropping into a problem set, an exam, or lecture notes written in LaTeX. Each step is an aligned display with the formula in symbols, the same formula with the observed numbers, and the result. Tables use `tabular`.

## Usage

```
to_latex(x, verbose = NULL, digits = NULL, standalone = FALSE)
```

## Arguments

| | |
|---|---|
| `x` | An `epibyhand_derivation`. |
| `verbose` | Detail level, as for `print.epibyhand_derivation()`. |
| `digits` | Decimal places, as for `print.epibyhand_derivation()`. |
| `standalone` | If `TRUE`, wrap the fragment in a minimal complete document. |

## Details

The fragment needs `\usepackage{amsmath}` in the preamble. Nothing else.

Set `standalone = TRUE` for a complete document that compiles on its own, which is useful for checking the output or for a one-page answer sheet.

## Value

A single character string of LaTeX.

## See Also

`to_markdown()`.

## Examples

```
d <- odds_ratio(epi2x2(36, 14, 30, 25))
cat(to_latex(d))

# Write a compilable answer sheet
f <- tempfile(fileext = ".tex")
writeLines(to_latex(d, standalone = TRUE), f)
```

`to_markdown` *Render a derivation as Markdown with embedded LaTeX mathematics*

## Description

Produces the same worked derivation the console shows, but as Markdown with the formulas typeset as mathematics: fractions, square roots and subscripts rather than slashes and underscores. Tables become Markdown tables. The output renders in Quarto and R Markdown to HTML (through MathJax or KaTeX), to PDF (through LaTeX), and to Word.

## Usage

```
to_markdown(x, verbose = NULL, digits = NULL)
```

## Arguments

| | |
|---|---|
| `x` | An `epibyhand_derivation`. |
| `verbose` | Detail level, as for `print.epibyhand_derivation()`. |
| `digits` | Decimal places, as for `print.epibyhand_derivation()`. |

## Details

You rarely need to call this directly. When a derivation is the result of a code chunk in a Quarto or R Markdown document, it is rendered this way automatically. Set `options(epibyhand.knit = "console")` to get the plain console output in documents instead.

## Value

A single character string of Markdown.

## See Also

`to_latex()` for standalone LaTeX.

## Examples

```
d <- odds_ratio(epi2x2(36, 14, 30, 25))
cat(to_markdown(d))
cat(to_markdown(d, verbose = 0))
```

# Index

attributable_fraction, 2

check_work, 3
collapse_strata, 4
confint.epibyhand_derivation, 5

derivation, 5
derivation_step, 6

epi2x2, 7
epi2x2(), *8*
epi_strata, 8
epi_strata(), *9–11*
estimate, 9

homogeneity, 9

mh_odds_ratio, 10
mh_odds_ratio(), *9*
mh_risk_ratio, 11

odds_ratio, 12

predictive_value, 12
print.epi2x2, 14
print.epi_strata, 15
print.epibyhand_derivation, 14
print.epibyhand_derivation(), *17, 18*

risk_difference, 15
risk_ratio, 16

steps_table, 16

to_latex, 17
to_latex(), *7*, *18*
to_markdown, 18
to_markdown(), *7*, *17*