Package: twoPhaseSens (via r-universe)

October 22, 2026

Title Sensitivity Analysis for Covariates Observed in Two-Phase Samples

Version 0.1.0

Description Provides a structured method-comparison sensitivity-analysis workflow for linear regression when important continuous adjustment covariates are observed only in a Phase-2 subsample. Implements naive analysis, complete-case analysis, fully conditional specification multiple imputation, joint-model multiple imputation, inverse probability weighting, and augmented inverse probability weighting, with standardized tables, diagnostics, and publication-style forest plots. Methodological background is provided by Little and Rubin (2019) <doi:10.1002/9781119482260> for general missing-data methods, White and Carlin (2010) <doi:10.1002/sim.3944> for complete-case analysis with missing covariates, White, Royston and Wood (2011) <doi:10.1002/sim.4067> and van Buuren and Groothuis-Oudshoorn (2011) <doi:10.18637/jss.v045.i03> for chained-equations multiple imputation, Quartagno, Grund and Carpenter (2019) <doi:10.32614/RJ-2019-028> for joint-model multiple imputation, Robins, Rotnitzky and Zhao (1994) <doi:10.1080/01621459.1994.10476818> for inverse-probability weighting with missing regressors, and Bang and Robins (2005) <doi:10.1111/j.1541-0420.2005.00377.x> for doubly robust estimation.

License MIT + file LICENSE

URL https://rophenceojiambo.github.io/twoPhaseSens/, https://github.com/rophenceojiambo/twoPhaseSens

BugReports https://github.com/rophenceojiambo/twoPhaseSens/issues

Encoding UTF-8

RoxygenNote 7.3.3

Imports ggplot2, gtsummary (>= 2.0.0), jomo, mice, rlang, tidyselect

Suggests knitr, rmarkdown, testthat (>= 3.0.0)

VignetteBuilder knitr

Config/testthat/edition 3

NeedsCompilation no

Author Rophence Ojiambo [aut, cre] (ORCID: <https://orcid.org/0000-0002-6798-9066>), Jemar Bather [aut] (ORCID: <https://orcid.org/0000-0002-0285-3678>)

Maintainer Rophence Ojiambo <rophence.ojiambo@nyu.edu>

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

Date/Publication 2026-10-22 19:24:17 UTC

RemoteUrl https://github.com/cran/twoPhaseSens

RemoteRef HEAD

RemoteSha ac3f393049a5e871e58b7ccab59657c7c54a34f9

Contents

diagnostics 2
plot_sensitivity 3
tbl_sensitivity 5
twophase_example_data 6
twophase_sensitivity 7

Index **9**

`diagnostics` *Extract diagnostics from a twoPhaseSens analysis*

Description

Returns method-level information from a `twophase_sensitivity()` fit. The default view contains analysis status, sample-size information, stochastic settings, multiple-imputation settings, and available IPW/AIPW weight diagnostics. More focused views can be requested with `type`.

Usage

```
diagnostics(x, ...)

## S3 method for class 'twophase_sensitivity'
diagnostics(
  x,
  type = c("all", "status", "weights", "mi"),
  outcome = NULL,
  method = NULL,
  ...
)
```

Arguments

`x`	An object.
`...`	Additional arguments passed to methods.
`type`	Diagnostic view. One of `"all"`, `"status"`, `"weights"`, or `"mi"`.
`outcome`	Optional character vector selecting one or more outcomes.
`method`	Optional character vector selecting one or more methods. Method names are normalized in the same way as in `twophase_sensitivity()`.

Details

The function reports diagnostics but does not automatically label values as acceptable or unacceptable. Interpretation of weight variability, effective sample size, imputation behavior, and method failures should be based on the study design and scientific context.

Value

A data frame with one row per selected outcome-method combination.

Examples

```
dat <- twophase_example_data(n = 200, seed = 2026)
fit <- twophase_sensitivity(
  data = dat,
  outcome = "outcome1",
  exposure = "exposure",
  covariates = c("age", "sex"),
  phase2_covariates = c("marker1", "marker2", "marker3"),
  phase2 = "phase2",
  methods = c("naive", "cca", "ipw"),
  seed = 1001
)
diagnostics(fit)
```

`plot_sensitivity` *Plot sensitivity-analysis estimates and confidence intervals*

Description

Creates a publication-style forest plot from a `twophase_sensitivity()` object. Each method is shown with a method-specific color and symbol, horizontal confidence interval, and optional right-side text columns for the formatted coefficient/CI and p value. With multiple outcomes, outcomes are shown in separate facets.

Usage

```
plot_sensitivity(
  x,
  outcome = NULL,
  method = NULL,
  outcome_labels = NULL,
  method_labels = NULL,
  reference_line = 0,
  facet_scales = c("free_x", "fixed"),
  x_label = NULL,
  estimate_label = "Coefficient",
  title = NULL,
  annotate = TRUE,
  estimate_digits = 3L,
  p_digits = 3L
)

## S3 method for class 'twophase_sensitivity'
plot(x, ...)
```

Arguments

`x`	A `twophase_sensitivity` object.
`outcome`	Optional character vector selecting one or more outcomes.
`method`	Optional character vector selecting one or more methods.
`outcome_labels`	Optional character vector used to relabel outcomes. Supply either a named vector whose names are the outcome variable names or an unnamed vector in the same order as the selected outcomes.
`method_labels`	Optional named character vector used to relabel methods.
`reference_line`	Numeric scalar giving the vertical reference line. Use `NULL` to omit the line. The default is 0.
`facet_scales`	Facet x-axis scaling for multiple outcomes: `"free_x"` (default) or `"fixed"`.
`x_label`	Optional x-axis label. If `NULL`, a label is generated from `estimate_label` and the confidence level stored in the analysis object.
`estimate_label`	Label used for the numeric estimate column. The default is `"Coefficient"`. For a manuscript that reports adjusted betas, for example, use `estimate_label = "Adjusted beta"`.
`title`	Optional overall plot title. Outcome labels are shown in facet strips, including when a single outcome is plotted.
`annotate`	Logical; if `TRUE` (default), add manuscript-style right-side columns for the coefficient with confidence interval and p value.
`estimate_digits`	Number of digits used in the coefficient and confidence-interval annotation. The default is 3.
`p_digits`	Number of digits used in the p-value annotation.
`...`	Arguments passed to `plot_sensitivity()`.

Value

A ggplot object.

Examples

```
dat <- twophase_example_data(n = 200, seed = 2026)
fit <- twophase_sensitivity(
  data = dat,
  outcome = "outcome1",
  exposure = "exposure",
  covariates = c("age", "sex"),
  phase2_covariates = c("marker1", "marker2", "marker3"),
  phase2 = "phase2",
  methods = c("naive", "cca", "ipw"),
  seed = 1001
)
plot_sensitivity(fit, outcome_labels = c(outcome1 = "Outcome 1"))
```

---

tbl_sensitivity *Create a publication-ready sensitivity-analysis table*

Description

Formats method-comparison results from `twophase_sensitivity()` with methods in rows and one group of columns per outcome. By default, the function returns a `gtsummary` object with outcome spanning headers. A plain data frame can be requested for downstream custom formatting.

Usage

```
tbl_sensitivity(
  x,
  estimate_digits = 3L,
  p_digits = 3L,
  include_se = FALSE,
  method_labels = NULL,
  outcome_labels = NULL,
  caption = NULL,
  output = c("gtsummary", "data.frame")
)
```

Arguments

`x`	A `twophase_sensitivity` object.
`estimate_digits`	Number of digits for estimates and confidence limits. The default is 3.
`p_digits`	Number of digits for p values.
`include_se`	Logical; include a separate standard-error column for each outcome.

method_labels	Optional named character vector used to relabel methods. Names should be package method labels such as "FCS-MI" or "AIPW".
outcome_labels	Optional character vector used to relabel outcomes. Supply either a named vector whose names are the outcome variable names or an unnamed vector in the same order as x$specification$outcome.
caption	Optional table caption. Markdown is supported for output = "gtsummary".
output	Output type: "gtsummary" (default) or "data.frame".

Value

A gtsummary object by default, or a formatted data frame when output = "data.frame".

Examples

```
dat <- twophase_example_data(n = 200, seed = 2026)
fit <- twophase_sensitivity(
  data = dat,
  outcome = "outcome1",
  exposure = "exposure",
  covariates = c("age", "sex"),
  phase2_covariates = c("marker1", "marker2", "marker3"),
  phase2 = "phase2",
  methods = c("naive", "cca", "ipw"),
  seed = 1001
)
tbl_sensitivity(fit, output = "data.frame")
```

---

twophase_example_data *Generate a synthetic two-phase example dataset*

Description

Creates a small reproducible dataset for examples, teaching, and package documentation. Phase-1 variables are fully observed, while the continuous Phase-2 covariates are observed as a complete block only when phase2 = 1.

Usage

```
twophase_example_data(n = 300L, seed = 2026L)
```

Arguments

n	Number of observations.
seed	Non-negative integer seed used only while generating the example data. The caller's pre-existing global random-number state is restored.

Details

The generated data are synthetic and are not derived from MIDUS or any other participant-level study data.

Value

A data frame containing two continuous outcomes, one continuous exposure, two Phase-1 covariates, three continuous Phase-2 covariates, and a binary Phase-2 indicator.

Examples

```
dat <- twophase_example_data(n = 200, seed = 2026)
head(dat)
table(dat$phase2)
```

---

twophase_sensitivity *Compare methods for Phase-2 covariates missing by design*

Description

Runs a prespecified set of analytic approaches for a linear-regression exposure coefficient when a block of continuous adjustment covariates is observed only in a Phase-2 subsample.

Usage

```
twophase_sensitivity(
  data,
  outcome,
  exposure,
  covariates = character(),
  phase2_covariates,
  phase2,
  methods = c("naive", "cca", "fcs_mi", "jm_mi", "ipw", "aipw"),
  n_imputations = .twophasesens_defaults$n_imputations,
  mice_maxit = .twophasesens_defaults$mice_maxit,
  jomo_nburn = .twophasesens_defaults$jomo_nburn,
  jomo_nbetween = .twophasesens_defaults$jomo_nbetween,
  seed = NULL,
  conf_level = .twophasesens_defaults$conf_level
)
```

Arguments

data	A data frame.
outcome	Character vector naming one or more continuous outcomes. A single outcome is fully supported. Multiple outcomes are analyzed separately using the same exposure, covariates, Phase-2 variables, and requested methods.

`exposure`	Character scalar naming the numeric primary exposure.
`covariates`	Character vector naming fully observed Phase-1 covariates.
`phase2_covariates`	Character vector naming continuous Phase-2 covariates.
`phase2`	Character scalar naming the 0/1 Phase-2 indicator.
`methods`	Methods to run. Defaults to all six validated approaches.
`n_imputations`	Number of imputations for FCS-MI and JM-MI.
`mice_maxit`	Maximum FCS iterations.
`jomo_nburn`	JM-MI burn-in iterations.
`jomo_nbetween`	JM-MI iterations between saved imputations.
`seed`	Optional random-number seed. Supply either one non-negative integer, which is expanded sequentially across outcomes, or one integer per outcome. A named vector may be supplied using the outcome names. The previous global RNG state is restored after each outcome analysis.
`conf_level`	Confidence level for intervals.

Value

An object of class `twophase_sensitivity`.

Examples

```
dat <- twophase_example_data(n = 200, seed = 2026)
fit <- twophase_sensitivity(
  data = dat,
  outcome = "outcome1",
  exposure = "exposure",
  covariates = c("age", "sex"),
  phase2_covariates = c("marker1", "marker2", "marker3"),
  phase2 = "phase2",
  methods = c("naive", "cca", "ipw"),
  seed = 1001
)
fit
```

Index

diagnostics, 2

plot.twophase_sensitivity
(plot_sensitivity), 3
plot_sensitivity, 3
plot_sensitivity(), *4*

tbl_sensitivity, 5
twophase_example_data, 6
twophase_sensitivity, 7
twophase_sensitivity(), *2, 3*, *5*