Package 'AIBias'

April 4, 2026

Title Longitudinal Bias Auditing for Sequential Decision Systems

Version 0.1.0

Description Provides tools for detecting, quantifying, and visualizing algorithmic bias as a longitudinal process in repeated decision systems. Existing fairness metrics treat bias as a single-period snapshot; this package operationalizes the view that bias in sequential systems must be measured over time. Implements group-specific decision-rate trajectories, standardized disparity measures analogous to the standardized mean difference (Cohen, 1988, ISBN:0-8058-0283-5), cumulative bias burden, Markov-based transition disparity (recovery and retention gaps), and a dynamic amplification index that quantifies whether prior decisions compound current group inequality. The amplification framework extends longitudinal causal inference ideas from Robins (1986) <doi:10.1016/0270-0255(86)90088-6> and the sequential decision-process perspective in the fairness literature (see <https://fairmlbook.org>) to the audit setting. Covariate-adjusted trajectories are estimated via logistic regression, generalized additive models (Wood, 2017, <doi:10.1201/9781315370279>), or generalized linear mixed models (Bates, 2015, <doi:10.18637/jss.v067.i01>). Uncertainty quantification uses the cluster bootstrap (Cameron, 2008, <doi:10.1162/rest.90.3.414>).

License MIT + file LICENSE

Encoding UTF-8

Language en-US

RoxygenNote 7.3.3

Imports dplyr (>= 1.1.0), tidyr (>= 1.3.0), ggplot2 (>= 3.4.0), rlang (>= 1.1.0), cli (>= 3.6.0), purrr (>= 1.0.0), tibble (>= 3.2.0)

Suggests mgcv, lme4, boot, knitr, rmarkdown, testthat (>= 3.0.0)

Config/testthat/edition 3

VignetteBuilder knitr

LazyData true

Depends R (>= 4.1.0)

URL https://github.com/causalfragility-lab/AIBias

BugReports https://github.com/causalfragility-lab/AIBias/issues

NeedsCompilation no

Author Subir Hait [aut, cre] (ORCID: <https://orcid.org/0009-0004-9871-9677>)

Maintainer Subir Hait <haitsubi@msu.edu>

Repository CRAN

Date/Publication 2026-04-04 09:30:08 UTC

Contents

`aib_adjust` *Covariate-adjusted bias trajectories*

Description

Estimates covariate-adjusted bias trajectories by fitting a model for $Pr(D_{it} = 1 \mid X_{it}, A_i, t)$ and computing marginal predicted disparities by group and time.

Usage

```
aib_adjust(
  object,
  formula,
  method = c("glm", "gam", "mixed"),
  ref_group = NULL,
  verbose = TRUE
)
```

Arguments

object	An aibias object from aib_build().
formula	A one-sided formula specifying covariates, e.g. ~ age + income + prior_score. The group and time variables are added automatically.
method	One of "glm", "gam", or "mixed". • "glm": logistic regression • "gam": generalized additive model (requires mgcv) • "mixed": GLMM with random intercept for id (requires lme4)
ref_group	Character. Reference group.
verbose	Logical.

Value

The aibias object with $adjusted populated, containing:

- trajectory: adjusted bias trajectory
- marginal_rates: marginal predicted rates by group and time
- model: the fitted model object
- formula_used: the full formula passed to the model

Examples

```
data(lending_panel)
obj <- aib_build(lending_panel, "applicant_id", "year", "race", "approved")
obj <- aib_adjust(obj, formula = ~ income + credit_score, method = "glm",
                  ref_group = "White")
```

---

aib_amplify *Compute dynamic bias amplification indices*

Description

Estimates the amplification index $A_{g,r}(t) = B_{g,r}(t|1) - B_{g,r}(t|0)$, which measures how conditioning on prior decision state changes the group disparity at time t. Non-zero amplification indicates that prior decisions are shaping current disparities—the hallmark of compounding bias.

Usage

```
aib_amplify(object, ref_group = NULL, verbose = TRUE)
```

Arguments

object	An aibias object from aib_build(). Recommended to run aib_transition() first.
ref_group	Character. Reference group.
verbose	Logical.

Details

A decision system exhibits **bias amplification** if:

1. $|B_{g,r}(t)| > |B_{g,r}(s)|$ for some $t > s$ (disparity grows), AND
2. $A_{g,r}(t) \neq 0$ (prior decisions drive current disparity), OR
3. $P_g(t) \neq P_r(t)$ (transition matrices are unequal)

Value

The `aibias` object with `$amplification` populated. Contains:

- `lagged_disparity`: $B_{g,r}(t|d)$ for d equal to 0 or 1
- `index`: Amplification index $A_{g,r}(t)$
- `cumulative`: $A_{g,r}(T)$ summed over time
- `matrix_norm`: $\sum_t \|P_g(t) - P_r(t)\|$

Examples

```
data(lending_panel)
obj <- aib_build(lending_panel, "applicant_id", "year", "race", "approved")
obj <- aib_transition(obj, ref_group = "White")
obj <- aib_amplify(obj, ref_group = "White")
obj$amplification$index
```

`aib_audit` *Run the full AIBias audit pipeline*

Description

A convenience wrapper that runs the complete audit pipeline: `aib_describe()` → `aib_transition()` → `aib_amplify()`. Optionally runs `aib_bootstrap()` for uncertainty quantification.

Usage

```
aib_audit(
  object,
  ref_group = NULL,
  bootstrap = FALSE,
  B = 500,
  seed = NULL,
  verbose = TRUE,
  ...
)
```

Arguments

object	An aibias object from aib_build(), OR a data frame (in which case ... arguments are passed to aib_build()).
ref_group	Character. Reference group.
bootstrap	Logical. Run bootstrap CIs? Default FALSE.
B	Integer. Bootstrap replicates if bootstrap = TRUE.
seed	Integer. Random seed.
verbose	Logical.
...	If object is a data frame, additional arguments passed to aib_build().

Value

A fully-populated aibias object.

Examples

```
data(lending_panel)
result <- aib_audit(lending_panel,
                    id        = "applicant_id",
                    time      = "year",
                    group     = "race",
                    decision  = "approved",
                    ref_group = "White")

summary(result)
plot(result, type = "trajectory")
```

---

aib_bootstrap *Bootstrap confidence intervals for bias trajectories*

Description

Computes bootstrap confidence intervals for the bias trajectory and cumulative burden by resampling units (cluster bootstrap).

Usage

```
aib_bootstrap(
  object,
  B = 500,
  conf = 0.95,
  ref_group = NULL,
  seed = NULL,
  verbose = TRUE
)
```

Arguments

object	An aibias object with $bias populated.
B	Integer. Number of bootstrap replicates. Default 500.
conf	Numeric. Confidence level. Default 0.95.
ref_group	Character. Reference group.
seed	Integer. Random seed for reproducibility.
verbose	Logical.

Details

Uses the cluster (unit-level) bootstrap to preserve the panel structure. Units are resampled with replacement; all their time observations are retained.

Value

The aibias object with $bootstrap populated.

Examples

```
data(lending_panel)
obj <- aib_build(lending_panel, "applicant_id", "year", "race", "approved")
obj <- aib_describe(obj, ref_group = "White")
obj <- aib_bootstrap(obj, B = 200, seed = 42)
```

---

aib_build *Build an AIBias audit object*

Description

Constructs the core aibias S3 object from a panel dataset. Validates the panel structure and prepares internal data for downstream analysis.

Usage

```
aib_build(data, id, time, group, decision, outcome = NULL, verbose = TRUE)
```

Arguments

data	A data frame in long (panel) format.
id	Character. Name of the unit identifier column.
time	Character. Name of the time/wave column (integer or factor).
group	Character. Name of the protected group column.
decision	Character. Name of the binary decision column (0/1).
outcome	Character or NULL. Optional downstream outcome column.
verbose	Logical. Print validation messages. Default TRUE.

Details

The function expects a balanced or unbalanced panel where:

- `id` indexes units observed over multiple periods
- `time` is an ordered index (will be coerced to integer rank)
- `group` is a categorical variable indicating protected group membership
- `decision` is a binary 0/1 variable (1 = favorable decision)

Value

An object of class `"aibias"`.

Examples

```
data(lending_panel)
obj <- aib_build(lending_panel,
                 id       = "applicant_id",
                 time     = "year",
                 group    = "race",
                 decision = "approved")
```

---

aib_describe *Describe longitudinal disparity trajectories*

Description

Computes group decision rate trajectories, raw and standardized bias trajectories, and cumulative bias burden.

Usage

```
aib_describe(object, ref_group = NULL, weights = NULL, verbose = TRUE)
```

Arguments

`object`	An `aibias` object from `aib_build()`.
`ref_group`	Character. Reference group label. If NULL, uses the first group level.
`weights`	Numeric vector of time weights for cumulative burden. Length must equal the number of time points. If NULL, equal weights.
`verbose`	Logical. Print summary output.

Value

The `aibias` object with `$bias` populated. The bias element is a list with components:

- `trajectory`: raw bias trajectory $B_{g,r}(t)$
- `trajectory_smd`: standardized bias trajectory $B^*_{g,r}(t)$
- `cumulative`: cumulative bias burden $CB_{g,r}(T)$
- `slope`: first differences $\Delta B_{g,r}(t)$
- `curvature`: second differences $\Delta^2 B_{g,r}(t)$
- `ref_group`: reference group used

Examples

```
data(lending_panel)
obj <- aib_build(lending_panel, "applicant_id", "year", "race", "approved")
obj <- aib_describe(obj, ref_group = "White")
obj$bias$cumulative
```

---

aib_figures *Copy and run the paper figures script*

Description

Copies the paper figures script to your working directory and optionally runs it. The script produces four publication-ready figures illustrating bias trajectory, transition asymmetry, amplification index, and cumulative burden from a toy simulation (N=20, T=3).

Usage

```
aib_figures(run = TRUE, dest = file.path(tempdir(), "paper_figures.R"))
```

Arguments

`run`	Logical. If TRUE (default), runs the script immediately. If FALSE, just copies the file for you to inspect and edit first.
`dest`	Character. Destination filename. Default `file.path(tempdir(), "paper_figures.R")`.

Value

The path to the copied script, invisibly.

Examples

```
# Copy and run immediately
aib_figures()

# Just copy to inspect first
aib_figures(run = FALSE, dest = file.path(tempdir(), "paper_figures.R"))
```

`aib_panel_info` *Check panel balance*

Description

Check panel balance

Usage

```
aib_panel_info(object)
```

Arguments

`object`	An `aibias` object.

Value

A tibble summarizing observation counts per unit.

`aib_persistence` *Compute bias persistence above a threshold*

Description

Compute bias persistence above a threshold

Usage

```
aib_persistence(object, threshold = 0.05)
```

Arguments

`object`	An `aibias` object with `$bias` populated.
`threshold`	Numeric. Minimum absolute disparity to count. Default 0.05.

Value

A tibble with group-level persistence counts.

aib_transition *Compute group-specific decision transition matrices and disparities*

Description

Estimates group-specific Markov transition probabilities $p_g^{ab}(t) = Pr(D_{it} = b \mid D_{i,t-1} = a, A_i = g)$ and derives transition disparities, advantage retention, and recovery gaps.

Usage

```
aib_transition(object, ref_group = NULL, verbose = TRUE)
```

Arguments

object	An aibias object from aib_build().
ref_group	Character. Reference group. If NULL, uses first level.
verbose	Logical. Print summary.

Value

The aibias object with $transitions populated. Contains:

- probs: Transition probabilities by group and time
- pooled: Pooled transition probabilities (time-averaged)
- disparity: Transition disparity $\Delta_{g,r}^{ab}(t)$
- recovery_gap: Disparity in 0->1 transitions (recovery)
- retention_gap: Disparity in 1->1 transitions (retention)
- matrices: Named list of 2x2 transition matrices per group

Examples

```
data(lending_panel)
obj <- aib_build(lending_panel, "applicant_id", "year", "race", "approved")
obj <- aib_transition(obj, ref_group = "White")
obj$transitions$pooled
```

lending_panel *Synthetic Lending Panel Dataset*

Description

A synthetic panel dataset simulating loan application decisions over six years for applicants from three racial groups. Designed to illustrate longitudinal bias analysis with AIBias.

The data are generated so that Black and Hispanic applicants face lower approval rates, lower recovery probabilities after denial, and lower retention probabilities after approval — producing compounding disparities over time.

Usage

```
lending_panel
```

Format

A data frame with 3,600 rows and 6 columns:

applicant_id Character. Unique applicant identifier.

year Integer. Year of application (2015–2020).

race Factor. Racial group: White, Black, Hispanic.

income Numeric. Annual income (thousands USD).

credit_score Numeric. Credit score (300–850).

approved Integer. Loan approval decision (1 = approved, 0 = denied).

Details

Transition parameters used in data generation:

| Group | P(approve | prev approved) | P(approve | prev denied) | |———-|———————|———————| | White | 0.82 | 0.65 | | Black | 0.62 | 0.38 | | Hispanic | 0.68 | 0.44 |

Source

Synthetic data generated via `data-raw/lending_panel.R`.

Examples

```
data(lending_panel)
head(lending_panel)
table(lending_panel$race, lending_panel$year)
```

`plot.aibias` *Plot an aibias object*

Description

Visualizes audit results. Supports four plot types:

- `"trajectory"`: Bias trajectory $B_{g,r}(t)$ over time
- `"heatmap"`: Group-time disparity surface
- `"transition"`: Group-specific transition probabilities
- `"amplification"`: Amplification index $A_{g,r}(t)$ over time

Usage

```
## S3 method for class 'aibias'
plot(
  x,
  type = c("trajectory", "heatmap", "transition", "amplification"),
  show_ci = TRUE,
  color_palette = NULL,
  ...
)
```

Arguments

`x`	An `aibias` object.
`type`	Character. Plot type. One of `"trajectory"`, `"heatmap"`, `"transition"`, `"amplification"`.
`show_ci`	Logical. Show bootstrap CIs if available. Default TRUE.
`color_palette`	Character vector of colors for groups. If NULL, uses a sensible default.
`...`	Ignored.

Value

A `ggplot2` object.

Examples

```
data(lending_panel)
obj <- aib_audit(lending_panel,
                 id = "applicant_id", time = "year",
                 group = "race",     decision = "approved",
                 ref_group = "White", verbose = FALSE)
plot(obj, type = "trajectory")
plot(obj, type = "heatmap")
```

print.aibias *Print an aibias object*

Description

Print an aibias object

Usage

```
## S3 method for class 'aibias'
print(x, ...)
```

Arguments

x	An aibias object.
...	Ignored.

Value

Invisibly returns x, called for its side effect of printing a concise summary of the audit object to the console.

summary.aibias *Summarize an aibias object*

Description

Produces a comprehensive audit summary including trajectory statistics, transition gaps, amplification indices, and narrative interpretation.

Usage

```
## S3 method for class 'aibias'
summary(object, ...)
```

Arguments

object	An aibias object.
...	Ignored.

Value

Invisibly returns object, called for its side effect of printing a comprehensive audit summary to the console.

Index