Package: BSET (via r-universe)

October 31, 2026

Type Package

Title A Bayesian Surrogate Evaluation Test

Version 1.0

Date 2026-05-15

Description An implementation of the Bayesian Surrogate Evaluation Test (BSET) for assessing the validity of surrogate markers in clinical trials. Provides hypothesis testing tools to evaluate whether a surrogate can reliably estimate the causal effect of a treatment on a primary outcome. Implements the imputation-based Bayesian methodology of Carlotti and Parast (2026) <doi:10.48550/arXiv.2603.14381>, extending the frequentist rank-based approach of Parast et al. (2024) <doi:10.1093/biomtc/ujad035>. Addresses key limitations of the frequentist method, including the lack of causal interpretability and the inability to adjust for covariates in the estimation process.

License GPL-3

URL https://pietrocarlotti.github.io/BSET/, https://github.com/PietroCarlotti/BSET

BugReports https://github.com/PietroCarlotti/BSET/issues

Encoding UTF-8

LazyData true

RoxygenNote 7.3.3

Imports rstan, ggplot2, future, parallel, mvtnorm, dplyr, SurrogateRank, Rdpack, rlang

RdMacros Rdpack

Suggests knitr, rmarkdown, bookdown,

VignetteBuilder knitr, bookdown

Depends R (>= 3.5)

Config/pak/sysreqs cmake make libicu-dev libuv1-dev

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

Date/Publication 2026-10-01 20:03:43 UTC

RemoteUrl https://github.com/pietrocarlotti/bset

RemoteRef HEAD

RemoteSha 4d026834db1e5af52d125ca5bc0d719315ad06d3

Contents

Bayesian_test . . . 2
BSET . . . 3
Carlotti_and_Parast_2026_simulations_grid . . . 7
compute_BF_distribution . . . 9
compute_delta . . . 10
compute_estimands_Carlotti_and_Parast_2026 . . . 11
compute_estimands_Parast_et_al_2024 . . . 12
compute_theta . . . 14
compute_V_S_star . . . 15
DGP_no_X . . . 16
DGP_X_binary . . . 18
DGP_X_Gaussian . . . 19
estimands_Carlotti_and_Parast_2026 . . . 21
estimands_Parast_et_al_2024 . . . 21
frequentist_test . . . 22
Parast_et_al_2024_simulations_grid . . . 23

Index **24**

`Bayesian_test` *Bayesian Test from Carlotti and Parast (2026)*

Description

This function performs a Bayesian test for surrogate evaluation based on the imputation-based methodology of Carlotti and Parast (2026). It calculates credible intervals, the η threshold used, and determines if the surrogate is valid. This function is generally not intended to be called directly by the user and is instead used internally within `BSET_no_X` and `BSET_X`.

Usage

```
Bayesian_test(
  P_MCMC,
  alpha = 0.05,
  V_S_star,
  theta_true = NULL,
  V_Y_true = NULL
)
```

Arguments

`P_MCMC`	A three-dimensional array of potential outcomes obtained via MCMC sampling with dimensions `[n_subjects, variables, n_samples]`. The variables should correspond to (Y_1, S_1, Y_0, S_0).
`alpha`	Numeric. Significance level for the credible interval (default is 0.05).
`V_S_star`	Numeric. The value of v_S that satisfies the power constraint for the surrogate validation test.
`theta_true`	Numeric (optional). The true value of η, used to calculate frequentist coverage during simulations.
`V_Y_true`	Numeric (optional). The true value of V_Y, used to calculate the surrogate validation threshold η during simulations. If not provided, it will be estimated from the MCMC samples.

Value

A list containing:

- `V_Y_MCMC`: Posterior draws for V_Y.
- `V_S_MCMC`: Posterior draws for V_S.
- `theta_MCMC`: Posterior draws for $\theta = V_Y - V_S$.
- `CI`: The calculated credible interval for θ.
- `threshold`: The η threshold value used in the test.
- `coverage`: Logical indicating if `theta_true` falls within the `CI` (if `theta_true` is provided).
- `power`: Logical indicating if the upper bound of `CI` is below `threshold`, which indicates that the test identifies the surrogate as valid.

References

Carlotti P, Parast L (2026). "A Bayesian Critique of Rank-Based Methods for Surrogate Marker Evaluation." *arXiv preprint arXiv:2603.14381*.

`BSET` *Bayesian Surrogate Evaluation Test from Carlotti and Parast (2026)*

Description

This function implements the Bayesian Surrogate Evaluation Test (BSET) as proposed by Carlotti and Parast (2026). When `X` is `NULL` (the default), the model is fit without covariates; when `X` is provided, a multivariate regression model is used to adjust for baseline covariates. In both cases the function fits a Bayesian model using Stan to generate posterior samples for the parameters of interest, which are then used to conduct a Bayesian hypothesis test for evaluating the validity of the surrogate marker. A frequentist test is also performed for comparison.

Usage

```
BSET(
  data,
  Y,
  S,
  Z,
  X = NULL,
  delta_true = NULL,
  theta_true = NULL,
  seed = NULL,
  n_chains = 4,
  n_iter = 2000,
  burn_in_ratio = 0.25,
  a = 1,
  b = 1,
  alpha = 0.05,
  beta = 0.2,
  V_S_zero = 0.5,
  BF_alternative = "greater",
  root_tolerance = 1e-16,
  mu_0 = rep(0, 4),
  Sigma_0 = diag(4),
  intercept = TRUE,
  mu_beta = NULL,
  Sigma_beta = NULL,
  s = rep(1, 4),
  tau = 1,
  plot = FALSE,
  mute = TRUE,
  parallel = TRUE
)
```

Arguments

`data`	A data frame containing the observed data.
`Y`	Character. Name of the outcome variable.
`S`	Character. Name of the surrogate variable.
`Z`	Character. Name of the treatment assignment variable.
`X`	Character vector. Names of the covariate columns to include in the model. When `NULL` (the default), the model is fit without covariates and the `mu_0` / `Sigma_0` prior parameters are used. When provided, the covariate-adjusted model is fit and the `intercept` / `mu_beta` / `Sigma_beta` prior parameters are used instead.
`delta_true`	The true value of delta, used to calculate frequentist coverage during simulations (optional).
`theta_true`	The true value of theta, used to calculate Bayesian coverage during simulations (optional).

seed	Random seed for reproducibility (optional. If not provided, a random seed will be generated).
n_chains	Number of MCMC chains to run (default is 4).
n_iter	Number of iterations per MCMC chain (default is 2000).
burn_in_ratio	Proportion of iterations to discard as burn-in (default is 0.25).
a, b	Prior parameters for prior Beta distribution on V_S (default is a = 1, b = 1).
alpha	Type I error rate for the test (default is 0.05).
beta	Type II error rate for the test (default is 0.2).
V_S_zero	Value of V_S under the null hypothesis (default is 0.5).
BF_alternative	Alternative hypothesis for the Bayes factor test ("greater", "less" or "two.sided").
root_tolerance	Numerical tolerance for root-finding algorithms (default is 1e-16).
mu_0	Prior mean vector for the mean parameters. Used only when X = NULL (default is a vector of zeros of length 4).
Sigma_0	Prior covariance matrix for the mean vector. Used only when X = NULL (default is a 4x4 identity matrix).
intercept	Logical. Whether to include an intercept in the regression model. Used only when X is provided (default is TRUE).
mu_beta	Prior mean vector for the regression coefficients. Used only when X is provided (default is a vector of zeros with length equal to the number of covariates, including the intercept if intercept = TRUE).
Sigma_beta	Prior covariance matrix for the regression coefficients. Used only when X is provided (default is a diagonal matrix with variance 10, with dimensions equal to the number of covariates, including the intercept if intercept = TRUE).
s	Prior scale parameters for the error variance (default is a vector of ones of length 4).
tau	Prior parameter for the LKJ correlation distribution (default is 1).
plot	Logical. Whether to plot the posterior distribution of theta (default is FALSE).
mute	Logical. Whether to suppress Stan output during model fitting (default is TRUE).
parallel	Logical. Whether to use parallel processing for MCMC sampling (default is TRUE).

Details

Without covariates (X = NULL), the Bayesian model is:

$$(Y_i, S_i) = \begin{cases} P_{1i}, & \text{if } Z_i = 1 \\ P_{0i}, & \text{if } Z_i = 0 \end{cases}, \quad i = 1, \ldots, n,$$

$$P_i \overset{\text{ind.}}{\sim} \mathcal{N}_4(\mu, \Sigma), \quad i = 1, \ldots, n,$$

$$\mu \sim \mathcal{N}_4(\mu_0, \Sigma_0),$$

$$\Sigma = \text{diag}(\sigma_{1:4})\, \Omega\, \text{diag}(\sigma_{1:4}),$$

$$\sigma_k \sim \text{Half-Normal}(0, s_k), \quad k = 1, \ldots, 4,$$
$$\Omega \sim \text{LKJ}(\tau).$$

With covariates (`X` provided), the Bayesian model is:

$$P_i \overset{\text{ind.}}{\sim} \mathcal{N}_4(BX_i, \Sigma), \quad i = 1, \ldots, n,$$
$$B = \begin{bmatrix} \beta_1 & \beta_2 & \beta_3 & \beta_4 \end{bmatrix}^\top,$$
$$\beta_k \sim \mathcal{N}_d(\mu_\beta, \Sigma_\beta), \quad k = 1, \ldots, 4,$$

with Σ, σ_k, and Ω as above.

This is a primary user-facing function of the package and includes working examples below.

Value

A list containing:

- `stan_cores`: The number of CPU cores used for MCMC sampling.
- `Bayesian_test`: A list containing the results of the Bayesian test, including posterior samples and credible intervals.
- `frequentist_test`: A list containing the results of the frequentist test, including point estimates and confidence intervals.
- `theta_posterior_plot`: A `ggplot` object showing the posterior distribution of θ, with vertical lines indicating the credible interval, the η threshold, and the true values of δ and θ (if provided).

References

Carlotti P, Parast L (2026). "A Bayesian Critique of Rank-Based Methods for Surrogate Marker Evaluation." *arXiv preprint arXiv:2603.14381*.

Parast L, Cai T, Tian L (2024). "A rank-based approach to evaluate a surrogate marker in a small sample setting." *Biometrics*, **80**(1), ujad035.

Examples

```
# Generate data from the perfect surrogate setting of Parast et al. (2024)
set.seed(123)
data_no_X <- DGP_no_X(
  n = 100,
  p = 0.5,
  mu_star = c(6, 6, 2.5, 2.5),
  Sigma_star = kronecker(diag(2), matrix(c(3, 3, 3, 3.1), 2, 2)),
  model = "Gaussian"
)

# Prepare the data frame
df_no_X <- data.frame(
  Y = data_no_X$P_observed[, "Y"],
  S = data_no_X$P_observed[, "S"],
```

```
  Z = data_no_X$Z
)

# Run BSET without covariates (requires Stan compilation, ~3 minutes)

result_no_X <- BSET(
  data = df_no_X,
  Y = "Y",
  S = "S",
  Z = "Z",
  seed = 123,
  n_chains = 2,
  n_iter = 500
)


# Generate data from the setting of Carlotti and Parast (2026) with a binary covariate
set.seed(123)
data_X <- DGP_X_binary(
  n = 100,
  p = 0.5,
  q = 0.5,
  mu_0 = c(5, 5, 0, 0),
  mu_1 = c(5, -5, 0, -10),
  Sigma_0 = kronecker(diag(2), matrix(c(1, 1, 1, 2), 2, 2)),
  Sigma_1 = kronecker(diag(2), matrix(c(1, 1, 1, 2), 2, 2))
)

# Prepare the data frame
df_X <- data.frame(
  Y = data_X$P_observed[, "Y"],
  S = data_X$P_observed[, "S"],
  Z = data_X$Z,
  X = data_X$X
)

# Run BSET with covariates (requires Stan compilation, ~3 minutes)

result_X <- BSET(
  data = df_X,
  Y = "Y",
  S = "S",
  Z = "Z",
  X = "X",
  seed = 123,
  n_chains = 2,
  n_iter = 500
)
```

`Carlotti_and_Parast_2026_simulations_grid`
Simulation grid for Carlotti and Parast (2026)

Description

A dataset containing the grid of simulation results for Carlotti and Parast (2026) across two covariate settings: a binary covariate setting (setting 1) and a Gaussian covariate setting (setting 2). Each setting is run for 500 simulations, for a total of 1000 rows.

Usage

```
Carlotti_and_Parast_2026_simulations_grid
```

Format

A data frame with 1000 rows and 17 columns:

setting The index of the simulation setting: 1 for the binary covariate setting (`X_binary`) and 2 for the Gaussian covariate setting (`X_Gaussian`).

simulation The index of the simulation run.

n_sample The sample size used.

n_chains The number of MCMC chains used.

n_iterations The number of MCMC iterations.

n_simulations The total number of simulations per setting.

timestamp The timestamp of execution.

seed The random seed used.

BF_alternative The alternative hypothesis used for the Bayes factor.

Bayesian_epsilon The threshold for the Bayesian test.

frequentist_epsilon The threshold for the frequentist test.

Bayesian_CI_upper The upper bound of the Bayesian credible interval.

frequentist_CI_upper The upper bound of the frequentist confidence interval.

Bayesian_coverage Logical. Indicates if the true value is in the credible interval.

frequentist_coverage Logical. Indicates if the true value is in the confidence interval.

Bayesian_power Logical. Indicates if the Bayesian test rejected the null.

frequentist_power Logical. Indicates if the frequentist test rejected the null.

Source

Generated using the code in the `simulations` folder of the BSET GitHub repository (https://github.com/PietroCarlotti/BSET).

`compute_BF_distribution` *Compute the Distribution of the Bayes Factor from Carlotti and Parast (2026)*

Description

This function calculates the probability mass function and cumulative distribution function of the Bayes factor defined in Carlotti and Parast (2026) for the following hypothesis test:

$$\begin{cases} H_0 : V_S = V_S^0 \\ H_1 : V_S > V_S^0 \end{cases}$$

where V_S is the surrogate's treatment effect on S measured as the probability

$$V_S = P(S_{1i} > S_{0i})$$

and V_S^0 is a hypothesized value under the null hypothesis. These hypotheses can be tested by fitting the following Beta-binomial model to the data:

$$\hat{V}_S \mid V_S \sim \text{Binomial}(n, V_S)$$

$$p(V_S) = \frac{\text{Beta}(V_S \mid a, b)}{\int_{1/2}^{1} \text{Beta}(v \mid a, b)\, dv}, \quad V_S \in \left(\tfrac{1}{2}, 1\right),$$

where $\hat{V}_S$ is the sample estimate of the surrogate's treatment effect on S computed as

$$\hat{V}_S = \frac{1}{n} \sum_{i=1}^{n} I(S_{1i} > S_{0i}).$$

The Bayes factor is then computed as the ratio of the marginal likelihoods under the alternatives:

$$BF_n = \frac{B(a + n\hat{V}_S,\, b + n - n\hat{V}_S) \left(1 - F_{\text{Beta}(a+n\hat{V}_S,\, b+n-n\hat{V}_S)}\left(\frac{1}{2}\right)\right)}{B(a, b) \left(1 - F_{\text{Beta}(a,b)}\left(\frac{1}{2}\right)\right)} \cdot 2^n,$$

where $B(a, b)$ is the Beta function, $F_{\text{Beta}(a,b)}$ is the cumulative distribution function of the $\text{Beta}(a, b)$ distribution, and a and b are the shape parameters of the Beta prior. Given the true value of V_S, the distribution of the Bayes factor can be computed by evaluating BF_n for all possible values of $\hat{V}_S$ and their corresponding probabilities under the Binomial distribution with parameters n and the true value of V_S. This function is generally not intended to be called directly by the user and is instead used internally within `BSET_no_X` and `BSET_X`.

Usage

```
compute_BF_distribution(
  n,
  V_S_true,
  V_S_zero = 0.5,
  a = 1,
  b = 1,
  BF_alternative
)
```

Arguments

n	Integer. The sample size.
V_S_true	Numeric. The true value of treatment effect on the surrogate.
V_S_zero	Numeric. The hypothesized value of the surrogate's treatment effect under the null hypothesis (default is 0.5).
a	Numeric. First shape parameter alpha for the Beta prior (default is 1).
b	Numeric. Second shape parameter beta for the Beta prior (default is 1).
BF_alternative	Character. The type of alternative hypothesis: either "two_sided" or "greater".

Value

A data frame containing:

- BF_values: The possible values of the Bayes Factor.
- BF_PMF: The probability mass function for the Bayes Factor.
- BF_CDF: The cumulative distribution function for the Bayes Factor.

References

Carlotti P, Parast L (2026). "A Bayesian Critique of Rank-Based Methods for Surrogate Marker Evaluation." *arXiv preprint arXiv:2603.14381*.

---

compute_delta *Monte Carlo Computation of the Parameter δ from Parast et al. (2024)*

Description

This function implements a Monte Carlo approach to estimate the parameter δ from Parast et al. (2024). This parameter represents the difference in treatment effects between the primary and surrogate outcomes, both measured using the Mann-Whitney statistic.

Usage

```
compute_delta(MC_data)
```

Arguments

MC_data	A list containing: • P_observed: A data frame or matrix with columns "Y" and "S". • Z: Treatment assignment vector. • n1: Number of treated units. • n0: Number of control units.

Details

The function processes data from a chosen data generating process, computing the Mann-Whitney U statistic for both the primary outcome Y and the surrogate S:

$$\hat{U}_Y = \frac{1}{n_1 n_0} \sum_{i:Z_i=1} \sum_{j:Z_j=0} I(Y_i > Y_j),$$

$$\hat{U}_S = \frac{1}{n_1 n_0} \sum_{i:Z_i=1} \sum_{j:Z_j=0} I(S_i > S_j).$$

Then, it calculates

$$\hat{\delta} = \hat{U}_Y - \hat{U}_S.$$

This function is generally not intended to be called directly by the user and is instead used internally within `BSET_no_X` and `BSET_X`.

Value

A list containing:

- `U_Y`: Mann-Whitney U statistic for the primary outcome Y computed on `P_observed`.
- `U_S`: Mann-Whitney U statistic for the surrogate S computed on `P_observed`.
- `delta`: The difference `U_Y - U_S`.

References

Parast L, Cai T, Tian L (2024). "A rank-based approach to evaluate a surrogate marker in a small sample setting." *Biometrics*, **80**(1), ujad035.

`compute_estimands_Carlotti_and_Parast_2026`

Monte Carlo Computation of the Estimands for the Simulation Study in Carlotti and Parast (2026)

Description

This function iterates through the simulation settings defined in Carlotti and Parast (2026) and estimates the true values of U_Y, U_S, δ, V_Y, V_S, and θ using a Monte Carlo dataset generated according to the specified data-generating processes.

Usage

```
compute_estimands_Carlotti_and_Parast_2026(MC_samples)
```

Arguments

`MC_samples`	Integer. The number of Monte Carlo samples to generate per setting.

Details

The settings are defined as follows:

- Setting 1: **X binary**
  - If $X = 1$:
$$Y_1 \sim \mathcal{N}(5,1), \quad Y_0 \sim \mathcal{N}(0,1)$$
$$S_1 = Y_1 + \mathcal{N}(0,1), \quad S_0 = Y_0 + \mathcal{N}(0,1)$$
  - If $X = 0$:
$$Y_1 \sim \mathcal{N}(5,1), \quad Y_0 \sim \mathcal{N}(0,1)$$
$$S_1 = Y_1 + \mathcal{N}(-10,1), \quad S_0 = Y_0 + \mathcal{N}(-10,1)$$
- Setting 2: **X Gaussian**
$$(Y_1, S_1, Y_0, S_0) \mid X = x \sim \mathcal{N}(x(1,7,0,6)', \frac{1}{2} I_4),$$

This function is generally not intended to be called directly by the user. It is provided as a utility for computing the true parameter values for the simulation settings described in Carlotti and Parast (2026).

Value

A data frame containing the Monte Carlo estimates for each setting:

- `setting`: The index of the simulation setting.
- `U_Y_MC`, `U_S_MC`, `delta_MC`: Parameters of interest from Parast et al. (2024).
- `V_Y_MC`, `V_S_MC`, `theta_MC`: Parameters of interest from Carlotti and Parast (2026).

References

Carlotti P, Parast L (2026). "A Bayesian Critique of Rank-Based Methods for Surrogate Marker Evaluation." *arXiv preprint arXiv:2603.14381*.

`compute_estimands_Parast_et_al_2024`
Monte Carlo Computation of the Estimands for the Simulation Study in Parast et al. (2024)

Description

This function iterates through the four simulation settings defined in Parast et al. (2024) and estimates the true values of U_Y, U_S, δ, V_Y, V_S, and θ using a Monte Carlo dataset generated according to the specified data-generating processes.

Usage

```
compute_estimands_Parast_et_al_2024(MC_samples)
```

Arguments

`MC_samples`	Integer. The number of Monte Carlo samples to generate per setting.

Details

The settings are defined as follows:

- Setting 1: **Useless surrogate (Gaussian model)**

$$Y_1 \sim \mathcal{N}(5,3), \quad Y_0 \sim \mathcal{N}(3,3),$$

$$S_1 \sim \mathcal{N}(2/3,1), \quad S_0 \sim \mathcal{N}(2/3,1).$$

- Setting 2: **Perfect surrogate (Gaussian model)**

$$Y_1 \sim \mathcal{N}(6,3), \quad Y_0 \sim \mathcal{N}(5/2,3),$$

$$S_1 = Y_1 + \mathcal{N}(0,1/10), \quad S_0 = Y_0 + \mathcal{N}(0,1/10).$$

- Setting 3: **Imperfect surrogate (Gaussian model)**

$$S_1 \sim \mathcal{N}(5,3), \quad S_0 \sim \mathcal{N}(3,3),$$

$$Y_1 = S_1 + \mathcal{N}(1.5,0.6), \quad Y_0 = S_0 + \mathcal{N}(0,0.6).$$

- Setting 4: **Misspecified model (non-Gaussian model)**

$$S_1 \sim \exp(\mathcal{N}(2.5,1.5)), \quad S_0 \sim \exp(\mathcal{N}(0.5,1.5)),$$

$$Y_1 = 2 + 6/5\sqrt{S_1} + 3/10\exp(S_1/500) + \exp(\mathcal{N}(0,0.3)),$$

$$Y_0 = 4/5\sqrt{S_0} + 1/5\exp(S_0/50) + \exp(\mathcal{N}(0,0.3)).$$

This function is generally not intended to be called directly by the user. It is provided as a utility for computing the true parameter values for the simulation settings described in Parast et al. (2024).

Value

A data frame containing the Monte Carlo estimates for each setting:

- `setting`: The index of the simulation setting.
- `U_Y_MC`, `U_S_MC`, `delta_MC`: Parameters of interest from Parast et al. (2024).
- `V_Y_MC`, `V_S_MC`, `theta_MC`: Parameters of interest from Carlotti and Parast (2026).

References

Parast L, Cai T, Tian L (2024). "A rank-based approach to evaluate a surrogate marker in a small sample setting." *Biometrics*, **80**(1), ujad035.

`compute_theta` *Monte Carlo Computation of the Parameter θ from Carlotti and Parast (2026)*

Description

This function implements a Monte Carlo approach to estimate the parameter θ from Carlotti and Parast (2026). This parameter represents the difference in treatment effects between the primary and surrogate outcomes, both measured using the probability that the treated outcome is larger than the control outcome.

Usage

```
compute_theta(MC_data)
```

Arguments

`MC_data`	A list containing: • `P`: A matrix or data frame of potential outcomes with columns "Y1", "Y0", "S1", and "S0".

Details

The function processes data from a chosen data generating process, computing the sample probabilities for both the primary outcome Y and the surrogate S:

$$\hat{V}_Y = \frac{1}{n}\sum_{i=1}^{n} I(Y_{1i} > Y_{0i}),$$

$$\hat{V}_S = \frac{1}{n}\sum_{i=1}^{n} I(S_{1i} > S_{0i}).$$

Then, it calculates

$$\hat{\theta} = \hat{V}_Y - \hat{V}_S.$$

This function is generally not intended to be called directly by the user and is instead used internally within `BSET_no_X` and `BSET_X`.

Value

A list containing the true values:

- `V_Y`: The Monte Carlo estimate of $P(Y_{1i} > Y_{0i})$ computed on `P`.
- `V_S`: The Monte Carlo estimate of $P(S_{1i} > S_{0i})$ computed on `P`.
- `theta`: The difference `V_Y - V_S`.

References

Carlotti P, Parast L (2026). "A Bayesian Critique of Rank-Based Methods for Surrogate Marker Evaluation." *arXiv preprint arXiv:2603.14381*.

`compute_V_S_star` *Compute v_S from Carlotti and Parast (2026)*

Description

This function determines the value v_S that is used to compute the surrogate validation threshold η from Carlotti and Parast (2026):

$$\eta = \max\{v_Y - v_S, 0\},$$

where v_Y is the hypothesized value of the treatment effect on the primary outcome (typically set equal to the estimate computed on the available data) and v_S is the value that satisfies the following power constraint:

$$P(\text{BF}_n \geq \text{BF}_{n,\alpha} \mid V_S = v_S) = 1 - \beta,$$

where $\text{BF}_{n,\alpha}$ is the $(1-\alpha)$ quantile of the Bayes factor distribution under the null hypothesis $V_S = V_S^0$, and $1-\beta$ is the desired power of the test. The function computes the distribution of the Bayes factor under the null hypothesis, derives the critical value $\text{BF}_{n,\alpha}$, and then uses a root-finding algorithm to solve for the value of v_S that satisfies the power constraint. This function is generally not intended to be called directly by the user and is instead used internally within `BSET_no_X` and `BSET_X`.

Usage

```
compute_V_S_star(
  n,
  alpha = 0.05,
  beta = 0.2,
  V_S_zero = 0.5,
  a = 1,
  b = 1,
  BF_alternative = "greater",
  root_tolerance = 1e-16
)
```

Arguments

`n`	Integer. Sample size.
`alpha`	Numeric. Type I error rate (default is 0.05).
`beta`	Numeric. Type II error rate (default is 0.2).
`V_S_zero`	Numeric. The hypothesized value of the surrogate's treatment effect under the null hypothesis (default is 0.5).
`a`	Numeric. First shape parameter alpha for the Beta prior (default is 1).
`b`	Numeric. Second shape parameter beta for the Beta prior (default is 1).
`BF_alternative`	Character. The type of alternative hypothesis: either `"two_sided"` or `"greater"`.
`root_tolerance`	Numeric. Tolerance level for the root-finding algorithm (default is 1e-16).

Value

A list containing:

- `BF_alpha`: The critical value of the Bayes factor corresponding to the specified alpha level.
- `V_S_star`: The value of v_S that satisfies the power constraint for the surrogate validation test.

References

Carlotti P, Parast L (2026). "A Bayesian Critique of Rank-Based Methods for Surrogate Marker Evaluation." *arXiv preprint arXiv:2603.14381*.

`DGP_no_X` *Data Generating Process without Baseline Covariates*

Description

This function generates potential outcomes from the simulation settings described in Parast et al. (2024). It creates a dataset of potential outcomes

$$P = (Y_1, S_1, Y_0, S_0)$$

and observed outcomes

$$P_{observed} = (Y, S)$$

based on a random treatment assignment Z.

Usage

```
DGP_no_X(
  n,
  p,
  mu_star = NULL,
  Sigma_star = NULL,
  model = c("Gaussian", "misspecified")
)
```

Arguments

`n`	Integer. Total sample size.
`p`	Numeric. Probability of being assigned to the treatment group (Z=1).
`mu_star`	Numeric vector. The mean vector for P. Required if `model = "Gaussian"`.
`Sigma_star`	Matrix. The covariance matrix for P. Required if `model = "Gaussian"`.
`model`	Character. The type of data generation: `"Gaussian"` or `"misspecified"`.

Details

The function supports two types of data-generating processes:

- **Gaussian model:** Potential outcomes are drawn from a multivariate normal distribution:

$$P \sim \mathcal{N}_4(\mu^*, \Sigma^*).$$

- **Non-linear model:** Potential outcomes for the surrogate are generated from a non-Gaussian distribution, and the potential outcomes for the primary outcome are generated from a non-linear function of the surrogate plus non-Gaussian noise.

Value

A list containing:

- `Z`: Treatment assignment vector.
- `n1`: Number of treated units.
- `n0`: Number of control units.
- `P`: Full matrix of potential outcomes.
- `P_observed`: Observed outcomes (Y, S) corresponding to the assigned treatment Z.
- `P_unobserved`: Counterfactual outcomes under the opposite treatment.

This function is useful for generating synthetic data to test or explore the method, for instance to verify the behavior of `BSET_no_X` under known simulation settings.

References

Parast L, Cai T, Tian L (2024). "A rank-based approach to evaluate a surrogate marker in a small sample setting." *Biometrics*, **80**(1), ujad035.

Examples

```
set.seed(123)
data <- DGP_no_X(
  n = 100,
  p = 0.5,
  mu_star = c(6, 6, 2.5, 2.5),
  Sigma_star = kronecker(diag(2), matrix(c(3, 3, 3, 3.1), 2, 2)),
  model = "Gaussian"
)
```

DGP_X_binary *Data Generating Process with a Binary Covariate*

Description

This function generates potential outcomes from a data generating process similar to the one described in Parast et al. (2024), but with the addition of a binary covariate X. It creates a dataset of potential outcomes

$$P = (Y_1, S_1, Y_0, S_0)$$

and observed outcomes

$$P_{observed} = (Y, S)$$

based on a random treatment assignment Z.

Usage

```
DGP_X_binary(n, p, q, mu_0, mu_1, Sigma_0, Sigma_1)
```

Arguments

n	Integer. Total sample size.
p	Numeric. Probability of being assigned to the treatment group ($Z = 1$).
q	Numeric. Probability of the binary covariate X being 1.
mu_0	Numeric vector. Mean vector for P when $X = 0$.
mu_1	Numeric vector. Mean vector for P when $X = 1$.
Sigma_0	Matrix. Covariance matrix for P when $X = 0$.
Sigma_1	Matrix. Covariance matrix for P when $X = 1$.

Details

The potential outcomes are generated from multivariate normal distributions with different mean vectors and covariance matrices depending on the value of X. Specifically:

$$P \mid X = 0 \sim \mathcal{N}_4(\mu^0, \Sigma^0), P \mid X = 1 \sim \mathcal{N}_4(\mu^1, \Sigma^1).$$

Value

A list containing:

- X: The binary covariate vector.
- n_X1: Number of units with $X = 1$.
- n_X0: Number of units with $X = 0$.
- Z: Treatment assignment vector.
- n1: Number of treated units.
- n0: Number of control units.

- P: Full matrix of potential outcomes.
- P_observed: Observed outcomes (Y, S) corresponding to the assigned treatment Z.
- P_unobserved: Counterfactual outcomes under the opposite treatment.

This function is useful for generating synthetic data to test or explore the method, for instance to verify the behavior of BSET_X under known simulation settings.

References

Carlotti P, Parast L (2026). "A Bayesian Critique of Rank-Based Methods for Surrogate Marker Evaluation." *arXiv preprint arXiv:2603.14381*.

Examples

```
set.seed(123)
data <- DGP_X_binary(
  n = 100,
  p = 0.5,
  q = 0.5,
  mu_0 = c(5, 5, 0, 0),
  mu_1 = c(5, -5, 0, -10),
  Sigma_0 = kronecker(diag(2), matrix(c(1, 1, 1, 2), 2, 2)),
  Sigma_1 = kronecker(diag(2), matrix(c(1, 1, 1, 2), 2, 2))
)
```

DGP_X_Gaussian *Data Generating Process with a Gaussian Covariate*

Description

This function generates potential outcomes from a data generating process similar to the one described in Parast et al. (2024), but with the addition of a Gaussian covariate X. It creates a dataset of potential outcomes

$$P = (Y_1, S_1, Y_0, S_0)$$

and observed outcomes

$$P_{observed} = (Y, S)$$

based on a random treatment assignment Z.

Usage

```
DGP_X_Gaussian(n, p, beta, Sigma, m, s)
```

Arguments

n	Integer. Total sample size.
p	Numeric. Probability of being assigned to the treatment group ($Z = 1$).
beta	Numeric vector. Coefficients for the linear function of the covariate X.
Sigma	Matrix. Covariance matrix for the potential outcomes.
m	Numeric. Mean of the Gaussian covariate X.
s	Numeric. Standard deviation of the Gaussian covariate X.

Details

The potential outcomes are generated from a multivariate normal distribution with mean vector and covariance matrix that depend on the value of X. Specifically, the mean vector is a linear function of X:

$$\mu(X) = x \cdot (\beta_{Y1}, \beta_{S1}, \beta_{Y0}, \beta_{S0})^T,$$

and the covariance matrix is constant across values of X:

$$\Sigma(X) = \Sigma.$$

Value

A list containing:

- X: The Gaussian covariate vector.
- Z: Treatment assignment vector.
- n1: Number of treated units.
- n0: Number of control units.
- P: Full matrix of potential outcomes.
- P_observed: Observed outcomes (Y, S) corresponding to the assigned treatment Z.
- P_unobserved: Counterfactual outcomes under the opposite treatment.

This function is useful for generating synthetic data to test or explore the method, for instance to verify the behavior of BSET_X under known simulation settings.

References

Carlotti P, Parast L (2026). "A Bayesian Critique of Rank-Based Methods for Surrogate Marker Evaluation." *arXiv preprint arXiv:2603.14381*.

Examples

```
set.seed(123)
data <- DGP_X_Gaussian(
  n = 100,
  p = 0.5,
  beta = c(1, 7, 0, 6),
  Sigma = 0.5 * diag(4),
  m = 3,
  s = 1
)
```

estimands_Carlotti_and_Parast_2026

True estimands for Carlotti and Parast (2026)

Description

A dataset containing the Monte Carlo estimates of the parameters of interest for the two simulation settings considered in Carlotti and Parast (2026): setting 1 (binary covariate) and setting 2 (Gaussian covariate).

Usage

```
estimands_Carlotti_and_Parast_2026
```

Format

A data frame with 2 rows and 7 columns:

setting The index of the simulation setting.

U_Y_MC Monte Carlo estimate of $U_Y = P(Y_{1i} > Y_{0j})$.

U_S_MC Monte Carlo estimate of $U_S = P(S_{1i} > S_{0j})$.

delta_MC Monte Carlo estimate of $\delta = U_Y - U_S$.

V_Y_MC Monte Carlo estimate of $V_Y = P(Y_{1i} > Y_{0i})$.

V_S_MC Monte Carlo estimate of $V_S = P(S_{1i} > S_{0i})$.

theta_MC Monte Carlo estimate of $\theta = V_Y - V_S$.

Source

Computed using the function `compute_estimands_Carlotti_and_Parast_2026(1000000)`.

estimands_Parast_et_al_2024

True estimands for Parast et al. (2024)

Description

A dataset containing the Monte Carlo estimates of the parameters of interest for the simulation settings considered in Parast et al. (2024).

Usage

```
estimands_Parast_et_al_2024
```

Format

A data frame with 4 rows and 7 columns:

setting The index of the simulation setting.
U_Y_MC Numeric vector of Monte Carlo estimates.
U_S_MC Numeric vector of Monte Carlo estimates.
delta_MC Numeric vector of Monte Carlo estimates.
V_Y_MC Numeric vector of Monte Carlo estimates.
V_S_MC Numeric vector of Monte Carlo estimates.
theta_MC Numeric vector of Monte Carlo estimates.

Source

Computed using the function `compute_estimands_Parast_et_al_2024(1000000)`.

`frequentist_test` *Frequentist Test from Parast et al. (2024)*

Description

This function performs a frequentist test for surrogate evaluation based on the rank-based methodology of Parast et al. (2024). It calculates confidence intervals, the ε threshold used, and determines if the surrogate is valid. This function is generally not intended to be called directly by the user and is instead used internally within `BSET_no_X` and `BSET_X`.

Usage

```
frequentist_test(P_observed, Z, alpha = 0.05, beta = 0.2, delta_true = NULL)
```

Arguments

`P_observed`	Observed outcomes (Y, S) corresponding to the assigned treatment Z.
`Z`	Treatment assignment vector.
`alpha`	Numeric. Significance level for the confidence interval (default is 0.05).
`beta`	Numeric. Type II error rate (default is 0.2).
`delta_true`	Numeric (optional). The true value of δ, used to calculate frequentist coverage during simulations.

Value

A list containing:

- `CI`: The calculated confidence interval for δ.
- `threshold`: The ε threshold value used in the test.
- `coverage`: Logical indicating if `delta_true` falls within the `CI` (if `delta_true` is provided).
- `power`: Logical indicating if the upper bound of `CI` is below `threshold`, which indicates that the test identifies the surrogate as valid.

References

Parast L, Cai T, Tian L (2024). "A rank-based approach to evaluate a surrogate marker in a small sample setting." *Biometrics*, **80**(1), ujad035.

`Parast_et_al_2024_simulations_grid`
Simulation grid for Parast et al. (2024)

Description

A dataset containing the grid of simulation results for Parast et al. (2024).

Usage

```
Parast_et_al_2024_simulations_grid
```

Format

A data frame with multiple columns:

setting The index of the simulation setting.

simulation The index of the simulation run.

n_sample The sample size used.

n_chains The number of MCMC chains used.

n_iterations The number of MCMC iterations.

n_simulations The total number of simulations.

timestamp The timestamp of execution.

seed The random seed used.

BF_alternative The alternative hypothesis used for the Bayes factor.

Bayesian_epsilon The threshold for the Bayesian test.

frequentist_epsilon The threshold for the frequentist test.

Bayesian_CI_upper The upper bound of the Bayesian credible interval.

frequentist_CI_upper The upper bound of the frequentist confidence interval.

Bayesian_coverage Logical. Indicates if the true value is in the credible interval.

frequentist_coverage Logical. Indicates if the true value is in the confidence interval.

Bayesian_power Logical. Indicates if the Bayesian test rejected the null.

frequentist_power Logical. Indicates if the frequentist test rejected the null.

Source

Generated using the code in the `simulations` folder of the BSET GitHub repository (https://github.com/PietroCarlotti/BSET).

Index

∗ **datasets**
Carlotti_and_Parast_2026_simulations_grid, 8
estimands_Carlotti_and_Parast_2026, 21
estimands_Parast_et_al_2024, 21
Parast_et_al_2024_simulations_grid, 23

Bayesian_test, 2
BSET, 3

Carlotti_and_Parast_2026_simulations_grid, 7
compute_BF_distribution, 9
compute_delta, 10
compute_estimands_Carlotti_and_Parast_2026, 11
compute_estimands_Parast_et_al_2024, 12
compute_theta, 14
compute_V_S_star, 15

DGP_no_X, 16
DGP_X_binary, 18
DGP_X_Gaussian, 19

estimands_Carlotti_and_Parast_2026, 21
estimands_Parast_et_al_2024, 21

frequentist_test, 22

Parast_et_al_2024_simulations_grid, 23