Package 'HOIFCar'

June 25, 2025

Type Package

Title Covariate Adjustment in RCT by Higher-Order Influence Functions

Version 0.2.1

Maintainer Xinbo Wang <cinbo_w@sjtu.edu.cn>

Description Estimates treatment effects using covariate adjustment methods in Randomized Clinical Trials (RCT) motivated by higher-order influence functions (HOIF). Provides point estimates, oracle bias, variance, and approximate variance for HOIF-adjusted estimators. For methodology details, see Zhao et al. (2024) <doi:10.48550/arXiv.2411.08491>.

License MIT + file LICENSE

Suggests mvtnorm

Encoding UTF-8

RoxygenNote 7.3.2

NeedsCompilation no

Author Sihui Zhao [aut],
Xinbo Wang [cre, aut],
Liu Liu [ctb]

Repository CRAN

Date/Publication 2025-06-25 15:40:06 UTC

Contents

`esti_mean_treat` *Estimate treatment effect and the corresponding variance estimation on the treatment arm using different covariate adjustment methods.*

Description

Implements a unified framework for comparing covariate adjustment method for completely randomized experiments under randomization-based framework.

Usage

```
esti_mean_treat(X, Y, A, H = NULL)
```

Arguments

`X`	The n by p covariates matrix.
`Y`	Vector of n dimensional observed response.
`A`	Vector of n dimensional treatment assignment.
`H`	The n by n hat projection matrix corresponding to X.

Value

A list with two named vectors:

point_est Point estimates for all estimators:

- `unadj`: Unadjusted estimator
- `db`: Debiased estimator (Lu et al., 2023)
- `adj2c`: HOIF-inspired debiased estimator (Zhao et al., 2024), the same as `db`
- `adj2`: HOIF-motivated adjusted estimator (Zhao et al., 2024)
- `adj3`: Bias-free adjusted estimator based on `adj2`
- `lin`: Covariate-adjusted estimator (Lin, 2013)
- `lin_db`: Debiased estimator with population leverage scores (Lei, 2020)

var_est Variance estimates corresponding to each estimator:

- `unadj`: Variance estimate for unadjusted estimator
- `db`: Variance estimate for debiased estimator (Lu et al., 2023)
- `adj2c`: Variance for `adj2c`, using formulas given in (Lu et al., 2023)
- `adj2c_v2`: Conservative variance for `adj2c` (Zhao et al., 2024)
- `adj2`: Variance for `adj2`, with formulas motivated by (Lu et al., 2023)
- `adj2_v2`: Conservative variance for `adj2` (Zhao et al., 2024)
- `adj3`: Variance for `adj3`, with formulas motivated by (Lu et al., 2023)
- `adj3_v2`: Conservative variance for `adj3` (Zhao et al., 2024)
- `lin`: HC3-type variance for Lin's (2013) estimator
- `lin_db`: HC3-type variance for Lei's (2020) estimator

References

Lin, W. (2013). *Agnostic notes on regression adjustments to experimental data: Reexamining Freedman's critique. The Annals of Statistics, Vol. 7(1), 295–318*, doi:10.1214/12AOAS583.
Lei, L. and Ding, P. (2020) *Regression adjustment in completely randomized experiments with a diverging number of covariates. Biometrika, Vol. 108(4), 815–828*, doi:10.1093/biomet/asaa103.
Lu, X., Yang, F. and Wang, Y. (2023) *Debiased regression adjustment in completely randomized experiments with moderately high-dimensional covariates. arXiv preprint, arXiv:2309.02073*, doi:10.48550/arXiv.2309.02073.
Zhao, S., Wang, X., Liu, L. and Zhang, X. (2024) *Covariate Adjustment in Randomized Experiments Motivated by Higher-Order Influence Functions. arXiv preprint, arXiv:2411.08491*, doi:10.48550/arXiv.2411.08491.

Examples

```
set.seed(100)
n <- 500
p <- n * 0.3
beta <- runif(p, -1 / sqrt(p), 1 / sqrt(p))

X <- mvtnorm::rmvt(n, sigma = diag(1, p), df = 3)
Y1 <- as.numeric(X %*% beta)
Y0 <- rep(0, n)

pi1 <- 2/3
n1 <- ceiling(n * pi1)
ind <- sample(n, size = n1)
A <- rep(0, n)
A[ind] <- 1
Y <- Y1 * A + Y0 * (1 - A)

Xc_svd <- svd(X)
H <- Xc_svd$u %*% t(Xc_svd$u)

result_ls <- esti_mean_treat(X, Y, A, H)
point_est <- result_ls$point_est
var_est <- result_ls$var_est
print(paste0('True mean treat:', round(mean(Y1), digits = 3), '.'))
print('Absolute bias:')
print(abs(point_est - mean(Y1)))
print('Estimate variance:')
print(var_est)
```

`fit.adj2.adj2c.Random` *Covariate-Adjusted Treatment Effect Estimation under the Randomization-based Framework*

Description

Implements the (HOIF-inspired) debiased estimators for average treatment effect (ATE) or treatment effect on the treatment/control arm with variance estimation using asymptotic-variance. Designed for randomized experiments with moderately high-dimensional covariates.

Usage

```
fit.adj2.adj2c.Random(Y, X, A, pi1 = NULL, target = "ATE")
```

Arguments

`Y`	Numeric vector of length n containing observed responses.
`X`	Numeric matrix (n x p) of covariates. Centering is required. Intercept term can include or not.
`A`	Binary vector of length n indicating treatment assignment (1 = treatment, 0 = control).
`pi1`	Default is NULL. The assignment probability for the randomization assignment.
`target`	A character string specifying the target estimand. Must be one of: - ‘"ATE"‘ (default): Average Treatment Effect (difference between treatment and control arms). - ‘"EY1"‘: Expected outcome under treatment (estimates the effect for the treated group). - ‘"EY0"‘: Expected outcome under control (estimates the effect for the control group).

Value

A list containing three named vectors, including point estimates and variance estimates:

tau_vec Point estimates:

- `adj2`: Point estimation of the HOIF-inspired debiased estimator given by Zhao et al.(2024).
- `adj2c`: Point estimation of the debiased estimator given by Lu et al. (2023), which is also the HOIF-inspired debiased estimator given by Zhao et al.(2024).

var_vec_v1 Variance estimates for adj2 and adj2c, with formulas inspired by Lu et al. (2023).:

- `adj2`: Variance for `adj2`.
- `adj2c`: Variance for `adj2c`.

var_vec_v2 Variance estimates for adj2 and adj2c, with formulas given in Zhao et al. (2024), which is more conservative.

- `adj2`: Variance for `adj2`.
- `adj2c`: Variance for `adj2c`.

References

Lu, X., Yang, F. and Wang, Y. (2023) *Debiased regression adjustment in completely randomized experiments with moderately high-dimensional covariates. arXiv preprint, arXiv:2309.02073*, doi:10.48550/arXiv.2309.02073.
Zhao, S., Wang, X., Liu, L. and Zhang, X. (2024) *Covariate Adjustment in Randomized Experiments Motivated by Higher-Order Influence Functions. arXiv preprint, arXiv:2411.08491*, doi:10.48550/arXiv.2411.08491.

Examples

```
set.seed(100)
n <- 500
p <- n * 0.3
beta <- runif(p, -1 / sqrt(p), 1 / sqrt(p))

X <- mvtnorm::rmvt(n, sigma = diag(1, p), df = 3)
Y1 <- as.numeric(X %*% beta)
Y0 <- as.numeric(X %*% beta - 1)

pi1 <- 2/3
n1 <- ceiling(n * pi1)
ind <- sample(n, size = n1)
A <- rep(0, n)
A[ind] <- 1
Y <- Y1 * A + Y0 * (1 - A)

Xc <- cbind(1, scale(X, scale = FALSE))
result.adj2.adj2c.random.ate.ls <- fit.adj2.adj2c.Random(Y, Xc, A, target = 'ATE')
result.adj2.adj2c.random.ate.ls

result.adj2.adj2c.random.treat.ls <- fit.adj2.adj2c.Random(Y, Xc, A, target = 'EY1')
result.adj2.adj2c.random.treat.ls

result.adj2.adj2c.random.control.ls <- fit.adj2.adj2c.Random(Y, Xc, A, target = 'EY0')
result.adj2.adj2c.random.control.ls
```

---

fit.adj2.adj2c.Super *Covariate-Adjusted Treatment Effect Estimation under the Super-Population Framework*

Description

Implements HOIF-inspired debiased estimators for average treatment effect (ATE) or treatment effect on the treatment/control arm with variance estimation using influence function-based and asymptotic-variance. Designed for randomized experiments with moderately high-dimensional covariates.

Usage

```
fit.adj2.adj2c.Super(
  Y,
  X,
  A,
  intercept = TRUE,
  pi1 = NULL,
  target = "ATE",
  lc = FALSE
)
```

Arguments

`Y`	Numeric vector of length n containing observed responses.
`X`	Numeric matrix (n x p) of covariates. Centering is required. May include intercept column.
`A`	Binary vector of length n indicating treatment assignment (1 = treatment, 0 = control).
`intercept`	Logical. If TRUE (default), X already contains intercept. Set FALSE if X does not contain intercept.
`pi1`	Default is NULL. The assignment probability for the randomization assignment.
`target`	A character string specifying the target estimand. Must be one of: - ‘"ATE"‘ (default): Average Treatment Effect (difference between treatment and control arms). - ‘"EY1"‘: Expected outcome under treatment (estimates the effect for the treated group). - ‘"EY0"‘: Expected outcome under control (estimates the effect for the control group).
`lc`	Default is FALSE. If TRUE, then performs linear calibration to achieve efficiency gain using $\hat{\mu}_0(X_i)$ and $\hat{\mu}_1(X_i)$.

Value

A list containing three named vectors, including point estimates and variance estimates:

tau_vec Point estimates:

- `adj2`: Point estimation of the HOIF-inspired debiased estimator (Zhao et al., 2024).
- `adj2c`: Point estimation of the the HOIF-inspired debiased estimator (Zhao et al., 2024), which is also the debiased estimator given by Lu et al. (2023).

var_infl_vec Influence function-based variance estimates:

- `adj2`: Variance for `adj2` via the sample variance of its influence function formula.
- `adj2c`: Variance for `adj2c` via the sample variance of its influence function formula.

var_rb_vec Variance estimates inspired by Bannick et al. (2025):

- `adj2`: Variance for `adj2` following the asymptotic variance given by Bannick et al. (2025).
- `adj2c`: Variance for `adj2c` following the asymptotic variance given by Bannick et al. (2025).

References

Bannick, M. S., Shao, J., Liu, J., Du, Y., Yi, Y. and Ye, T. (2025) *A General Form of Covariate Adjustment in Clinical Trials under Covariate-Adaptive Randomization. Biometrika, Vol. xx(x), 1-xx*, doi:10.1093/biomet/asaf029.

Lu, X., Yang, F. and Wang, Y. (2023) *Debiased regression adjustment in completely randomized experiments with moderately high-dimensional covariates. arXiv preprint, arXiv:2309.02073*, doi:10.48550/arXiv.2309.02073.

Zhao, S., Wang, X., Liu, L. and Zhang, X. (2024) *Covariate Adjustment in Randomized Experiments Motivated by Higher-Order Influence Functions. arXiv preprint, arXiv:2411.08491*, doi:10.48550/arXiv.2411.08491.

Examples

```
set.seed(120)
alpha0 <- 0.1;
n <- 400;

p0 <- ceiling(n * alpha0)
beta0_full <- 1 / (1:p0) ^ (1 / 2) * (-1) ^ c(1:p0)
beta <- beta0_full / norm(beta0_full,type='2')

Sigma_true <- matrix(0, nrow = p0, ncol = p0)
for (i in 1:p0) {
  for (j in 1:p0) {
    Sigma_true[i, j] <- 0.1 ** (abs(i - j))
  }
}

X <- mvtnorm::rmvt(n, sigma = Sigma_true, df = 3)

lp0 <- X %*% beta
delta_X <- 1  -  1/4 * X[, 2] -  1/8 * X[, 3]
lp1 <- lp0 + delta_X

Y0 <- lp0 + rnorm(n)
Y1 <- lp1 + rnorm(n)


pi1 <- 1 / 2
A <- rbinom(n, size = 1, prob = pi1)
Y <- A * Y1 + (1 - A) * Y0

Xc <- cbind(1, scale(X, scale = FALSE))
result.adj2.adj2c.sp.ate.ls <- fit.adj2.adj2c.Super(Y, Xc, A, intercept = TRUE,
                                                    target = 'ATE', lc = TRUE)
result.adj2.adj2c.sp.ate.ls
result.adj2.adj2c.sp.treat.ls <- fit.adj2.adj2c.Super(Y, Xc, A, intercept = TRUE,
                                                      target = 'EY1', lc = TRUE)
result.adj2.adj2c.sp.treat.ls
result.adj2.adj2c.sp.control.ls <- fit.adj2.adj2c.Super(Y, Xc, A, intercept = TRUE,
                                                        target = 'EY0', lc = TRUE)
result.adj2.adj2c.sp.control.ls
```

`get_oracle_bias_var_adj2c` *Estimate the oracle bias, the exact variance and approximated variance of the debiased estimator tau_adj2c inspired by HOIF (Zhao et al.(2024)).*

Description

Estimate the oracle bias, the exact variance and approximated variance of the debiased estimator tau_adj2c inspired by HOIF (Zhao et al.(2024)).

Usage

```
get_oracle_bias_var_adj2c(X, Y1, n1 = NULL)
```

Arguments

X	The n by p covariates matrix.
Y1	Vector of n dimensional potential response Y(1).
n1	The number of subjects in the treatment group.

Value

A list of oracle bias and variance of the debised adjusted estimator tau_adj2c.

bias_adj2c	The oracle bias of the debiased estimator tau_adj2c.
variance_exact_adj2c	The oracle exact bias of the debiased estimator tau_adj2c.
variance_approx_adj2c	The oracle approximated variance of the debiased estimator tau_adj2c which omits the term of order o(1/n).
variance_unadj	The oracle variance of the unadjusted estimator.

References

Zhao, S., Wang, X., Liu, L., & Zhang, X. (2024). Covariate adjustment in randomized experiments motivated by higher-order influence functions. arXiv preprint. https://arxiv.org/abs/2411.08491.

Examples

```
NULL
```

---

get_oracle_bias_var_adj_2_3

Estimate the oracle bias, the exact variance and approximated variance of the debiased estimator and the bias-free estimator motivated by HOIF (Zhao et al.(2024)).

Description

Estimate the oracle bias, the exact variance and approximated variance of the debiased estimator and the bias-free estimator motivated by HOIF (Zhao et al.(2024)).

Usage

```
get_oracle_bias_var_adj_2_3(X, Y1, n1 = NULL)
```

Arguments

X	The n by p covariates matrix.
Y1	Vector of n dimensional potential response Y(1).
n1	The number of subjects in the treatment group.

Value

A list of oracle bias and variance of the adjusted estimator motivated by HOIF and the bias-free estimator.

bias_adj2	The oracle bias of the estimator tau_adj2.
variance_exact_adj2	The oracle exact variance of the estimator tau_adj2.
variance_approx_adj2	The oracle approximated variance of the estimator tau_adj2 which omits the term of order o(1/n).
variance_exact_adj3	The oracle exact variance of the bias-free estimator tau_adj3.
variance_unadj	The oracle variance of the unadjusted estimator.

References

Zhao, S., Wang, X., Liu, L., & Zhang, X. (2024). Covariate adjustment in randomized experiments motivated by higher-order influence functions. arXiv preprint. https://arxiv.org/abs/2411.08491

Examples

```
# Linear setting
set.seed(100)
n <- 500
p <- 50
beta <- rt(p,3)

X <- mvtnorm::rmvt(n, sigma = diag(1, p), df = 3)
Y1 <- as.numeric(X %*% beta)
pi1 <- 0.50
n1 <- ceiling(n*pi1)

result_adj_db <- get_oracle_bias_var_adj_db(X = X,Y1=Y1,n1=n1)
result_adj2c <- get_oracle_bias_var_adj2c(X = X,Y1=Y1,n1=n1)
result_adj2_3 <- get_oracle_bias_var_adj_2_3(X = X,Y1=Y1,n1=n1)
unlist(result_adj_db)
unlist(result_adj2c)
unlist(result_adj2_3)



# Nonlinear setting
n <- 500;
```

```
alpha <- 0.2;
set.seed(1000)
p <- ceiling(n*alpha)
Sigma_true <- matrix(0,nrow=p,ncol=p)
for(i in 1:p){
  for(j in 1:p){
    Sigma_true[i,j] <- 0.1**(abs(i-j))
  }
}

X <- mvtnorm::rmvt(n, sigma = Sigma_true, df = 3)
beta <- rt(p,3)
or_baseline <- sign(X %*% beta) * abs(X %*% beta)^(1/2) + sin(X %*% beta)
epsilon1 <- epsilon0 <- rt(n,3)
Y1 <- 1 + as.numeric(or_baseline) + epsilon1


pi1 <- 0.50
n1 <- ceiling(n*pi1)

result_adj_db <- get_oracle_bias_var_adj_db(X = X,Y1=Y1,n1=n1) # from LYW paper
result_adj2c <- get_oracle_bias_var_adj2c(X = X,Y1=Y1,n1=n1)
result_adj2_3 <- get_oracle_bias_var_adj_2_3(X = X,Y1=Y1,n1=n1)
unlist(result_adj_db)
unlist(result_adj2c)
unlist(result_adj2_3)
```

`get_oracle_bias_var_adj_db` *Estimate the oracle bias, the oracle variance of the unadjusted estimator, the adjusted estimator by Lei's (2020) and the debiased estimator tau_db by Lu et al.(2023).*

Description

Estimate the oracle bias, the oracle variance of the unadjusted estimator, the adjusted estimator by Lei's (2020) and the debiased estimator tau_db by Lu et al.(2023).

Usage

```
get_oracle_bias_var_adj_db(X, Y1, n1 = NULL)
```

Arguments

`X`	The n by p covariates matrix.
`Y1`	Vector of n dimensional potential response Y(1).
`n1`	The number of subjects in the treatment group.

Value

A list of the oracle bias and variance of .

`bias_adj`	The oracle bias of the adjusted estimator tau_adj we proposed.
`variance_unadj`	The oracle variance of the unadjusted estimator.
`variance_adj_lin`	The oracle variance of Lei's (2020) debiased estimator with linear working model.
`variance_db`	The oracle variance of the debiased estimator tau_db by Lu et al.(2023).

References

Lihua Lei, Peng Ding. Regression adjustment in completely randomized experiments with a diverging number of covariates. Biometrika, 815–828, 2020.

Xin Lu, Fan Yang, and Yuhao Wang. Debiased regression adjustment in completely randomized experiments with moderately high-dimensional covariates. arXiv preprint arXiv:2309.02073, 2023.

Examples

```
NULL
```

Index