Isotonic subgroup selection

CCIMI 7$^{\text{th}}$ Annual Academic Conference

May 2023

Manuel Müller

University of Cambridge, CCIMI

Collaborators

Henry Reeve

Timothy Cannings

Richard Samworth

- M., M. M., Reeve, H. W. J., Cannings, T. I. and Samworth, R. J. (2023). Isotonic subgroup selection. arXiv:2305.04852.

Subgroup selection

Task:

Subgroup selection in regression: Identify a subset of the covariate domain on which the regression function satisfies a particular property (e.g. exceeds a threshold).

Subgroup selection

Task:

Subgroup selection in regression: Identify a subset of the covariate domain on which the regression function satisfies a particular property (e.g. exceeds a threshold).

Setting:

Fix $\sigma > 0$. Let $\mathcal{P}_{\mathrm{Mon},d}(\sigma)$ be the family of distributions P on $\mathbb{R}^d \times \mathbb{R}$ such that for $(X, Y) \sim P$,

Subgroup selection

Task:

Subgroup selection in regression: Identify a subset of the covariate domain on which the regression function satisfies a particular property (e.g. exceeds a threshold).

Setting:

Fix $\sigma > 0$. Let $\mathcal{P}_{\mathrm{Mon},d}(\sigma)$ be the family of distributions P on $\mathbb{R}^d \times \mathbb{R}$ such that for $(X, Y) \sim P$,

(i) the regression function $\eta(x) := \mathbb{E}(Y|X = x)$ is increasing* on $\mathbb{R}^d$.

*$x_0 \preccurlyeq x_1 \implies \eta(x_0) \leq \eta(x_1)$.

Subgroup selection

Task:

Subgroup selection in regression: Identify a subset of the covariate domain on which the regression function satisfies a particular property (e.g. exceeds a threshold).

Setting:

Fix $\sigma > 0$. Let $\mathcal{P}_{\mathrm{Mon},d}(\sigma)$ be the family of distributions P on $\mathbb{R}^d \times \mathbb{R}$ such that for $(X, Y) \sim P$,

(i) the regression function $\eta(x) := \mathbb{E}(Y|X = x)$ is increasing* on $\mathbb{R}^d$.

(ii) $Y - \eta(X) \mid X$ is sub-Gaussian with variance parameter σ^2.

---

*$x_0 \preccurlyeq x_1 \implies \eta(x_0) \leq \eta(x_1)$.

Statistical setting

Notation:

- Fix $\tau \in \mathbb{R}$. Define τ-superlevel set by $\mathcal{X}_\tau(\eta) := \{x \in \mathbb{R}^d : \eta(x) \geq \tau\}$.

Statistical setting

Notation:

- Fix $\tau \in \mathbb{R}$. Define τ-superlevel set by
  $\mathcal{X}_\tau(\eta) := \{x \in \mathbb{R}^d : \eta(x) \geq \tau\}$.
- μ denotes the marginal distribution of X.

Statistical setting

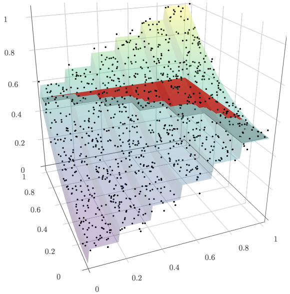
Notation:

- Fix $\tau \in \mathbb{R}$. Define τ-superlevel set by $\mathcal{X}_\tau(\eta) := \{x \in \mathbb{R}^d : \eta(x) \geq \tau\}$.
- μ denotes the marginal distribution of X.
- $\mathcal{D} := \big((X_1, Y_1), \ldots, (X_n, Y_n)\big)$.

Statistical setting

Notation:

- Fix $\tau \in \mathbb{R}$. Define τ-superlevel set by $\mathcal{X}_\tau(\eta) := \{x \in \mathbb{R}^d : \eta(x) \geq \tau\}$.
- μ denotes the marginal distribution of X.
- $\mathcal{D} := \big((X_1, Y_1), \ldots, (X_n, Y_n)\big)$.

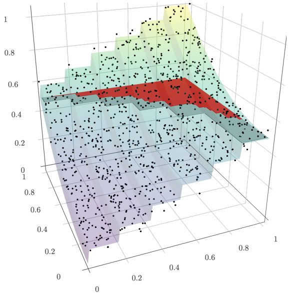


Goal: algorithm that returns a selected set $\hat{A}(\mathcal{D}) \subseteq \mathbb{R}^d$

Statistical setting

Notation:

- Fix $\tau \in \mathbb{R}$. Define τ-superlevel set by $\mathcal{X}_\tau(\eta) := \{x \in \mathbb{R}^d : \eta(x) \geq \tau\}$.
- μ denotes the marginal distribution of X.
- $\mathcal{D} := \big((X_1, Y_1), \ldots, (X_n, Y_n)\big)$.



Goal: algorithm that returns a selected set $\hat{A}(\mathcal{D}) \subseteq \mathbb{R}^d$ with:

- Type I error control: Fix $\alpha \in (0, 1)$. Require

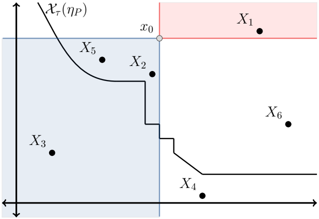
$$\inf_{P \in \mathcal{P}_{\mathrm{Mon},d}(\sigma)} \mathbb{P}_P\{\hat{A}(\mathcal{D}) \subseteq \mathcal{X}_\tau(\eta)\} \geq 1 - \alpha.$$

- Power: Want small expected regret

$$\mathbb{E}\{\mu(\mathcal{X}_\tau(\eta) \setminus \hat{A}(\mathcal{D}))\}.$$

High-level strategy

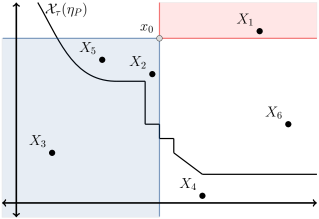
For $x_0 \in \mathbb{R}^d$, define null hypothesis $H_0(x_0) := \{P \in \mathcal{P}_{\mathrm{Mon},d}(\sigma) : \eta(x_0) < \tau\}$.



2. Construct p-values $\hat{p}_\ell$ for $H_0(X_\ell)$, i.e. $\mathbb{P}(\hat{p}_\ell(\mathcal{D}) \leq \alpha \mid (X_i)_{i=1}^n) \leq \alpha$ for all $P \in H_0(X_\ell)$, $\alpha \in (0,1)$, $\ell \in \{1,\dots,m\}$;
3. Apply a multiple testing procedure to reject $\mathcal{R}_\alpha \subseteq \{1,\dots,m\}$ with $\mathbb{P}_P(\mathcal{R}_\alpha \cap \{\ell \in \{1,\dots,m\} : P \in H_0(X_\ell)\} \neq \emptyset \mid (X_i)_{i=1}^m) \leq \alpha$;
4. Output $\hat{A} := \{x \in \mathbb{R}^d : X_\ell \preccurlyeq x \text{ for some } \ell \in \mathcal{R}_\alpha\}$.

High-level strategy

For $x_0 \in \mathbb{R}^d$, define null hypothesis $H_0(x_0) := \{P \in \mathcal{P}_{\mathrm{Mon},d}(\sigma) : \eta(x_0) < \tau\}$.

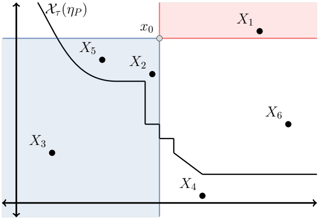


High-level strategy:

1. Sub-sample m covariate vectors $X_1, \ldots, X_m$ with $m \leq n$;

High-level strategy

For $x_0 \in \mathbb{R}^d$, define null hypothesis $H_0(x_0) := \{P \in \mathcal{P}_{\mathrm{Mon},d}(\sigma) : \eta(x_0) < \tau\}$.

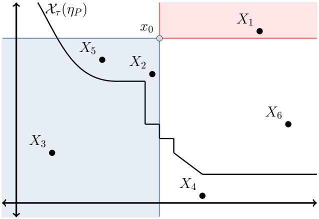


High-level strategy:

1. Sub-sample m covariate vectors $X_1, \ldots, X_m$ with $m \leq n$;
2. Construct p-values $\hat{p}_\ell$ for $H_0(X_\ell)$, i.e. $\mathbb{P}\big(\hat{p}_\ell(\mathcal{D}) \leq \alpha \mid (X_i)_{i=1}^n\big) \leq \alpha$ for all $P \in H_0(X_\ell)$, $\alpha \in (0,1)$, $\ell \in \{1, \ldots, m\}$;

High-level strategy

For $x_0 \in \mathbb{R}^d$, define null hypothesis $H_0(x_0) := \{P \in \mathcal{P}_{\mathrm{Mon},d}(\sigma) : \eta(x_0) < \tau\}$.

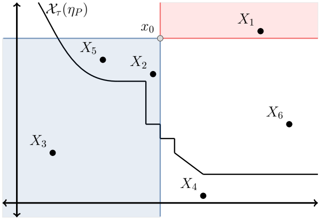


High-level strategy:

1. Sub-sample m covariate vectors $X_1, \ldots, X_m$ with $m \leq n$;
2. Construct p-values $\hat{p}_\ell$ for $H_0(X_\ell)$, i.e. $\mathbb{P}\big(\hat{p}_\ell(\mathcal{D}) \leq \alpha \mid (X_i)_{i=1}^n\big) \leq \alpha$ for all $P \in H_0(X_\ell)$, $\alpha \in (0,1)$, $\ell \in \{1, \ldots, m\}$;
3. Apply a multiple testing procedure to reject $\mathcal{R}_\alpha \subseteq \{1, \ldots, m\}$ with $\mathbb{P}_P\big(\mathcal{R}_\alpha \cap \{\ell \in \{1, \ldots, m\} : P \in H_0(X_\ell)\} \neq \emptyset \mid (X_i)_{i=1}^m\big) \leq \alpha$;

High-level strategy

For $x_0 \in \mathbb{R}^d$, define null hypothesis $H_0(x_0) := \{P \in \mathcal{P}_{\mathrm{Mon},d}(\sigma) : \eta(x_0) < \tau\}$.



High-level strategy:

1. Sub-sample m covariate vectors $X_1, \ldots, X_m$ with $m \leq n$;
2. Construct p-values $\hat{p}_\ell$ for $H_0(X_\ell)$, i.e. $\mathbb{P}\big(\hat{p}_\ell(\mathcal{D}) \leq \alpha \mid (X_i)_{i=1}^n\big) \leq \alpha$ for all $P \in H_0(X_\ell)$, $\alpha \in (0,1)$, $\ell \in \{1, \ldots, m\}$;
3. Apply a multiple testing procedure to reject $\mathcal{R}_\alpha \subseteq \{1, \ldots, m\}$ with $\mathbb{P}_P\big(\mathcal{R}_\alpha \cap \{\ell \in \{1, \ldots, m\} : P \in H_0(X_\ell)\} \neq \emptyset \mid (X_i)_{i=1}^m\big) \leq \alpha$;
4. Output $\hat{A} := \big\{x \in \mathbb{R}^d : X_\ell \preccurlyeq x \text{ for some } \ell \in \mathcal{R}_\alpha\big\}$.

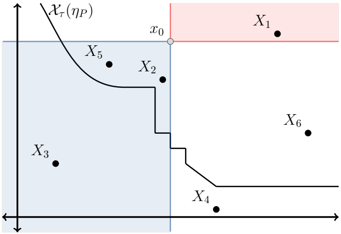
p-value construction

Given $x_0 \in \mathbb{R}^d$, we seek a p-value for
$H_0(x_0) := \{P \in \mathcal{P}_{\mathrm{Mon},d}(\sigma) : \eta(x_0) < \tau\}$.

p-value construction

Given $x_0 \in \mathbb{R}^d$, we seek a p-value for $H_0(x_0) := \{P \in \mathcal{P}_{\mathrm{Mon},d}(\sigma) : \eta(x_0) < \tau\}$.

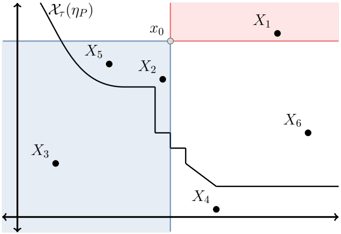
Write $\mathcal{I}(x_0) := \{i \in [n] : X_i \preccurlyeq x_0\}$. Let $X_{(j)}$ be the jth nearest neighbor in sup-norm of x_0 among X_i, $i \in \mathcal{I}(x_0)$, and let $Y_{(j)}$ be the corresponding response.



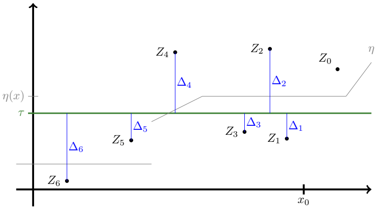
p-value construction

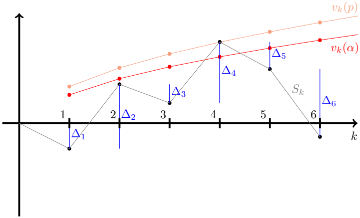
Given $x_0 \in \mathbb{R}^d$, we seek a p-value for
$H_0(x_0) := \{P \in \mathcal{P}_{\mathrm{Mon},d}(\sigma) : \eta(x_0) < \tau\}$.

Write $\mathcal{I}(x_0) := \{i \in [n] : X_i \preccurlyeq x_0\}$. Let $X_{(j)}$ be the jth nearest neighbor in sup-norm of x_0 among X_i, $i \in \mathcal{I}(x_0)$, and let $Y_{(j)}$ be the corresponding response.



Then, $S_k := \sum_{j=1}^{|\mathcal{I}(x_0)|} (Y_{(j)} - \tau)/\sigma$ is a supermartingale under $P \in H_0(x_0)$.
Combination with time-uniform bounds by Howard et al. (2021) gives p-values.





(Here: $Z_j := (X_{(j)}, Y_{(j)})$, $\Delta_j := (Y_{(j)} - \tau)/\sigma$.)

Theory: Type I error control

Combining the presented p-value construction with the just illustrated multiple testing procedure defines the proposed procedure $\hat{A}^{\mathrm{ISS}} \equiv \hat{A}^{\mathrm{ISS}}_{\sigma,\tau,\alpha,m}$.

Theorem. For any $n \in \mathbb{N}$, $m \in [n]$, $\alpha \in (0,1)$, $\sigma > 0$, and $P \in \mathcal{P}_{\mathrm{Mon},d}(\sigma)$, along with $\mathcal{D} \sim P^n$, we have

$$\mathbb{P}_P\big(\hat{A}^{\mathrm{ISS}}_{\sigma,\tau,\alpha,m}(\mathcal{D}) \subseteq \mathcal{X}_\tau(\eta) \,\big|\, \mathcal{D}_X\big) \geq 1 - \alpha.$$

Ideas: martingale test procedures (Duan et al., 2020), time-uniform boundaries on martingales with sub-Gaussian increments (Howard et al., 2021) and the sequential rejection principle (Goeman and Solari, 2010).

Theory: Power

Theorem. Let $\sigma, \gamma > 0, \theta > 1$ and $\lambda \in (0,1)$. There exists $C \geq 1$, depending only on (d, θ), such that for any $P \in \mathcal{P}_{\mathrm{Mon},d}(\sigma) \cap \mathcal{P}_{\mathrm{Reg},d}(\tau, \theta, \gamma, \lambda)$, $n \in \mathbb{N}$, $m \in [n]$, $\alpha \in (0,1)$ and $\mathcal{D} \sim P^n$ that[a]

$$\mathbb{E}_P\big\{\mu\big(\mathcal{X}_\tau(\eta) \setminus \hat{A}^{\mathrm{ISS}}(\mathcal{D})\big)\big\}$$
$$\leq 1 \wedge C\left\{\left(\frac{\sigma^2}{n\lambda^2}\log_+\left(\frac{m\log_+ n}{\alpha}\right)\right)^{1/(2\gamma+d)} + \left(\frac{\log_+ m}{m}\right)^{1/d}\right\}.$$

[a] $\log_+ x := \log(x \vee e)$.

Theory: Power

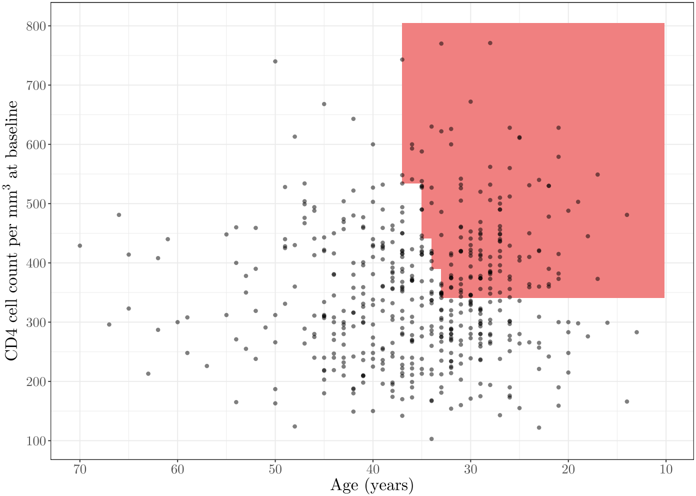
Theorem. Let $\sigma, \gamma > 0$, $\theta > 1$ and $\lambda \in (0,1)$. There exists $C \geq 1$, depending only on (d, θ), such that for any $P \in \mathcal{P}_{\mathrm{Mon},d}(\sigma) \cap \mathcal{P}_{\mathrm{Reg},d}(\tau, \theta, \gamma, \lambda)$, $n \in \mathbb{N}$, $m \in [n]$, $\alpha \in (0,1)$ and $\mathcal{D} \sim P^n$ that[a]

$$\mathbb{E}_P\big\{\mu\big(\mathcal{X}_\tau(\eta) \setminus \hat{A}^{\mathrm{ISS}}(\mathcal{D})\big)\big\} \leq 1 \wedge C\left\{\left(\frac{\sigma^2}{n\lambda^2}\log_+\left(\frac{m\log_+ n}{\alpha}\right)\right)^{1/(2\gamma+d)} + \left(\frac{\log_+ m}{m}\right)^{1/d}\right\}.$$

[a] $\log_+ x := \log(x \vee e)$.

Theorem. Our procedure is minimax-optimal up to poly-logarithmic factors among procedures with Type I error control over $\mathcal{P}_{\mathrm{Mon},d}(\sigma) \cap \mathcal{P}_{\mathrm{Reg},d}(\tau, \theta, \gamma, \lambda)$.

AIDS Clinical Trials Group Study 175 (Hammer et al., 1996)



Summary

- We addressed the problem of subgroup selection in isotonic regression.

Summary

- We addressed the problem of subgroup selection in isotonic regression.
- We propose a computationally-feasible algorithm controlling Type I error.

Summary

- We addressed the problem of subgroup selection in isotonic regression.
- We propose a computationally-feasible algorithm controlling Type I error.
- We show minimax-optimality up to poly-logarithmic factors of our method.

Summary

- We addressed the problem of subgroup selection in isotonic regression.
- We propose a computationally-feasible algorithm controlling Type I error.
- We show minimax-optimality up to poly-logarithmic factors of our method.

Extensions and further results (see full paper for details):

Summary

- We addressed the problem of subgroup selection in isotonic regression.
- We propose a computationally-feasible algorithm controlling Type I error.
- We show minimax-optimality up to poly-logarithmic factors of our method.

Extensions and further results (see full paper for details):

- Variation tailored to bounded responses and classification.

Summary

- We addressed the problem of subgroup selection in isotonic regression.
- We propose a computationally-feasible algorithm controlling Type I error.
- We show minimax-optimality up to poly-logarithmic factors of our method.

Extensions and further results (see full paper for details):

- Variation tailored to bounded responses and classification.
- Variation tailored to isotonic quantile regression.

Summary

- We addressed the problem of subgroup selection in isotonic regression.
- We propose a computationally-feasible algorithm controlling Type I error.
- We show minimax-optimality up to poly-logarithmic factors of our method.

Extensions and further results (see full paper for details):

- Variation tailored to bounded responses and classification.
- Variation tailored to isotonic quantile regression.
- Use in heterogeneous treatment effects in randomised controlled trials.

Summary

- We addressed the problem of subgroup selection in isotonic regression.
- We propose a computationally-feasible algorithm controlling Type I error.
- We show minimax-optimality up to poly-logarithmic factors of our method.

Extensions and further results (see full paper for details):

- Variation tailored to bounded responses and classification.
- Variation tailored to isotonic quantile regression.
- Use in heterogeneous treatment effects in randomised controlled trials.
- Further applications and simulations.

Summary

- We addressed the problem of subgroup selection in isotonic regression.
- We propose a computationally-feasible algorithm controlling Type I error.
- We show minimax-optimality up to poly-logarithmic factors of our method.

Extensions and further results (see full paper for details):

- Variation tailored to bounded responses and classification.
- Variation tailored to isotonic quantile regression.
- Use in heterogeneous treatment effects in randomised controlled trials.
- Further applications and simulations.
- ...

References

Duan, B., Ramdas, A., Balakrishnan, S., and Wasserman, L. (2020). Interactive martingale tests for the global null. *Electronic Journal of Statistics*, 14(2):4489–4551.

Goeman, J. J. and Solari, A. (2010). The sequential rejection principle of familywise error control. *The Annals of Statistics*, 38(6):3782–3810.

Hammer, S. M., Katzenstein, D. A., Hughes, M. D., Gundacker, H., Schooley, R. T., Haubrich, R. H., Henry, W. K., Lederman, M. M., Phair, J. P., Niu, M., et al. (1996). A trial comparing nucleoside monotherapy with combination therapy in HIV-infected adults with CD4 cell counts from 200 to 500 per cubic millimeter. *New England Journal of Medicine*, 335(15):1081–1090.

Howard, S. R., Ramdas, A., McAuliffe, J., and Sekhon, J. (2021). Time-uniform, nonparametric, nonasymptotic confidence sequences. *The Annals of Statistics*, 49:1055–1080.

Meijer, R. J. and Goeman, J. J. (2015). A multiple testing method for hypotheses structured in a directed acyclic graph. *Biometrical Journal*, 57(1):123–143.

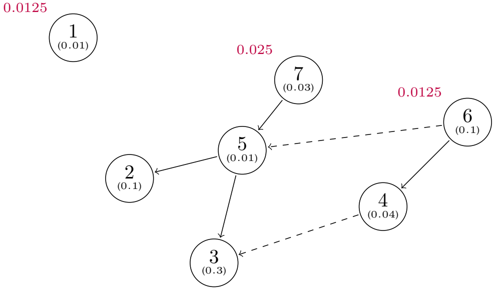
Multiple testing procedures for DAGs

Key idea: logical relationships of hypotheses $H_0(x_i), i \in [m]$, induce DAG with vertex set $[m]$. Careful α-budget allocation and sequential rejections are then performed.

Multiple testing procedures for DAGs

Key idea: logical relationships of hypotheses $H_0(x_i)$, $i \in [m]$, induce DAG with vertex set $[m]$. Careful α-budget allocation and sequential rejections are then performed.

Example: In the first iteration, no hypothesis has been rejected yet and only root nodes are assigned positive α-budget.

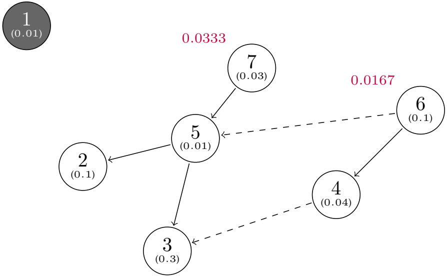


Here, nodes $1, 6$ and 7 are current rejection candidates, and 1 will be rejected, as $p_1 = 0.01 \leq 0.0125$.

Multiple testing procedures for DAGs

Key idea: logical relationships of hypotheses $H_0(x_i), i \in [m]$, induce DAG with vertex set $[m]$. Careful α-budget allocation and sequential rejections are then performed.

Example: After rejection of node 1 in the first step, we reallocate the α-budget.

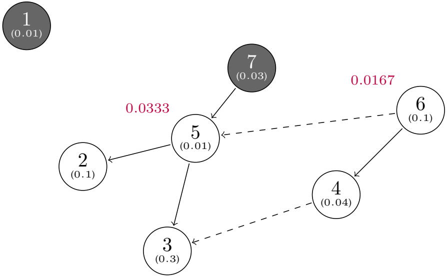


Here, node 7 will be rejected.

Multiple testing procedures for DAGs

Key idea: logical relationships of hypotheses $H_0(x_i), i \in [m]$, induce DAG with vertex set $[m]$. Careful α-budget allocation and sequential rejections are then performed.

Example: Now that node 7 has been rejected, its child 5 receives α-budget sufficiently large for it to be rejected.

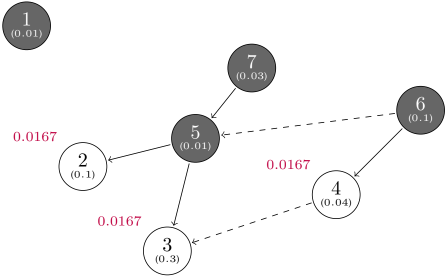


Although p_6 is quite large, 6 is an ancestor of 5 in the induced DAG and will hence also be rejected.

Multiple testing procedures for DAGs

Key idea: logical relationships of hypotheses $H_0(x_i), i \in [m]$, induce DAG with vertex set $[m]$. Careful α-budget allocation and sequential rejections are then performed.

Example: None of the remaining three nodes (coincidentally, the leaf nodes), have a p-value smaller than their respective α-budgets.



Nodes 1, 5, 6 and 7 have been rejected.

Theory: Power

Given $d \in \mathbb{N}$, $\tau \in \mathbb{R}$, $\theta > 1$, $\gamma > 0$ and $\lambda \in (0,1)$, we let $\mathcal{P}_{\mathrm{Reg},d}(\tau, \theta, \gamma, \lambda)$ denote the class of all distributions P on $\mathbb{R}^d \times \mathbb{R}$ with marginal μ on $\mathbb{R}^d$ and associated regression function η such that

(i) $\theta^{-1} \cdot r^d \leq \mu\big(B_\infty(x,r)\big) \leq \theta \cdot (2r)^d$ for $x \in \mathcal{X}_\tau(\eta) \cap \mathrm{supp}(\mu)$ and $r \in (0,1]$;

(ii) $B_\infty(x,r) \cap \mathcal{X}_{\tau+\lambda \cdot r^\gamma}(\eta) \neq \emptyset$ for $x \in \mathcal{X}_\tau(\eta) \cap \mathrm{supp}(\mu)$ and $r \in (0,1]$,

where $B_\infty(x,r)$ is the closed sup-norm ball around x of radius r.

Theory: Power

Given $d \in \mathbb{N}$, $\tau \in \mathbb{R}$, $\theta > 1$, $\gamma > 0$ and $\lambda \in (0,1)$, we let $\mathcal{P}_{\mathrm{Reg},d}(\tau,\theta,\gamma,\lambda)$ denote the class of all distributions P on $\mathbb{R}^d \times \mathbb{R}$ with marginal μ on $\mathbb{R}^d$ and associated regression function η such that

(i) $\theta^{-1} \cdot r^d \leq \mu\big(B_\infty(x,r)\big) \leq \theta \cdot (2r)^d$ for $x \in \mathcal{X}_\tau(\eta) \cap \mathrm{supp}(\mu)$ and $r \in (0,1]$;

(ii) $B_\infty(x,r) \cap \mathcal{X}_{\tau+\lambda\cdot r^\gamma}(\eta) \neq \emptyset$ for $x \in \mathcal{X}_\tau(\eta) \cap \mathrm{supp}(\mu)$ and $r \in (0,1]$,

where $B_\infty(x,r)$ is the closed sup-norm ball around x of radius r.

The first condition ensures that μ is genuinely d-dimensional.

Theory: Power

Given $d \in \mathbb{N}$, $\tau \in \mathbb{R}$, $\theta > 1$, $\gamma > 0$ and $\lambda \in (0, 1)$, we let $\mathcal{P}_{\mathrm{Reg},d}(\tau, \theta, \gamma, \lambda)$ denote the class of all distributions P on $\mathbb{R}^d \times \mathbb{R}$ with marginal μ on $\mathbb{R}^d$ and associated regression function η such that

(i) $\theta^{-1} \cdot r^d \leq \mu\big(B_\infty(x, r)\big) \leq \theta \cdot (2r)^d$ for $x \in \mathcal{X}_\tau(\eta) \cap \mathrm{supp}(\mu)$ and $r \in (0, 1]$;

(ii) $B_\infty(x, r) \cap \mathcal{X}_{\tau+\lambda \cdot r^\gamma}(\eta) \neq \emptyset$ for $x \in \mathcal{X}_\tau(\eta) \cap \mathrm{supp}(\mu)$ and $r \in (0, 1]$,

where $B_\infty(x, r)$ is the closed sup-norm ball around x of radius r.

The first condition ensures that μ is genuinely d-dimensional.

The second controls the way in which η grows around the τ-boundary.

Simulations: multiple testing procedure

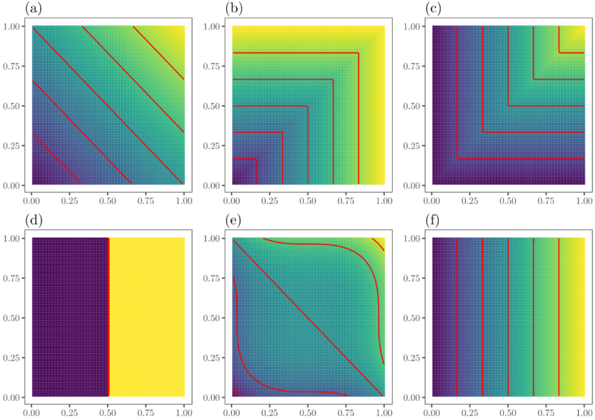
We conduct a simulation study to compare with other choices of multiple testing procedure.

Label	Function f	τ	$\gamma(P)$
(a)	$\sum_{j=1}^{d} x^{(j)}$	$1/2$	1
(b)	$\max_{1\leq j\leq d} x^{(j)}$	$1/2^{1/d}$	1
(c)	$\min_{1\leq j\leq d} x^{(j)}$	$1 - 1/2^{1/d}$	1
(d)	$\mathbb{1}_{(0.5,1]}\big(x^{(1)}\big)$	$1/2$	0
(e)	$\sum_{j=1}^{d}\big(x^{(j)} - 0.5\big)^3$	$1/2$	3
(f)	$x^{(1)}$	$1/2$	1

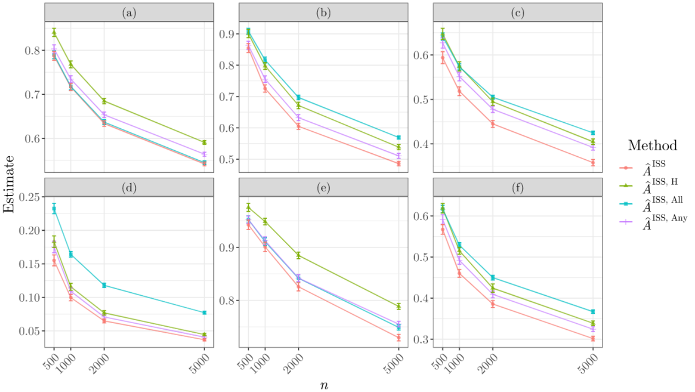
Our regression functions η are obtained by rescaling f.

Simulations: multiple testing procedure

We conduct a simulation study to compare with other choices of multiple testing procedure.



Simulations: multiple testing procedure



Presented is the MC-based estimate of $\mathbb{E}_P\{\mu(\mathcal{X}_\tau(\eta) \setminus \hat{A})\}$ for $\hat{A} \in \{\hat{A}^{\text{ISS}}, \hat{A}^{\text{ISS,H}}, \hat{A}^{\text{ISS,All}}, \hat{A}^{\text{ISS,Any}}\}$. The last two use methods due to Meijer and Goeman (2015).