Optimal classes of e-variables

Eugenio Clerico

Universitat Pompeu Fabra, Barcelona

BIRS-CMI workshop
Chennai, June-July 2025

Mainly based upon the arXiv preprints:
On the optimality of coin-betting for mean estimation
Optimal e-value testing for properly constrained hypotheses

Setting

- $\mathcal{Z} \subseteq \mathbb{R}^d$
- $\mathcal{P}_{\mathcal{Z}}$: Borel probability measures on $\mathcal{Z}$
- $\mathcal{H} \subseteq \mathcal{P}_{\mathcal{Z}}$ a hypothesis
- $\mathcal{E}_{\mathcal{H}} = \big\{E \geq 0 \text{ Borel } : \mathbb{E}_P[E] \leq 1\,, \forall P \in \mathcal{H}\big\}$

$\mathcal{E}_{\mathcal{H}}$ is the set of all the e-variables for $\mathcal{H}$.

What are the *best* e-variables?

Several notions of *optimality* in the literature...

- ⋆ Growth Rate Optimality
  Grünwald, de Heide, Koolen, *Safe testing* (2019)
- ⋆ Numéraire
  Larsson, Ramdas, Ruf, *The numeraire e-variable and reverse information projection* (2024)
- ⋆ Admissibility
  Ramdas, Ruf, Larsson, Koolen, *Admissible anytime-valid sequential inference must rely on nonnegative martingales* (2020)

We'll consider a slightly *weaker* notion, very close to admissibility.

Motivation (?)

P a probability distribution on $[-1,1]$, $(x_t)_{t\geq 1}$ independent draws from P.
Goal: test if $\mathbb{E}_P[X] = 0$.

***Coin-better*'s test**

At each round t:

- Pick $\lambda_t \in [-1,1]$ using *past* information
- Evaluate $W_t = \prod_{i=1}^{t}(1+\lambda_t x_t)$
- Reject and stop if $W_t > 1/\delta$

$x \mapsto 1 + \lambda x$ is an *e-variable* and this is a test via (single-round) e-variables!

Are we missing something restricting to coin-betting e-variables?

Majorising classes

Poset structure:

- ⋆ $E \geq E'$ if $E(z) \geq E'(z)$ for all $z \in \mathcal{Z}$
- ⋆ E is *maximal* if there is no $E' \in \mathcal{E}_{\mathcal{H}}$ such that $E' > E$

$\mathcal{E} \subseteq \mathcal{E}_{\mathcal{H}}$ is a *majorising* class if $\forall E \in \mathcal{E}_{\mathcal{H}}$ there is $E' \in \mathcal{E}$ such that $E' \geq E$

Picking e-variables outside of a majorising class is *pointless*!

Is there a *smallest* majorising class?

A majorising class contained in any other majorising class is *optimal*

Optimal classes do not always exist!

Lemma

An optimal class exists iff $\mathcal{E}_{\max} = \{E \in \mathcal{E}_{\mathcal{H}} : E \text{ is maximal}\}$ is majorising. In such case, the optimal class is unique and coincides with $\mathcal{E}_{\max}$.

Checking if $\mathcal{E}_{\text{opt}}$ exists comes down to check if every e-variable is dominated by a maximal e-variable!

When does the optimal class exist?

Characterising the hypotheses for which the optimal class exists is *open*

Some partial results...

- ⋆ $\mathcal{E}_{\mathrm{opt}}$ exists if $\mathcal{Z}$ is countable
- ⋆ If $\mathcal{Z}$ is uncountable $\mathcal{E}_{\mathrm{opt}}$ might not exist:
  e.g.: $\mathcal{Z} = [0, 1]$, $\mathcal{H} = \{P : P(\{0\}) \geq 1/2\} \cup \{\mathrm{Unif}_{[0,1]}\}$, $\nexists \mathcal{E}_{\mathrm{opt}}$
- ⋆ $\mathcal{E}_{\mathrm{opt}}$ exists when $\mathcal{H}$ is defined through finitely many linear constraints
  e.g.: $\mathcal{Z} = [-1, 1]$, $\mathcal{H} = \{P : \mathbb{E}_P[Z] = 0\}$

Testing for the mean

$\mathcal{Z} \subseteq \mathbb{R}^d$ compact
$0 \in \operatorname{int} \operatorname{conv} \mathcal{Z}$
$\mathcal{H} = \{P : \mathbb{E}_P[Z] = 0\}$

Proposition

$\mathcal{E}_{\mathrm{opt}}$ exists and is given by

$$\mathcal{E}_{\mathrm{opt}} = \{E_\lambda \geq 0\}\,, \quad E_\lambda(z) = 1 + \lambda \cdot z$$

Proof (maximality)

Each e-variable in the form $E_\lambda : x \mapsto 1 + \lambda \cdot x$ is maximal:

- ⋆ Fix an e-variable $E \geq E_\lambda$ and an arbitrary x
- ⋆ There is $P_x \in \mathcal{H}$ such that $P_x(\{x\}) > 0$
- ⋆ $\mathbb{E}_{P_x}[E_\lambda] = 1$
- ⋆ $0 \leq P_x(\{x\})(E(x) - E_\lambda(x)) \leq \mathbb{E}_{P_x}[E - E_\lambda] = \mathbb{E}_{P_x}[E] - 1 \leq 0$
- ⋆ So, $E = E_\lambda$!

Proof (maximality)

Each e-variable in the form $E_\lambda : x \mapsto 1 + \lambda \cdot x$ is maximal:

- ⋆ Fix an e-variable $E \geq E_\lambda$ and an arbitrary x
- ⋆ There is $P_x \in \mathcal{H}$ such that $P_x(\{x\}) > 0$
- ⋆ $\mathbb{E}_{P_x}[E_\lambda] = 1$
- ⋆ $0 \leq P_x(\{x\})(E(x) - E_\lambda(x)) \leq \mathbb{E}_{P_x}[E - E_\lambda] = \mathbb{E}_{P_x}[E] - 1 \leq 0$
- ⋆ So, $E = E_\lambda$!

Proof (maximality)

Each e-variable in the form $E_\lambda : x \mapsto 1 + \lambda \cdot x$ is maximal:

- ⋆ Fix an e-variable $E \geq E_\lambda$ and an arbitrary x
- ⋆ There is $P_x \in \mathcal{H}$ such that $P_x(\{x\}) > 0$
- ⋆ $\mathbb{E}_{P_x}[E_\lambda] = 1$
- ⋆ $0 \leq P_x(\{x\})(E(x) - E_\lambda(x)) \leq \mathbb{E}_{P_x}[E - E_\lambda] = \mathbb{E}_{P_x}[E] - 1 \leq 0$
- ⋆ So, $E = E_\lambda$!

Proof (maximality)

Each e-variable in the form $E_\lambda : x \mapsto 1 + \lambda \cdot x$ is maximal:

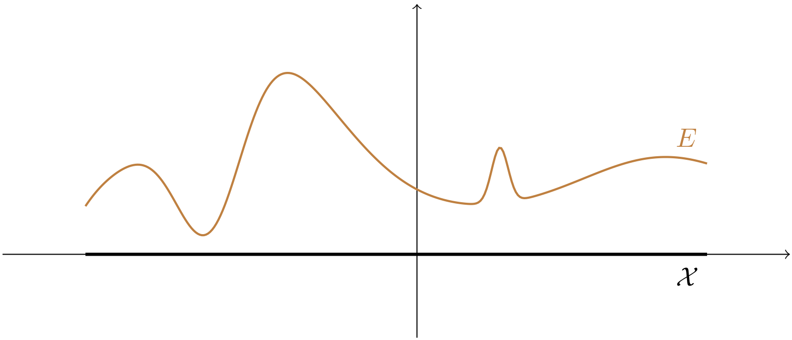
- ⋆ Fix an e-variable $E \geq E_\lambda$ and an arbitrary x
- ⋆ There is $P_x \in \mathcal{H}$ such that $P_x(\{x\}) > 0$
- ⋆ $\mathbb{E}_{P_x}[E_\lambda] = 1$
- ⋆ $0 \leq P_x(\{x\})(E(x) - E_\lambda(x)) \leq \mathbb{E}_{P_x}[E - E_\lambda] = \mathbb{E}_{P_x}[E] - 1 \leq 0$
- ⋆ So, $E = E_\lambda$!

Proof (maximality)

Each e-variable in the form $E_\lambda : x \mapsto 1 + \lambda \cdot x$ is maximal:

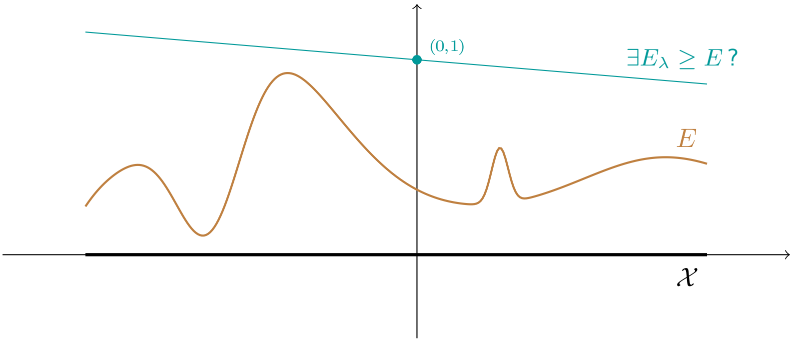
- ⋆ Fix an e-variable $E \geq E_\lambda$ and an arbitrary x
- ⋆ There is $P_x \in \mathcal{H}$ such that $P_x(\{x\}) > 0$
- ⋆ $\mathbb{E}_{P_x}[E_\lambda] = 1$
- ⋆ $0 \leq P_x(\{x\})(E(x) - E_\lambda(x)) \leq \mathbb{E}_{P_x}[E - E_\lambda] = \mathbb{E}_{P_x}[E] - 1 \leq 0$
- ⋆ So, $E = E_\lambda$!

Proof (majorising property)



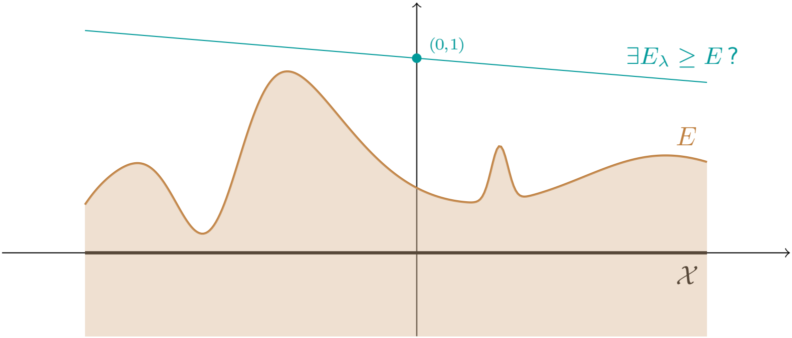
Fix an e-variable E.

Proof (majorising property)



Fix an e-variable E. We want to show that E is dominated by some E_λ.

Proof (majorising property)

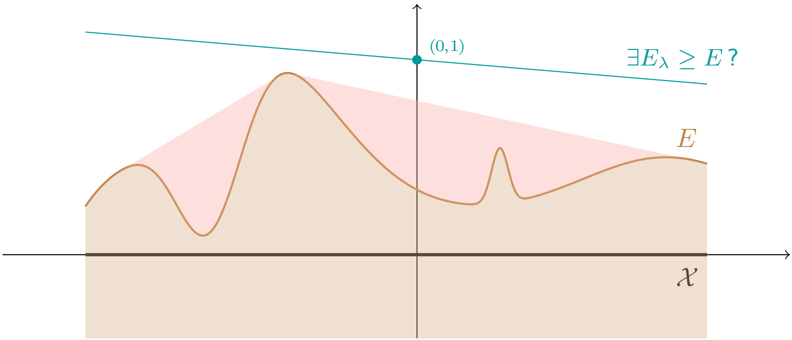


Fix an e-variable E. We want to show that E is dominated by some E_λ.

Strategy

We show that the *hypograph* of E lies entirely in a half-space bounded by a hyperplane passing through $(0, 1)$.

Proof (majorising property)



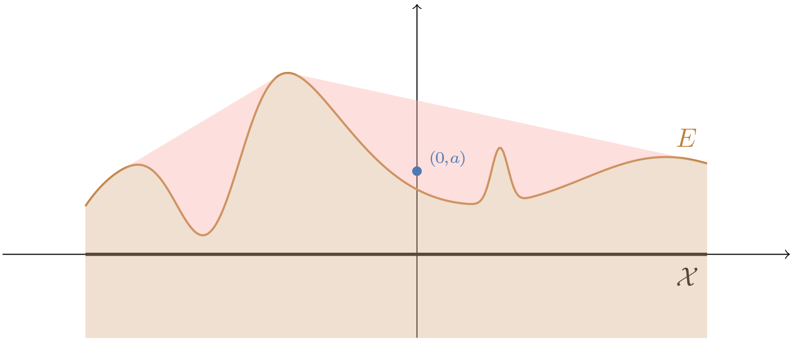
Fix an e-variable E. We want to show that E is dominated by some E_λ.

Strategy

We show that the *hypograph* of E lies entirely in a half-space bounded by a hyperplane passing through $(0, 1)$.

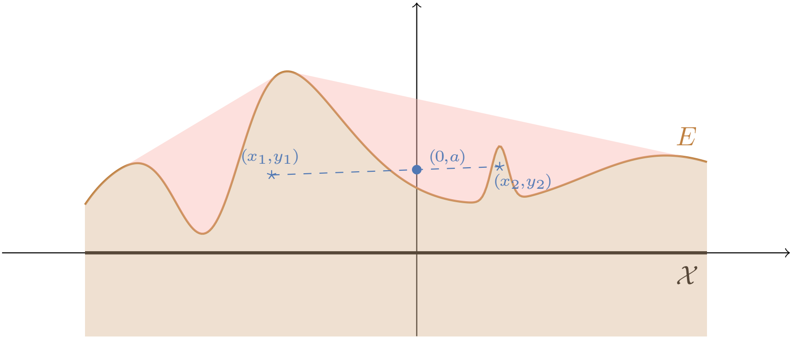
We show that $(0, 1)$ is above the convex hull of the hypograph of E.

Proof (majorising property)



Let $(0, a)$ be in the convex hull of the hypograph.

Proof (majorising property)

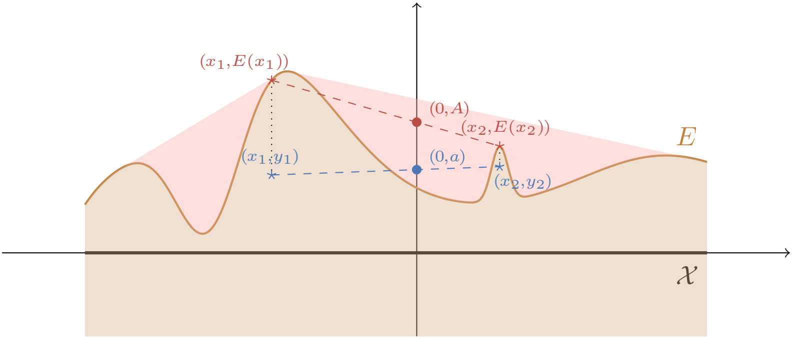


Let $(0, a)$ be in the convex hull of the hypograph.

There are $x_1, \ldots, x_N$, $y_1, \ldots, y_N$, $\alpha_1, \ldots, \alpha_N$ such that:

1) $\sum_i \alpha_i = 1$ and $\alpha_i \geq 0$;
2) $\sum_i \alpha_i x_i = 0$;
3) $\sum_i \alpha_i y_i = a$;
4) each (x_i, y_i) is in the hypograph.

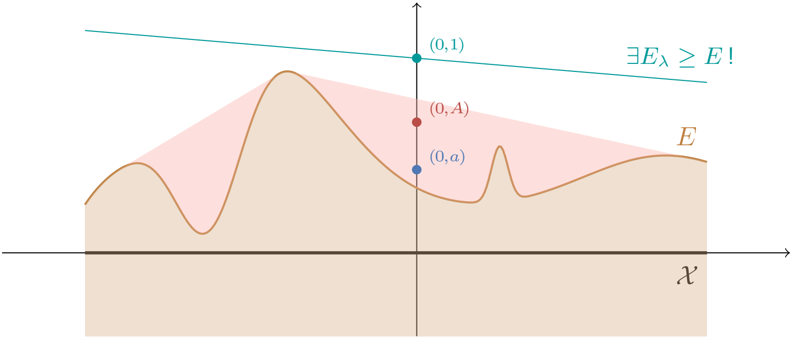
Proof (majorising property)



By construction $y_i \leq E(x_i)$, and $P = \sum_i \alpha_i \delta_{x_i} \in \mathcal{H}$.

$$a \leq A = \sum_i \alpha_i E(x_i) = \mathbb{E}_P[E] \leq 1\,.$$

Proof (majorising property)



By construction $y_i \leq E(x_i)$, and $P = \sum_i \alpha_i \delta_{x_i} \in \mathcal{H}$.

$$a \leq A = \sum_i \alpha_i E(x_i) = \mathbb{E}_P[E] \leq 1\,.$$

So, $(0,1)$ lies on the boundary or above the convex hull!

Testing for the conditional mean

$\mathcal{Z} = [-1, 1]^T$
$\mathcal{H} = \{P : \mathbb{E}_P[Z_t \,|\, Z_1, \dots, Z_{t-1}] = 0 \,, \forall t = 1 \dots T\}$

Proposition

$\mathcal{E}_{\mathrm{opt}}$ exists and is given by

$$\mathcal{E}_{\mathrm{opt}} = \{E_{\lambda_1,\dots,\lambda_T} \geq 0\}\,, \quad E_{\lambda_1,\dots,\lambda_T}(z_1, \dots z_T) = \prod_{t=1}^{T} \big(1 + \lambda_t(z_1 \dots z_{t-1})\, z_t\big)$$

$\lambda_1 \in [-1, 1]$, $\lambda_t : [-1, 1]^{t-1} \to [-1, 1]$ Borel

Proof's sketch ($T = 2$)

$\mathcal{H}^1 = \{Q \text{ on } [-1,1] : \mathbb{E}_Q[X] = 0\}$
$\mathcal{H}^2 = \{P \text{ on } [-1,1]^2 : \mathbb{E}_P[X] = 0\,,\, \mathbb{E}_P[Y|X] = 0\}$

Goal: the optimal e-variables for $\mathcal{H}_2$ are $(1+\lambda_1 x)(1+\lambda_2(x)y)$

Majorising property: Fix an e-variable E for $\mathcal{H}^2$.

Step 1: Show that $\forall E \in \mathcal{E}_{\mathcal{H}^2}\,,\, \exists \lambda_1 \in [-1,1] : \forall x \in [-1,1], \forall Q \in \mathcal{H}_1$

$$\mathbb{E}_Q[E(x,Y)] \leq 1 + \lambda_1 x$$

Idea: Same argument as before, applied to convex hull Γ of the union (for all Q) of the hypographs of $\mathbb{E}_Q[E(x,Y)]$. For $(0,a)$ in Γ we find $y_1,\dots,y_n$, $\alpha_1,\dots,\alpha_n$, and $Q_1,\dots,Q_n$ such that $\sum_i \alpha_i y_i = 0$ and $a = \sum_i \alpha_i \mathbb{E}_{Q_i}[E(x,y_i)]$. Since $P = \sum_i \alpha_i \delta_{y_i} \otimes Q_i \in \mathcal{H}$, then $a \leq 1$.

Proof's sketch $(T = 2)$

$\mathcal{H}^1 = \{Q \text{ on } [-1,1] : \mathbb{E}_Q[X] = 0\}$
$\mathcal{H}^2 = \{P \text{ on } [-1,1]^2 : \mathbb{E}_P[X] = 0\,,\, \mathbb{E}_P[Y|X] = 0\}$

Goal: the optimal e-variables for $\mathcal{H}_2$ are $(1 + \lambda_1 x)(1 + \lambda_2(x)y)$

Majorising property: Fix an e-variable E for $\mathcal{H}^2$.

Step 1: $\forall E \in \mathcal{E}_{\mathcal{H}^2}\,,\, \exists \lambda_1 \in [-1,1] : \forall x \in [-1,1], \forall Q \in \mathcal{H}_1\,,\, \mathbb{E}_Q[E(x,Y)] \leq 1 + \lambda_1 x$

Step 2: There are λ_1, $\tilde{\lambda}_2$ such that $E(x_1,x_2) \leq (1 + \lambda_1 x)(1 + \tilde{\lambda}_2(x)y)$
Idea: Fix x. From step 1, for all $Q \in \mathcal{H}_1$, $\mathbb{E}_Q[E(x,Y)]/(1 + \lambda_1 x) \leq 1$, so $y \mapsto E(x,y)/(1 + \lambda_1 x)$ is an e-variable for $\mathcal{H}^1$. Step 2 follows from what we have proved for $T = 1$.

Proof's sketch ($T = 2$)

$\mathcal{H}^1 = \{Q \text{ on } [-1,1] : \mathbb{E}_Q[X] = 0\}$
$\mathcal{H}^2 = \{P \text{ on } [-1,1]^2 : \mathbb{E}_P[X] = 0\,, \mathbb{E}_P[Y|X] = 0\}$

Goal: the optimal e-variables for $\mathcal{H}_2$ are $(1 + \lambda_1 x)(1 + \lambda_2(x)y)$

Majorising property: Fix an e-variable E for $\mathcal{H}^2$.

Step 1: $\forall E \in \mathcal{E}_{\mathcal{H}^2}\,, \exists \lambda_1 \in [-1,1] : \forall x \in [-1,1], \forall Q \in \mathcal{H}_1\,, \mathbb{E}_Q[E(x,Y)] \leq 1 + \lambda_1 x$

Step 2: There are λ_1, $\tilde{\lambda}_2$ such that $E(x_1, x_2) \leq (1 + \lambda_1 x)(1 + \tilde{\lambda}_2(x)y)$
Idea: Fix x. From step 1, for all $Q \in \mathcal{H}_1$, $\mathbb{E}_Q[E(x,Y)]/(1 + \lambda_1 x) \leq 1$, so $y \mapsto E(x,y)/(1 + \lambda_1 x)$ is an e-variable for $\mathcal{H}^1$. Step 2 follows from what we have proved for $T = 1$.
Issue: $x \mapsto \tilde{\lambda}_2(x)$ might be non-measurable!

Proof's sketch ($T = 2$)

$\mathcal{H}^1 = \{Q$ on $[-1,1] : \mathbb{E}_Q[X] = 0\}$
$\mathcal{H}^2 = \{P$ on $[-1,1]^2 : \mathbb{E}_P[X] = 0\,, \mathbb{E}_P[Y|X] = 0\}$

Goal: the optimal e-variables for $\mathcal{H}_2$ are $(1+\lambda_1 x)(1+\lambda_2(x)y)$

Majorising property: Fix an e-variable E for $\mathcal{H}^2$.

Step 1: $\exists \lambda_1 \in [-1,1] : \forall x \in [-1,1], \forall Q \in \mathcal{H}_1\,, \mathbb{E}_Q[E(x,Y)] \leq 1 + \lambda_1 x$

Step 2: $\exists \lambda_1, \, \tilde{\lambda}_2 \, : \, E(x_1, x_2) \leq (1+\lambda_1 x)(1+\tilde{\lambda}_2(x)y)$

Step 3: There is λ_2 *Borel* such that $E \leq (1+\lambda_1 x)(1+\lambda_2(x)y)$
Idea: Use a functional corollary of Lusin separation theorem. From step 2

$$\sup_{y>0} \frac{1}{y}\left(\frac{E(x,y)}{1+\lambda_1 x} - 1\right) \leq \inf_{y<0} \frac{1}{y}\left(\frac{E(x,y)}{1+\lambda_1 x} - 1\right).$$

The LHS is upper-semianalytic and the RHS is lower-semianalytic, so they must be separated by a Borel function.

Proof's sketch ($T = 2$)

$\mathcal{H}^1 = \{Q \text{ on } [-1,1] : \mathbb{E}_Q[X] = 0\}$
$\mathcal{H}^2 = \{P \text{ on } [-1,1]^2 : \mathbb{E}_P[X] = 0\,,\, \mathbb{E}_P[Y|X] = 0\}$

Goal: the optimal e-variables for $\mathcal{H}_2$ are $(1 + \lambda_1 x)(1 + \lambda_2(x)y)$

Majorising property: Fix an e-variable E for $\mathcal{H}^2$.

Step 1: $\exists \lambda_1 \in [-1,1] : \forall x \in [-1,1], \forall Q \in \mathcal{H}_1\,,\, \mathbb{E}_Q[E(x,Y)] \leq 1 + \lambda_1 x$

Step 2: $\exists \lambda_1,\, \tilde{\lambda}_2 : E(x_1, x_2) \leq (1 + \lambda_1 x)(1 + \tilde{\lambda}_2(x)y)$

Step 3: $\exists \lambda_2$ Borel $: E \leq (1 + \lambda_1 x)(1 + \lambda_2(x)y)$

So the class considered is majorising!

Maximality: Same proof we did for $T = 1$ works here as well.

A few questions (from me) I can't yet fully answer...

- Conditional mean for $\mathcal{Z} = \mathcal{X}^T$, $\mathcal{X} \subseteq \mathbb{R}^d$?
  Everything should work similarly but there's a measurability issue in adapting the same proof...
- $\mathcal{H}$ hypothesis on $\mathcal{Z}$ with $\mathcal{E}_{\mathrm{opt}}$, can we characterise the optimal class for $\mathcal{H}^T = \{P \text{ on } \mathcal{Z}^T \text{ such that } P_t\big|_{Z^{t-1}} \in \mathcal{H}\,, \forall t \leq T\}$
  Do we have always product of maximal e-variables in $\mathcal{E}_{\mathrm{opt}}$?
- What can we say of optimal class of e-processes when testing the conditional mean?
  If domain is finite the *maximal* e-processes are $E_t = E_{\lambda^t}$ for a sequence $(\lambda_t)_{t\geq 1}$, but what if the domain is infinite?
- What can we say for multi-round e-variables when testing mean and i.i.d. assumption?
  If the domain has more than 2 elements, there are more e-variables than in the conditional case...

A few questions (from me) I can't yet fully answer...

- ⋆ Conditional mean for $\mathcal{Z} = \mathcal{X}^T$, $\mathcal{X} \subseteq \mathbb{R}^d$?
  Everything should work similarly but there's a measurability issue in adapting the same proof...
- ⋆ $\mathcal{H}$ hypothesis on $\mathcal{Z}$ with $\mathcal{E}_{\mathrm{opt}}$, can we characterise the optimal class for $\mathcal{H}^T = \{P \text{ on } \mathcal{Z}^T \text{ such that } P_t\big|_{Z^{t-1}} \in \mathcal{H}\,,\ \forall t \leq T\}$
  Do we have always product of maximal e-variables in $\mathcal{E}_{\mathrm{opt}}$?
- ⋆ What can we say of optimal class of e-processes when testing the conditional mean?
  If domain is finite the *maximal* e-processes are $E_t = E_{\lambda^t}$ for a sequence $(\lambda_t)_{t\geq 1}$, but what if the domain is infinite?
- ⋆ What can we say for multi-round e-variables when testing mean and i.i.d. assumption?
  If the domain has more than 2 elements, there are more e-variables than in the conditional case...

A few questions (from me) I can't yet fully answer...

- ⋆ Conditional mean for $\mathcal{Z} = \mathcal{X}^T$, $\mathcal{X} \subseteq \mathbb{R}^d$?
  Everything should work similarly but there's a measurability issue in adapting the same proof...
- ⋆ $\mathcal{H}$ hypothesis on $\mathcal{Z}$ with $\mathcal{E}_{\mathrm{opt}}$, can we characterise the optimal class for $\mathcal{H}^T = \{P \text{ on } \mathcal{Z}^T \text{ such that } P_t\big|_{Z^{t-1}} \in \mathcal{H}\,,\ \forall t \leq T\}$
  Do we have always product of maximal e-variables in $\mathcal{E}_{\mathrm{opt}}$?
- ⋆ What can we say of optimal class of e-processes when testing the conditional mean?
  If domain is finite the *maximal* e-processes are $E_t = E_{\lambda^t}$ for a sequence $(\lambda_t)_{t\geq 1}$, but what if the domain is infinite?
- ⋆ What can we say for multi-round e-variables when testing mean and i.i.d. assumption?
  If the domain has more than 2 elements, there are more e-variables than in the conditional case...

A few questions (from me) I can't yet fully answer...

- Conditional mean for $\mathcal{Z} = \mathcal{X}^T$, $\mathcal{X} \subseteq \mathbb{R}^d$?
  Everything should work similarly but there's a measurability issue in adapting the same proof...
- $\mathcal{H}$ hypothesis on $\mathcal{Z}$ with $\mathcal{E}_{\text{opt}}$, can we characterise the optimal class for $\mathcal{H}^T = \{P \text{ on } \mathcal{Z}^T \text{ such that } P_t\big|_{Z^{t-1}} \in \mathcal{H}\,, \forall t \leq T\}$
  Do we have always product of maximal e-variables in $\mathcal{E}_{\text{opt}}$?
- What can we say of optimal class of e-processes when testing the conditional mean?
  If domain is finite the *maximal* e-processes are $E_t = E_{\lambda^t}$ for a sequence $(\lambda_t)_{t\geq 1}$, but what if the domain is infinite?
- What can we say for multi-round e-variables when testing mean and i.i.d. assumption?
  If the domain has more than 2 elements, there are more e-variables than in the conditional case...

Thank you!

More questions? (from you...)