The suprema of selector processes

Rafał Meller (based on joint work with W. Bednorz and R. Martynek)

Polish Academy of Science

Bedlewo June 2023

Motivation

Problem

Let $(X_t)_{t \in T}$ *be a stochastic process.* $\mathbb{E} \sup_{t \in T} X_t \approx ?$

Motivation

Problem

Let $(X_t)_{t\in T}$ *be a stochastic process.* $\mathbb{E}\sup_{t\in T} X_t \approx ?$

Theorem (Talagrand 1987)

Let $(G_t)_{t\in T}$ *be a Gaussian process. Then* $\mathbb{E}\sup_{t\in T} G_t \approx \gamma(T)$ *and* $\gamma(T)$ *depends only on the geometry of the index set* T.

Motivation

Problem

Let $(X_t)_{t\in T}$ *be a stochastic process.* $\mathbb{E}\sup_{t\in T} X_t \approx ?$

Theorem (Talagrand 1987)

Let $(G_t)_{t\in T}$ *be a Gaussian process. Then* $\mathbb{E}\sup_{t\in T} G_t \approx \gamma(T)$ *and* $\gamma(T)$ *depends only on the geometry of the index set* T.

Some further generalization

- Variables with c.d.f $c_r \exp(-|x|^r)$, $1 \le r < \infty$, Talagrand 1994;
  Variables with LCT +technical condition on growth, Latała 1997;
  Variables which moments are not "too big" and not "too small", Latała, Tkocz 2015

Motivation

Problem

Let $(X_t)_{t\in T}$ *be a stochastic process.* $\mathbb{E}\sup_{t\in T} X_t \approx ?$

Theorem (Talagrand 1987)

Let $(G_t)_{t\in T}$ *be a Gaussian process. Then* $\mathbb{E}\sup_{t\in T} G_t \approx \gamma(T)$ *and* $\gamma(T)$ *depends only on the geometry of the index set* T.

Some further generalization

- Variables with c.d.f $c_r \exp(-|x|^r)$, $1 \le r < \infty$, Talagrand 1994; Variables with LCT +technical condition on growth, Latała 1997; Variables which moments are not "too big" and not "too small", Latała, Tkocz 2015
- Bernoulli r.v's, Bednorz, Latała 2014

Motivation

Problem

Let $(X_t)_{t\in T}$ *be a stochastic process.* $\mathbb{E}\sup_{t\in T} X_t \approx ?$

Theorem (Talagrand 1987)

Let $(G_t)_{t\in T}$ *be a Gaussian process. Then* $\mathbb{E}\sup_{t\in T} G_t \approx \gamma(T)$ *and* $\gamma(T)$ *depends only on the geometry of the index set* T.

Some further generalization

- Variables with c.d.f $c_r \exp(-|x|^r)$, $1 \leq r < \infty$, Talagrand 1994; Variables with LCT +technical condition on growth, Latała 1997; Variables which moments are not "too big" and not "too small", Latała, Tkocz 2015
- Bernoulli r.v's, Bednorz, Latała 2014

Conclusion

It's hard to describe exact size of $\mathbb{E}\sup_{t\in T} X_t$ *when the process is "small".*

Motivation

Problem (Very hard)

Let δ_i *be i.i.d,* $\mathbb{P}(\delta_i = 1) = p = 1 - \mathbb{P}(\delta_i = 0)$ *and* $T \subset \mathbb{R}^n$*.*

$$\mathbb{E} \sup_{t \in T} \sum_i t_i \delta_i \approx ?$$

Motivation

Problem (Very hard)

Let δ_i *be i.i.d,* $\mathbb{P}(\delta_i = 1) = p = 1 - \mathbb{P}(\delta_i = 0)$ *and* $T \subset \mathbb{R}^n$*.*

$$\mathbb{E} \sup_{t \in T} \sum_i t_i \delta_i \approx ?$$

Fact (Talagrand)

$g_1, \ldots$ *i.i.d* $\mathcal{N}(0,1)$*. Then*

$$\{\sup_{t \in T} \sum_i g_i t_i \geq L \mathbb{E} \sup_{t \in T} \sum_i g_i t_i\} \subset \bigcup_{k \geq 1} \{\sum_i g_i t_i^k \geq u_k\},$$ *where*

$$\sum_k \mathbb{P}(\sum_i g_i t_i^k \geq u_k) \leq 1/2.$$

Conjecture (Talagrand $\sim$ 2010)

Is generalization of the above statement true for selector process?

Formulation of the Problem

Let δ_i be i.i.d, $\mathbb{P}(\delta_i = 1) = p = 1 - \mathbb{P}(\delta_i = 0)$, $\delta(T) := \mathbb{E} \sup_{t \in T} \sum_{i=1}^{n} t_i \delta_i$.

Theorem (Park, Pham 2022)

There exists constant L such that for any $T \subset (\mathbb{R}_+)^n$, there exists family $\mathcal{G}$ of subsets of $[n]$ such that

$$\{\sup_{t \in T} \sum_{i=1}^{n} t_i \delta_i \geq L\delta(T)\} \subset \bigcup_{S \in \mathcal{G}} \{\delta_i = 1 \ \mathit{for} \ i \in S\} \tag{1}$$

$$\sum_{S \in \mathcal{G}} \mathbb{P}(\delta_i = 1 \ \mathit{for} \ i \in S) = \sum_{S \in \mathcal{G}} p^{|S|} \leq 1/2. \tag{2}$$

Formulation of the Problem

Let δ_i be i.i.d, $\mathbb{P}(\delta_i = 1) = p = 1 - \mathbb{P}(\delta_i = 0)$, $\delta(T) := \mathbb{E} \sup_{t \in T} \sum_{i=1}^{n} t_i \delta_i$.

Theorem (Park, Pham 2022)

There exists constant L *such that for any* $T \subset (\mathbb{R}_+)^n$, *there exists family* $\mathcal{G}$ *of subsets of* $[n]$ *such that*

$$\{\sup_{t \in T} \sum_{i=1}^{n} t_i \delta_i \geq L\delta(T)\} \subset \bigcup_{S \in \mathcal{G}} \{\delta_i = 1 \; for \; i \in S\} \tag{1}$$

$$\sum_{S \in \mathcal{G}} \mathbb{P}(\delta_i = 1 \; for \; i \in S) = \sum_{S \in \mathcal{G}} p^{|S|} \leq 1/2. \tag{2}$$

Remark

The Theorem is obvious for $p > c > 0$ *since*

$$\sup_{t \in T} \sum_{i} t_i \delta_i \leq \sup_{t \in T} t_i = p^{-1} \sup_{t \in T} \mathbb{E} \sum_{i=1}^{n}$$

Small families

A collection $\mathcal{F}$ of subsets of $[n] := \{1, \ldots, n\}$ is p-small if there exists a collection $\mathcal{G}$ of $[n]$ such that

$$\mathcal{F} \subset \langle \mathcal{G} \rangle := \bigcup_{S \in \mathcal{G}} \langle S \rangle = \bigcup_{S \in \mathcal{G}} \{I \subset \{1, \ldots, n\} : S \subset I\}, \ \sum_{S \in G} p^{|S|} \leq 1/2,$$

$\langle S \rangle =$ all subsets of $[n]$ that contains S. $\mathbb{P}(\langle S \rangle) = p^{|S|}$.

Small families

A collection $\mathcal{F}$ of subsets of $[n] := \{1, \ldots, n\}$ is p-small if there exists a collection $\mathcal{G}$ of $[n]$ such that

$$\mathcal{F} \subset \langle \mathcal{G} \rangle := \bigcup_{S \in \mathcal{G}} \langle S \rangle = \bigcup_{S \in \mathcal{G}} \{I \subset \{1, \ldots, n\} : S \subset I\}, \ \sum_{S \in G} p^{|S|} \leq 1/2,$$

$\langle S \rangle =$ all subsets of $[n]$ that contains S. $\mathbb{P}(\langle S \rangle) = p^{|S|}$. **Examples:**

- $2^{[n]} = \bigcup_{k=1}^{n} \langle k \rangle \cup \emptyset$. If $np \leq 1/2$ than any familly is small.

Small families

A collection $\mathcal{F}$ of subsets of $[n] := \{1, \ldots, n\}$ is *p*-small if there exists a collection $\mathcal{G}$ of $[n]$ such that

$$\mathcal{F} \subset \langle \mathcal{G} \rangle := \bigcup_{S \in \mathcal{G}} \langle S \rangle = \bigcup_{S \in \mathcal{G}} \{I \subset \{1, \ldots, n\} : S \subset I\}, \ \sum_{S \in G} p^{|S|} \leq 1/2,$$

$\langle S \rangle$ = all subsets of $[n]$ that contains S. $\mathbb{P}(\langle S \rangle) = p^{|S|}$. **Examples:**

- $2^{[n]} = \bigcup_{k=1}^{n} \langle k \rangle \cup \emptyset$. If $np \leq 1/2$ than any familly is small.
- $\mathcal{F}$ = all subsets of $[n]$ with cardinality $\geq k+1$. Then

$$\mathcal{F} \subset \bigcup_{|I|=k} \langle I \rangle \subset \bigcup_{|I|=k-1} \langle I \rangle \subset \ldots$$

If $\min_{l \leq k} p^l \binom{n}{l} \leq 1/2$ then $\mathcal{F}$ is small.

Refomulation of Theorem

The following is equivalent to the Theorem formulated on the previous slides.

Theorem

Let $\mathcal{F}$ be a family of subsets of $[n]$, which is not p-small and with each $I \in \mathcal{F}$ we have an associated probabilistic measure μ_I on $[n]$, $\mu_I(I) = 1$. Then $\mathbb{E} \sup_{I \in \mathcal{F}} \sum_{i \in I} \mu_I(i)\delta_i \geq \frac{1}{220}$.

Refomulation of Theorem

The following is equivalent to the Theorem formulated on the previous slides.

Theorem

Let $\mathcal{F}$ be a family of subsets of $[n]$, which is not p-small and with each $I \in \mathcal{F}$ we have an associated probabilistic measure μ_I on $[n]$, $\mu_I(I) = 1$. Then $\mathbb{E} \sup_{I \in \mathcal{F}} \sum_{i \in I} \mu_I(i)\delta_i \geq \frac{1}{220}$.

1. We may assume that $np \geq 1/2$

Refomulation of Theorem

The following is equivalent to the Theorem formulated on the previous slides.

Theorem

Let $\mathcal{F}$ be a family of subsets of $[n]$, which is not p-small and with each $I \in \mathcal{F}$ we have an associated probabilistic measure μ_I on $[n]$, $\mu_I(I) = 1$. Then $\mathbb{E} \sup_{I \in \mathcal{F}} \sum_{i \in I} \mu_I(i)\delta_i \geq \frac{1}{220}$.

1. We may assume that $np \geq 1/2$
2. Improvement over trivial argument for small p

$$\mathbb{E} \sup_{I \in \mathcal{F}} \sum_{i \in I} \mu_I(i)\delta_i \geq \sup_{I \in \mathcal{F}} \mathbb{E} \sum_{i \in I} \mu_I(i)\delta_i = p.$$

Proof - first step

Let $(\delta_i'), (\delta_i'')$ be independent Bernoulli r.v's,
$\mathbb{P}(\delta_i' = 1) = Cp, \ \mathbb{P}(\delta_i'' = 1) = \frac{1}{C}$.

Proof - first step

Let $(\delta_i'), (\delta_i'')$ be independent Bernoulli r.v's,
$\mathbb{P}(\delta_i' = 1) = Cp,\ \mathbb{P}(\delta_i'' = 1) = \frac{1}{C}$. By Jensen's inequality

$$\mathbb{E}\sup_{I\in\mathcal{F}}\sum_i \mu_I(i)\delta_i = \mathbb{E}\sup_{I\in\mathcal{F}}\sum_i \mu_I(i)\delta_i'\delta_i'' \geq \frac{1}{C}\mathbb{E}\sup_{I\in\mathcal{F}}\sum_i \mu_I(i)\delta_i'.$$

Proof - first step

Let $(\delta_i'), (\delta_i'')$ be independent Bernoulli r.v's,
$\mathbb{P}(\delta_i' = 1) = Cp, \ \mathbb{P}(\delta_i'' = 1) = \frac{1}{C}$. By Jensen's inequality

$$\mathbb{E} \sup_{I \in \mathcal{F}} \sum_i \mu_I(i)\delta_i = \mathbb{E} \sup_{I \in \mathcal{F}} \sum_i \mu_I(i)\delta_i'\delta_i'' \geq \frac{1}{C}\mathbb{E} \sup_{I \in \mathcal{F}} \sum_i \mu_I(i)\delta_i'.$$

Denote $S = \sum_{i=1}^{n} \delta_i'$ so that

$$E \sup_{I \in \mathcal{F}} \sum_i \mu_I(i)\delta_i' \geq \sum_{m \geq Cpn} \mathbb{E} \left(\sup_{I \in \mathcal{F}} \sum_i \mu_I(i)\delta_i' \mid S = m \right) \mathbb{P}(S = m).$$

Proof - first step

Let $(\delta_i'), (\delta_i'')$ be independent Bernoulli r.v's,
$\mathbb{P}(\delta_i' = 1) = Cp,\ \mathbb{P}(\delta_i'' = 1) = \frac{1}{C}$. By Jensen's inequality

$$\mathbb{E} \sup_{I \in \mathcal{F}} \sum_i \mu_I(i)\delta_i = \mathbb{E} \sup_{I \in \mathcal{F}} \sum_i \mu_I(i)\delta_i'\delta_i'' \geq \frac{1}{C}\mathbb{E} \sup_{I \in \mathcal{F}} \sum_i \mu_I(i)\delta_i'.$$

Denote $S = \sum_{i=1}^{n} \delta_i'$ so that

$$E \sup_{I \in \mathcal{F}} \sum_i \mu_I(i)\delta_i' \geq \sum_{m \geq Cpn} \mathbb{E}\left(\sup_{I \in \mathcal{F}} \sum_i \mu_I(i)\delta_i' \mid S = m \right) \mathbb{P}(S = m).$$

Since $np \geq 1/2$, we may assume that Cnp is an integer and by N. Lord's result $\mathbb{P}(S \geq Cnp) \geq 1/2$.

Proof - first step

Let $(\delta_i'),(\delta_i'')$ be independent Bernoulli r.v's,
$\mathbb{P}(\delta_i' = 1) = Cp,\ \mathbb{P}(\delta_i'' = 1) = \frac{1}{C}$. By Jensen's inequality

$$\mathbb{E}\sup_{I\in\mathcal{F}}\sum_i \mu_I(i)\delta_i = \mathbb{E}\sup_{I\in\mathcal{F}}\sum_i \mu_I(i)\delta_i'\delta_i'' \geq \frac{1}{C}\mathbb{E}\sup_{I\in\mathcal{F}}\sum_i \mu_I(i)\delta_i'.$$

Denote $S = \sum_{i=1}^n \delta_i'$ so that

$$E\sup_{I\in\mathcal{F}}\sum_i \mu_I(i)\delta_i' \geq \sum_{m\geq Cpn}\mathbb{E}\left(\sup_{I\in\mathcal{F}}\sum_i \mu_I(i)\delta_i' \mid S = m\right)\mathbb{P}(S = m).$$

Since $np \geq 1/2$, we may assume that Cnp is an integer and by N. Lord's result $\mathbb{P}(S \geq Cnp) \geq 1/2$.**Goal:** $\mathbb{E}\left(\sup_{I\in\mathcal{F}}\sum_i \mu_I(i)\delta_i' \mid S = m\right) \geq 0.1$ for $m \geq Cnp$.

Key Lemma

Definition

A set $X \subset [n]$ *is bad if* $\sup_{I \in \mathcal{F}} \mu_I(X \cap I) < 1/2$.

Key Lemma

Definition

A set $X \subset [n]$ *is bad if* $\sup_{I \in \mathcal{F}} \mu_I(X \cap I) < 1/2$.

By referring to X as a random set we mean $X = \{i \in [n] : \delta'_i = 1\} \subset [n]$.

Key Lemma

Definition

A set $X \subset [n]$ *is bad if* $\sup_{I \in \mathcal{F}} \mu_I(X \cap I) < 1/2$.

By referring to X as a random set we mean $X = \{i \in [n] : \delta'_i = 1\} \subset [n]$.

Lemma (key lemma)

Let $\mathcal{F}$ *be not small p-small. Then for any* $m \leq n$

$$\mathbb{P}(X \ \mathit{is\ bad} \mid |X| = m) \leq \sum_{k=1}^{n} \left(4\frac{np}{m}\right)^k .$$

Key Lemma

Definition

A set $X \subset [n]$ is bad if $\sup_{I \in \mathcal{F}} \mu_I(X \cap I) < 1/2$.

By referring to X as a random set we mean $X = \{i \in [n] : \delta'_i = 1\} \subset [n]$.

Lemma (key lemma)

Let $\mathcal{F}$ be not small p-small. Then for any $m \leq n$

$$\mathbb{P}(X \text{ is bad} \mid |X| = m) \leq \sum_{k=1}^{n} \left(4\frac{np}{m}\right)^k .$$

Proof of the main Theorem, assuming key lemma:

$$\mathbb{E}\left(2 \sup_{I \in \mathcal{F}} \sum_i \mu_I(i)\delta'_i \mid S = m\right) = \mathbb{E}\left(2 \sup_{I \in \mathcal{F}} \mu_I(X \cap I) \mid S = m\right)$$

$$\geq \mathbb{P}\left(X \text{ not bad} \mid S = m\right) \overset{\text{key lemma}}{\geq} 1 - \sum_{k=1}^{n} \left(4\frac{np}{m}\right)^k \overset{m \geq Cnp}{\geq} \frac{1}{5}.$$

Structure of a bad set

We write elements of $\mathcal{F} \ni I = (i_1, i_2, \ldots, i_{|I|})$ in such way that $\mu_I(i_1), \mu_I(i_2), \ldots$ is non-increasing. Define $I_j := (i_1, \ldots, i_j)$. So $I_j \subset I$ consisting of j elements with largest $\mu_I(i)$.

Structure of a bad set

We write elements of $\mathcal{F} \ni I = (i_1, i_2, \ldots, i_{|I|})$ in such way that $\mu_I(i_1), \mu_I(i_2), \ldots$ is non-increasing. Define $I_j := (i_1, \ldots, i_j)$. So $I_j \subset I$ consisting of j elements with largest $\mu_I(i)$.

Lemma (Bad sets intersects $\mathcal{F}$ sparsely)

Fix $X \subset [n]$ a bad set. Then for any $I \in \mathcal{F}$ there exists $j = j(I, X)$ such that $|I_j \cap X| < \frac{1}{2}|I_j|$.

Structure of a bad set

We write elements of $\mathcal{F} \ni I = (i_1, i_2, \dots, i_{|I|})$ in such way that $\mu_I(i_1), \mu_I(i_2), \dots$ is non-increasing. Define $I_j := (i_1, \dots, i_j)$. So $I_j \subset I$ consisting of j elements with largest $\mu_I(i)$.

Lemma (Bad sets intersects $\mathcal{F}$ sparsely)

Fix $X \subset [n]$ a bad set. Then for any $I \in \mathcal{F}$ there exists $j = j(I, X)$ such that $|I_j \cap X| < \frac{1}{2}|I_j|$.

Proof: Fix $I \in \mathcal{F}$. For $\varepsilon \in [0, 1]$ we define

$$f(\varepsilon) = \sum_{i \in I \cap X} \mu_I(i) \wedge \varepsilon - \frac{1}{2} \sum_{i \in I} \mu_I(i) \wedge \varepsilon.$$

Structure of a bad set

We write elements of $\mathcal{F} \ni I = (i_1, i_2, \ldots, i_{|I|})$ in such way that $\mu_I(i_1), \mu_I(i_2), \ldots$ is non-increasing. Define $I_j := (i_1, \ldots, i_j)$. So $I_j \subset I$ consisting of j elements with largest $\mu_I(i)$.

Lemma (Bad sets intersects $\mathcal{F}$ sparsely)

Fix $X \subset [n]$ a bad set. Then for any $I \in \mathcal{F}$ there exists $j = j(I, X)$ such that $|I_j \cap X| < \frac{1}{2}|I_j|$.

Proof: Fix $I \in \mathcal{F}$. For $\varepsilon \in [0, 1]$ we define

$$f(\varepsilon) = \sum_{i \in I \cap X} \mu_I(i) \wedge \varepsilon - \frac{1}{2} \sum_{i \in I} \mu_I(i) \wedge \varepsilon.$$

f is continuous, $f(1) < 0$ (X is bad), $f(0) = 0$. So exists largest $\varepsilon(I, X) \in [0, 1]$ such that $f(\varepsilon(I, X)) \geq 0$ and $f(x) < 0$ for $x > \varepsilon(I, X)$.

Structure of a bad set

$$\sum_{i\in I\cap X} \mu_I(i) \wedge \varepsilon(I,X) \geq \frac{1}{2} \sum_{i\in I} \mu_I(i) \wedge \varepsilon(I,X) \quad (\text{since } f(\varepsilon(I,X)) \geq 0), \quad (3)$$

Structure of a bad set

$$\sum_{i \in I \cap X} \mu_I(i) \wedge \varepsilon(I, X) \geq \frac{1}{2} \sum_{i \in I} \mu_I(i) \wedge \varepsilon(I, X) \quad (\text{since } f(\varepsilon(I, X)) \geq 0), \tag{3}$$

For $x > \varepsilon(I, X)$ reverse inequality holds i.e.

$$\sum_{i \in I \cap X} \mu_I(i) \wedge (\varepsilon(I, X) + \delta) < \frac{1}{2} \sum_{i \in I} \mu_I(i) \wedge (\varepsilon(I, X) + \delta). \tag{4}$$

Structure of a bad set

$$\sum_{i\in I\cap X}\mu_I(i)\wedge\varepsilon(I,X)\geq\frac{1}{2}\sum_{i\in I}\mu_I(i)\wedge\varepsilon(I,X)\ \ (\text{since } f(\varepsilon(I,X))\geq 0), \quad (3)$$

For $x>\varepsilon(I,X)$ reverse inequality holds i.e.

$$\sum_{i\in I\cap X}\mu_I(i)\wedge(\varepsilon(I,X)+\delta)<\frac{1}{2}\sum_{i\in I}\mu_I(i)\wedge(\varepsilon(I,X)+\delta). \quad (4)$$

If $\delta>0$ is sufficiently small

$$\sum_{i\in I}\mu_I(i)\wedge(\varepsilon(I,X)+\delta)=\sum_{i\in I}\mu_I(i)\wedge\varepsilon(I,X)+\delta\sum_{i\in I}\mathbf{1}_{\mu_I(i)>\varepsilon(I,X)},$$

and the same hold for I replaced by $I\cap X$.

Structure of a bad set

$$\sum_{i \in I \cap X} \mu_I(i) \wedge \varepsilon(I,X) \geq \frac{1}{2} \sum_{i \in I} \mu_I(i) \wedge \varepsilon(I,X) \quad (\text{since } f(\varepsilon(I,X)) \geq 0), \quad (3)$$

For $x > \varepsilon(I,X)$ reverse inequality holds i.e.

$$\sum_{i \in I \cap X} \mu_I(i) \wedge (\varepsilon(I,X) + \delta) < \frac{1}{2} \sum_{i \in I} \mu_I(i) \wedge (\varepsilon(I,X) + \delta). \qquad (4)$$

If $\delta > 0$ is sufficiently small

$$\sum_{i \in I} \mu_I(i) \wedge (\varepsilon(I,X) + \delta) = \sum_{i \in I} \mu_I(i) \wedge \varepsilon(I,X) + \delta \sum_{i \in I} \mathbf{1}_{\mu_I(i) > \varepsilon(I,X)},$$

and the same hold for I replaced by $I \cap X$.Substituting this to (4) and using (3) gives

$$\sum_{i \in I \cap X} \mathbf{1}_{\mu_I(i) > \varepsilon(I,X)} < \frac{1}{2} \sum_{i \in I} \mathbf{1}_{\mu_I(i) > \varepsilon(I,X)}$$

Witnesses

Since $\sum_{i \in I \cap X} \mathbf{1}_{\mu_I(i) > \varepsilon(I,X)} < \frac{1}{2} \sum_{i \in I} \mathbf{1}_{\mu_I(i) > \varepsilon(I,X)}$, we take $\tilde{I} := \{i \in I : \mu_I(i) > \varepsilon(I,X)\}$ and $j(I,X) := |\tilde{I}|$. The results follows, since $I_{j(I,X)} = \tilde{I}$. □

Witnesses

Since $\sum_{i \in I \cap X} \mathbf{1}_{\mu_I(i) > \varepsilon(I,X)} < \frac{1}{2} \sum_{i \in I} \mathbf{1}_{\mu_I(i) > \varepsilon(I,X)}$, we take $\tilde{I} := \{i \in I : \mu_I(i) > \varepsilon(I,X)\}$ and $j(I,X) := |\tilde{I}|$. The results follows, since $I_{j(I,X)} = \tilde{I}$. □

To remember: Set $I_{j(I,X)}$ consists of elements of I for which coefficient $\mu_I(i)$ exceeds certain level.

Witnesses

Since $\sum_{i \in I \cap X} \mathbf{1}_{\mu_I(i) > \varepsilon(I,X)} < \frac{1}{2} \sum_{i \in I} \mathbf{1}_{\mu_I(i) > \varepsilon(I,X)}$, we take $\tilde{I} := \{i \in I : \mu_I(i) > \varepsilon(I, X)\}$ and $j(I, X) := |\tilde{I}|$. The results follows, since $I_{j(I,X)} = \tilde{I}$. □

To remember: Set $I_{j(I,X)}$ consists of elements of I for which coefficient $\mu_I(i)$ exceeds certain level.

Definition (Pivotal definition)

Fix bad set $X \subset [n]$. To each $I \in \mathcal{F}$ we associate number $j(I, X)$ from the previous lemma. Fix $I \in \mathcal{F}$. Among all $I' \in \mathcal{F}$ such that $I'_{j(I',X)} \setminus X \subset I \setminus X$ we chose the ones for which $j(I', X)$ is smallest. Among the latter we pick any I' such that $I'_{j(I',X)} \setminus X$ has a minimal number of elements. We define (I, X) witness as

$$W(I, X) := I'_{j(I',X)}.$$

Remark: Park and Pham used $I'_{j(I',X)} \setminus X$ rather than $W(I, X)$.

The cover

Definition

The cover of $\mathcal{F}$ is given by $\mathcal{G}(X) = \{W(I,X) \setminus X : \; I \in \mathcal{F}\}$. It is a cover since for any $I \in \mathcal{F}$, $W(I,X) \setminus X \subset I$.

The cover

Definition

The cover of $\mathcal{F}$ is given by $\mathcal{G}(X) = \{W(I, X) \setminus X : \ I \in \mathcal{F}\}$. It is a cover since for any $I \in \mathcal{F}$, $W(I, X) \setminus X \subset I$.

Plan: $\mathcal{F}$ is not $p-$small so $\sum_{G \in \mathcal{G}(X)} p^{|G|} \geq 1/2$ and as a result

$$\mathbb{P}(X \text{ is bad} \ \mid |X| = m)\binom{n}{m} = \sum_{\substack{X \text{ is bad} \\ |X|=m}} 1 \leq 2 \sum_{\substack{X \text{ is bad} \\ |X|=m}} \sum_{G \in \mathcal{G}(X)} p^{|G|}$$

The cover

Definition

The cover of $\mathcal{F}$ is given by $\mathcal{G}(X) = \{W(I,X) \setminus X : \ I \in \mathcal{F}\}$. It is a cover since for any $I \in \mathcal{F}$, $W(I,X) \setminus X \subset I$.

Plan: $\mathcal{F}$ is not $p-$small so $\sum_{G \in \mathcal{G}(X)} p^{|G|} \geq 1/2$ and as a result

$$\mathbb{P}(X \text{ is bad } \mid |X| = m)\binom{n}{m} = \sum_{\substack{X \text{ is bad} \\ |X|=m}} 1 \leq 2 \sum_{\substack{X \text{ is bad} \\ |X|=m}} \sum_{G \in \mathcal{G}(X)} p^{|G|}$$

Questions:

1. How to control the above sum?
2. Can we somehow "parameterize" pairs (G, X)?

Properties of bad sets

Lemma

Let X, Y *be bad sets,* $I, J \in \mathcal{F}$. *Let* $I'_j = W(I, X)$, $J'_k = W(J, Y)$. *Assume that*

$$1) j := j(I', X) = j(J', Y) \; 2) Z := I'_j \cup X = J'_j \cup Y \; 3) t := |I'_j \setminus X| = |J'_j \setminus Y|.$$

Then

$$\sum_{i \in I' \cap Y} \mu_{I'}(i) \wedge \varepsilon(I', X) \geq \frac{1}{2} \sum_{i \in I'} \mu_{I'}(i) \wedge \varepsilon(I', X), \tag{5}$$

where $\varepsilon(\cdot, \cdot)$ *is the proper threshold from the key lemma.*

Properties of bad sets

Lemma

Let X, Y be bad sets, $I, J \in \mathcal{F}$. Let $I'_j = W(I, X)$, $J'_k = W(J, Y)$. Assume that

1)$j := j(I', X) = j(J', Y)$ 2)$Z := I'_j \cup X = J'_j \cup Y$ 3)$t := |I'_j \setminus X| = |J'_j \setminus Y|$.

Then

$$\sum_{i \in I' \cap Y} \mu_{I'}(i) \wedge \varepsilon(I', X) \geq \frac{1}{2} \sum_{i \in I'} \mu_{I'}(i) \wedge \varepsilon(I', X), \tag{5}$$

where $\varepsilon(\cdot, \cdot)$ is the proper threshold from the key lemma.

1. By the definition $\varepsilon(I', Y)$ is the greatest number for which (5) holds. Thus, $\varepsilon(I', Y) \geq \varepsilon(I', X)$. This is a key consequence of the above lemma.
2. In fact under the above assumptions, $W(I, X) \setminus Y = W(J, Y) \setminus Y$.

Proof of Lemma

$$\sum_{i \in I' \cap Y} \mu_{I'}(i) \wedge \varepsilon(I', X) = \sum_{i \in (I'_j)^c \cap Y} \mu_{I'}(i) + \sum_{i \in I'_j \cap Y} \varepsilon(I', X)$$

Proof of Lemma

$$\sum_{i \in I' \cap Y} \mu_{I'}(i) \wedge \varepsilon(I', X) = \sum_{i \in (I'_j)^c \cap Y} \mu_{I'}(i) + \sum_{i \in I'_j \cap Y} \varepsilon(I', X)$$

$$= \sum_{i \in (I'_j)^c \cap Z} \mu_{I'}(i) - \sum_{i \in (I'_j)^c \cap (Z \setminus Y)} \mu_{I'}(i) + \varepsilon(I', X)|I'_j \cap Y|$$

Proof of Lemma

$$\sum_{i \in I' \cap Y} \mu_{I'}(i) \wedge \varepsilon(I', X) = \sum_{i \in (I'_j)^c \cap Y} \mu_{I'}(i) + \sum_{i \in I'_j \cap Y} \varepsilon(I', X)$$

$$= \sum_{i \in (I'_j)^c \cap Z} \mu_{I'}(i) - \sum_{i \in (I'_j)^c \cap (Z \setminus Y)} \mu_{I'}(i) + \varepsilon(I', X) |I'_j \cap Y|$$

$$= \sum_{i \in (I'_j)^c \cap Z} \mu_{I'}(i) - \sum_{i \in (I'_j)^c \cap (J'_j \setminus Y)} \mu_{I'}(i) + \varepsilon(I', X) \left(|I'_j \cap Z| - |I'_j \cap (Z \setminus Y)| \right)$$

Proof of Lemma

$$\sum_{i\in I'\cap Y} \mu_{I'}(i) \wedge \varepsilon(I',X) = \sum_{i\in (I'_j)^c\cap Y} \mu_{I'}(i) + \sum_{i\in I'_j\cap Y} \varepsilon(I',X)$$

$$= \sum_{i\in (I'_j)^c\cap Z} \mu_{I'}(i) - \sum_{i\in (I'_j)^c\cap (Z\setminus Y)} \mu_{I'}(i) + \varepsilon(I',X)|I'_j \cap Y|$$

$$= \sum_{i\in (I'_j)^c\cap Z} \mu_{I'}(i) - \sum_{i\in (I'_j)^c\cap (J'_j\setminus Y)} \mu_{I'}(i) + \varepsilon(I',X)\left(|I'_j \cap Z| - |I'_j \cap (Z \setminus Y)|\right)$$

$$= \sum_{i\in (I'_j)^c\cap Z} \mu_{I'}(i) - \sum_{i\in (I'_j)^c\cap ((J'_j\setminus I'_j)\setminus Y)} \mu_{I'}(i) + \varepsilon(I',X)\left(j - |I'_j \cap (J'_j \setminus Y)|\right)$$

Proof of Lemma

$$\sum_{i\in I'\cap Y} \mu_{I'}(i) \wedge \varepsilon(I',X) = \sum_{i\in (I'_j)^c\cap Y} \mu_{I'}(i) + \sum_{i\in I'_j\cap Y} \varepsilon(I',X)$$

$$= \sum_{i\in (I'_j)^c\cap Z} \mu_{I'}(i) - \sum_{i\in (I'_j)^c\cap(Z\setminus Y)} \mu_{I'}(i) + \varepsilon(I',X)|I'_j \cap Y|$$

$$= \sum_{i\in (I'_j)^c\cap Z} \mu_{I'}(i) - \sum_{i\in (I'_j)^c\cap(J'_j\setminus Y)} \mu_{I'}(i) + \varepsilon(I',X)\left(|I'_j \cap Z| - |I'_j \cap (Z \setminus Y)|\right)$$

$$= \sum_{i\in (I'_j)^c\cap Z} \mu_{I'}(i) - \sum_{i\in (I'_j)^c\cap((J'_j\setminus I'_j)\setminus Y)} \mu_{I'}(i) + \varepsilon(I',X)\left(j - |I'_j \cap (J'_j \setminus Y)|\right)$$

$$= \sum_{i\in (I'_j)^c\cap Z} \mu_{I'}(i) - \sum_{i\in (I'_j)^c\cap\underline{((J'_j\setminus I'_j)\setminus Y)}} \mu_{I'}(i) + \varepsilon(I',X)\left(j - t + \underline{|(J'_j \setminus I'_j) \setminus Y|}\right)$$

Proof of Lemma

$$\sum_{i\in I'\cap Y} \mu_{I'}(i) \wedge \varepsilon(I',X) = \sum_{i\in (I'_j)^c\cap Y} \mu_{I'}(i) + \sum_{i\in I'_j\cap Y} \varepsilon(I',X)$$

$$= \sum_{i\in (I'_j)^c\cap Z} \mu_{I'}(i) - \sum_{i\in (I'_j)^c\cap (Z\setminus Y)} \mu_{I'}(i) + \varepsilon(I',X)|I'_j\cap Y|$$

$$= \sum_{i\in (I'_j)^c\cap Z} \mu_{I'}(i) - \sum_{i\in (I'_j)^c\cap (J'_j\setminus Y)} \mu_{I'}(i) + \varepsilon(I',X)\left(|I'_j\cap Z| - |I'_j\cap (Z\setminus Y)|\right)$$

$$= \sum_{i\in (I'_j)^c\cap Z} \mu_{I'}(i) - \sum_{i\in (I'_j)^c\cap ((J'_j\setminus I'_j)\setminus Y)} \mu_{I'}(i) + \varepsilon(I',X)\left(j - |I'_j\cap (J'_j\setminus Y)|\right)$$

$$= \sum_{i\in (I'_j)^c\cap Z} \mu_{I'}(i) - \sum_{i\in (I'_j)^c\cap \underline{((J'_j\setminus I'_j)\setminus Y)}} \mu_{I'}(i) + \varepsilon(I',X)\left(j - t + \underline{|(J'_j\setminus I'_j)\setminus Y|}\right)$$

$$\geq \sum_{i\in (I'_j)^c\cap Z} \mu_{I'}(i) + \varepsilon(I',X)(j-t) = \sum_{i\in I'\cap X} \mu_{I'}(i) \wedge \varepsilon(I',X)$$

Proof of Lemma

$$\sum_{i\in I'\cap Y}\mu_{I'}(i)\wedge\varepsilon(I',X)=\sum_{i\in(I'_j)^c\cap Y}\mu_{I'}(i)+\sum_{i\in I'_j\cap Y}\varepsilon(I',X)$$

$$=\sum_{i\in(I'_j)^c\cap Z}\mu_{I'}(i)-\sum_{i\in(I'_j)^c\cap(Z\setminus Y)}\mu_{I'}(i)+\varepsilon(I',X)|I'_j\cap Y|$$

$$=\sum_{i\in(I'_j)^c\cap Z}\mu_{I'}(i)-\sum_{i\in(I'_j)^c\cap(J'_j\setminus Y)}\mu_{I'}(i)+\varepsilon(I',X)\Big(|I'_j\cap Z|-|I'_j\cap(Z\setminus Y)|\Big)$$

$$=\sum_{i\in(I'_j)^c\cap Z}\mu_{I'}(i)-\sum_{i\in(I'_j)^c\cap((J'_j\setminus I'_j)\setminus Y)}\mu_{I'}(i)+\varepsilon(I',X)\Big(j-|I'_j\cap(J'_j\setminus Y)|\Big)$$

$$=\sum_{i\in(I'_j)^c\cap Z}\mu_{I'}(i)-\sum_{i\in(I'_j)^c\cap\underline{((J'_j\setminus I'_j)\setminus Y)}}\mu_{I'}(i)+\varepsilon(I',X)\Big(j-t+\underline{|(J'_j\setminus I'_j)\setminus Y|}\Big)$$

$$\geq\sum_{i\in(I'_j)^c\cap Z}\mu_{I'}(i)+\varepsilon(I',X)(j-t)=\sum_{i\in I'\cap X}\mu_{I'}(i)\wedge\varepsilon(I',X)$$

$\geq\frac{1}{2}\sum_{i\in I'}\mu_{I'}(i)\wedge\varepsilon(I',X)$. We used that on $(I'_j)^c,\mu_{I'}(i)\leq\varepsilon(I',X)$. $\square$

Structure of bad sets

Lemma

Under the assumption of the previous lemma, we have
$W(I, X) \setminus Y = W(J, Y) \setminus Y$.

Structure of bad sets

Lemma

Under the assumption of the previous lemma, we have $W(I, X) \setminus Y = W(J, Y) \setminus Y$.

As before let $I'_j = W(I, X), J'_j = W(J, Y)$.

$$I'_j \setminus Y \subset (I'_j \cup X) \setminus Y = (J'_j \cup Y) \setminus Y = J'_j \setminus Y \subset J \setminus Y. \tag{6}$$

Structure of bad sets

Lemma

Under the assumption of the previous lemma, we have $W(I,X) \setminus Y = W(J,Y) \setminus Y$.

As before let $I'_j = W(I,X), J'_j = W(J,Y)$.

$$I'_j \setminus Y \subset (I'_j \cup X) \setminus Y = (J'_j \cup Y) \setminus Y = J'_j \setminus Y \subset J \setminus Y. \tag{6}$$

By the construction, we must have $j_0 := j(I', Y) \geq j$, otwerwise

$$I'_{j_0} \setminus Y \subset I'_j \setminus Y \subset J \setminus Y, \text{ and } J'_j \text{ was not optimal for } Y, J.$$

Structure of bad sets

Lemma

Under the assumption of the previous lemma, we have $W(I,X) \setminus Y = W(J,Y) \setminus Y$.

As before let $I'_j = W(I,X), J'_j = W(J,Y)$.

$$I'_j \setminus Y \subset (I'_j \cup X) \setminus Y = (J'_j \cup Y) \setminus Y = J'_j \setminus Y \subset J \setminus Y. \tag{6}$$

By the construction, we must have $j_0 := j(I', Y) \geq j$, otwerwise

$$I'_{j_0} \setminus Y \subset I'_j \setminus Y \subset J \setminus Y, \text{ and } J'_j \text{ was not optimal for } Y, J.$$

Assume $j_0 = j(I', Y) > j = j(I', X)$. Thus,

$$I'_j = \{i \in I' : \mu_{I'}(i) \geq \varepsilon(I', X)\} \subsetneq I'_{j_0} = \{i \in I' : \mu_{I'}(i) \geq \varepsilon(I', Y)\}$$

so $\varepsilon(I', Y) < \varepsilon(I', X)$. This contradicts the previous lemma, so $j_0 = j$. So by construction $|I'_j \setminus Y| \geq |J'_j \setminus Y|$ and since (6) the assertion follows.

Parametrization of bad sets

Corollary

Fix m, t *and* $Z \subset [n]$, *where* $|Z| = m + t$. *Fix bad set* $X \subset [n]$, $|X| = m$. *For any* $t \leq j \leq n$ *there are at most* $\binom{j}{t}$ *sets of the form* $W(I, X) \setminus X$ *where* $I \in \mathcal{F}, \; |W(I, X)| = j, \; Z = W(I, X) \cup X, \; |W(I, X) \setminus X| = t$.

Parametrization of bad sets

Corollary

Fix m, t and $Z \subset [n]$, where $|Z| = m + t$. Fix bad set $X \subset [n]$, $|X| = m$. For any $t \leq j \leq n$ there are at most $\binom{j}{t}$ sets of the form $W(I, X) \setminus X$ where $I \in \mathcal{F},\ |W(I, X)| = j,\ Z = W(I, X) \cup X,\ |W(I, X) \setminus X| = t$.

Proof: Fix any X, I satisfying conditions in the corollary. Take any other pair Y, J which also satisfies the same conditions.

Parametrization of bad sets

Corollary

Fix m, t and $Z \subset [n]$, where $|Z| = m + t$. Fix bad set $X \subset [n]$, $|X| = m$. For any $t \leq j \leq n$ there are at most $\binom{j}{t}$ sets of the form $W(I, X) \setminus X$ where $I \in \mathcal{F}$, $|W(I, X)| = j$, $Z = W(I, X) \cup X$, $|W(I, X) \setminus X| = t$.

Proof: Fix any X, I satisfying conditions in the corollary. Take any other pair Y, J which also satisfies the same conditions. By previous lemma, $W(Y, J) \setminus Y \subset W(X, I)$ (since $W(Y, J) \setminus Y = W(X, I) \setminus X$). Cardinality of $W(X, I) = j$ so there are at most $\binom{j}{t}$ choices for $W(Y, J) \setminus Y$. □

Parametrization of bad sets

Corollary

Fix m, t and $Z \subset [n]$, where $|Z| = m + t$. Fix bad set $X \subset [n]$, $|X| = m$. For any $t \leq j \leq n$ there are at most $\binom{j}{t}$ sets of the form $W(I, X) \setminus X$ where $I \in \mathcal{F}$, $|W(I, X)| = j$, $Z = W(I, X) \cup X$, $|W(I, X) \setminus X| = t$.

Proof: Fix any X, I satisfying conditions in the corollary. Take any other pair Y, J which also satisfies the same conditions. By previous lemma, $W(Y, J) \setminus Y \subset W(X, I)$ (since $W(Y, J) \setminus Y = W(X, I) \setminus X$). Cardinality of $W(X, I) = j$ so there are at most $\binom{j}{t}$ choices for $W(Y, J) \setminus Y$. □

Remark

X is bad so by one of the previous lemma

$$j = |W(I, X)| = |W(I, X) \cap X| + |W(I, X) \setminus X| \leq \frac{1}{2}j + t, \text{ so } t \geq \frac{1}{2}j.$$

Proof of the key lemma

Goal: upper bound $\mathbb{P}(X \text{ is bad} \mid |X| = m)$.

Proof of the key lemma

Goal: upper bound $\mathbb{P}(X \text{ is bad} \mid |X| = m)$.

Reminder: $\mathbb{P}(X \text{ is bad} \mid |X| = m)\binom{n}{m} \leq 2 \sum_{|X|=m,\text{bad}} \sum_{G \in \mathcal{G}(X)} p^{|G|}$.

Proof of the key lemma

Goal: upper bound $\mathbb{P}(X \text{ is bad} \mid |X| = m)$.

Reminder: $\mathbb{P}(X \text{ is bad} \mid |X| = m)\binom{n}{m} \leq 2 \sum_{|X|=m,\text{bad}} \sum_{G \in \mathcal{G}(X)} p^{|G|}$.

The last hard thing to understand:

$$\sum_{|X|=m,\text{bad}} \sum_{G \in \mathcal{G}(X)} p^{|G|}$$

$$= \sum_{j \geq 1} \sum_{j/2 \leq t \leq j} \sum_{\substack{Z \subset [n] \\ |Z| = m+t}} p^t \Big|\Big\{ \underbrace{W(I,X) \setminus X}_{=G \text{ element from } \mathcal{G}(X)} : Z = W(I,X) \cup X,\ |W(I,X)| = j,\ |W(I,X) \setminus X| = t \Big\}\Big|$$

Proof of the key lemma

Goal: upper bound $\mathbb{P}(X \text{ is bad} \mid |X| = m)$.

Reminder: $\mathbb{P}(X \text{ is bad} \mid |X| = m)\binom{n}{m} \leq 2 \sum_{|X|=m,\text{bad}} \sum_{G \in \mathcal{G}(X)} p^{|G|}$.

The last hard thing to understand:

$$\sum_{|X|=m,\text{bad}} \sum_{G \in \mathcal{G}(X)} p^{|G|} = \sum_{j \geq 1} \sum_{j/2 \leq t \leq j} \sum_{\substack{Z \subset [n] \\ |Z|=m+t}} p^t |\{ \underbrace{W(I,X) \setminus X}_{=G \text{ element from } \mathcal{G}(X)} : Z = W(I,X) \cup X,\ |W(I,X)| = j,\ |W(I,X) \setminus X| = t \}|$$

1. We parameterise pairs (X, G) by $X \cup G$; $|G|$, ; $|G \setminus X|$.
2. The previous Corollary gives upper bound for the cardinality of the set in the above formula.

Proof of the key lemma

The assertion of our Corollary states that

$$|\{W(I,X)\backslash X : Z = W(I,X)\cup X,\ |W(I,X)| = j,\ |W(I,X)\backslash X| = t\}| \leq \binom{j}{t}.$$

Proof of the key lemma

The assertion of our Corollary states that

$$|\{W(I,X)\backslash X : Z = W(I,X)\cup X,\ |W(I,X)| = j,\ |W(I,X)\backslash X| = t\}| \le \binom{j}{t}.$$

$$\frac{1}{2}\mathbb{P}(X \text{ is bad} \mid |X| = m)\binom{n}{m} \le \sum_{|X|=m,\text{bad}} \sum_{G\in\mathcal{G}(X)} p^{|G|}$$

Proof of the key lemma

The assertion of our Corollary states that

$$|\{W(I,X)\setminus X : Z = W(I,X)\cup X,\ |W(I,X)| = j,\ |W(I,X)\setminus X| = t\}| \le \binom{j}{t}.$$

$$\frac{1}{2}\mathbb{P}(X \text{ is bad} \mid |X| = m)\binom{n}{m} \le \sum_{|X|=m,\text{bad}} \sum_{G\in\mathcal{G}(X)} p^{|G|}$$

$$\le \sum_{j\ge 1} \sum_{j/2\le t\le j} \sum_{|Z|=m+t} p^t \binom{j}{t}$$

Proof of the key lemma

The assertion of our Corollary states that

$$|\{W(I,X)\backslash X : Z = W(I,X) \cup X,\ |W(I,X)| = j,\ |W(I,X)\backslash X| = t\}| \leq \binom{j}{t}.$$

$$\frac{1}{2}\mathbb{P}(X \text{ is bad} \mid |X| = m)\binom{n}{m} \leq \sum_{|X|=m,\text{bad}} \sum_{G \in \mathcal{G}(X)} p^{|G|}$$

$$\leq \sum_{j \geq 1} \sum_{j/2 \leq t \leq j} \sum_{|Z|=m+t} p^t \binom{j}{t}$$

$$= \sum_{t=1}^{n} \sum_{j=t}^{2t} \binom{j}{t}\binom{n}{m+t} p^t = \sum_{t=1}^{n} \binom{n}{m+t} p^t \sum_{j=t}^{2t} \binom{j}{t}$$

Proof of the key lemma

The assertion of our Corollary states that

$$|\{W(I,X)\backslash X : Z = W(I,X)\cup X,\ |W(I,X)| = j,\ |W(I,X)\backslash X| = t\}| \le \binom{j}{t}.$$

$$\frac{1}{2}\mathbb{P}(X \text{ is bad} \mid |X| = m)\binom{n}{m} \le \sum_{|X|=m,\text{bad}} \sum_{G\in\mathcal{G}(X)} p^{|G|}$$

$$\le \sum_{j\ge 1} \sum_{j/2\le t\le j} \sum_{|Z|=m+t} p^t \binom{j}{t}$$

$$= \sum_{t=1}^{n} \sum_{j=t}^{2t} \binom{j}{t}\binom{n}{m+t} p^t = \sum_{t=1}^{n} \binom{n}{m+t} p^t \sum_{j=t}^{2t} \binom{j}{t}$$

$$= \sum_{t=1}^{n} \binom{n}{m+t} p^t \binom{2t+1}{t+1} \le \sum_{t=1}^{n} \binom{n}{m+t} (4p)^t \le \sum_{t=1}^{n} \left(4\frac{np}{m}\right)^t \binom{n}{m}$$

since $\binom{n}{m+t} \le \left(\frac{n}{m}\right)^t \binom{n}{m}$. □

Further results and questions

Theorem (Park, Pham; different proof Bednorz, Martynek, Meller)

(E, ρ) a metric space, fix $d \in \mathbb{N}$. Let $X_1, \ldots, X_d$ be i.i.d with values in E with distribution μ. Assume μ has no atoms. Let T be a finite class of nonnegative Borel measurable functions on E. Then there exists a family $\mathcal{F}$ of pairs (g, u) where each $g : E \to \mathbb{R}_+$ is μ-measurable and $u \geq 0$ such that

$$\left\{ \sup_{t \in T} \sum_{i=1}^{d} t(X_i) \geq K \mathbb{E} \sup_{t \in T} \sum_{i=1}^{d} t(X_i) \right\} \subset \bigcup_{(g,u) \in \mathcal{F}} \left\{ \frac{1}{d} \sum_{i=1}^{d} g(X_i) \geq u \right\}$$

$$\sum_{(g,u) \in \mathcal{F}} \mathbb{P} \left(\frac{1}{d} \sum_{i=1}^{d} g(X_i) \geq u \right) \leq \frac{1}{2}.$$

Questions: Can T be infinite? What for not i.i.d variables?
We know: We can skip the assumptions about atoms. Also similar result for invinitely divisable processes.