Stochastic Minimum Norm Combinatorial Optimization

Sharat Ibrahimpur

Chaitanya Swamy

University of Waterloo

Stochastic Minimum Norm Combinatorial Optimization

Sharat Ibrahimpur

Chaitanya Swamy

University of Waterloo

Stochastic Min Norm Optimization (StochNormOpt)

Stochastic Min Norm Optimization (StochNormOpt)

We *introduce* the following general model of 1-stage stochastic optimization:

- We have a combinatorial-optimization problem where costs involved are random variables with given distributions

Stochastic Min Norm Optimization (StochNormOpt)

We *introduce* the following general model of 1-stage stochastic optimization:

- ▷ We have a combinatorial-optimization problem where costs involved are random variables with given distributions
- ▷ Each feasible solution s induces a random m-dimensional cost vector Y^s such that $\{Y_i^s\}_{i\in[m]}$ are independent nonnegative random variables

Y^s follows a product distribution on $\mathbb{R}^m_{\geq 0}$

Stochastic Min Norm Optimization (StochNormOpt)

We *introduce* the following general model of 1-stage stochastic optimization:

- ▷ We have a combinatorial-optimization problem where costs involved are random variables with given distributions
- ▷ Each feasible solution s induces a random m-dimensional cost vector Y^s such that $\{Y_i^s\}_{i \in [m]}$ are <u>independent nonnegative random variables</u>

  Y^s follows a product distribution on $\mathbb{R}^m_{\geq 0}$

- ▷ We are given (a value oracle for) a monotone symmetric norm $f : \mathbb{R}^m_{\geq 0} \to \mathbb{R}_{\geq 0}$

Stochastic Min Norm Optimization (StochNormOpt)

We *introduce* the following general model of 1-stage stochastic optimization:

- ▷ We have a combinatorial-optimization problem where costs involved are random variables with given distributions
- ▷ Each feasible solution s induces a random m-dimensional cost vector Y^s such that $\{Y_i^s\}_{i\in[m]}$ are independent nonnegative random variables

  Y^s follows a product distribution on $\mathbb{R}^m_{\geq 0}$
- ▷ We are given (a value oracle for) a monotone symmetric norm $f : \mathbb{R}^m_{\geq 0} \to \mathbb{R}_{\geq 0}$

  $x \geq y \geq 0 \implies f(x) \geq f(y)$

Stochastic Min Norm Optimization (StochNormOpt)

We *introduce* the following general model of 1-stage stochastic optimization:

- We have a combinatorial-optimization problem where costs involved are random variables with given distributions
- Each feasible solution s induces a random m-dimensional cost vector Y^s such that $\{Y_i^s\}_{i \in [m]}$ are independent nonnegative random variables

  Y^s follows a product distribution on $\mathbb{R}^m_{\geq 0}$

- We are given (a value oracle for) a monotone symmetric norm $f : \mathbb{R}^m_{\geq 0} \to \mathbb{R}_{\geq 0}$

  monotone: $x \geq y \geq 0 \implies f(x) \geq f(y)$

  symmetric: invariant under permutations of coordinates

  $f(x) = f(x^{\downarrow})$, $\quad x^{\downarrow}$ is x sorted in decreasing order

Stochastic Min Norm Optimization (StochNormOpt)

We *introduce* the following general model of 1-stage stochastic optimization:

- ▷ We have a combinatorial-optimization problem where costs involved are random variables with given distributions
- ▷ Each feasible solution s induces a random m-dimensional cost vector Y^s such that $\{Y_i^s\}_{i\in[m]}$ are independent nonnegative random variables

  Y^s follows a product distribution on $\mathbb{R}^m_{\geq 0}$

- ▷ We are given (a value oracle for) a monotone symmetric norm $f : \mathbb{R}^m_{\geq 0} \to \mathbb{R}_{\geq 0}$

$x \geq y \geq 0 \implies f(x) \geq f(y)$

invariant under permutations of coordinates
$f(x) = f(x^{\downarrow}),\quad x^{\downarrow}$ is x sorted in decreasing order

Goal: Find an oblivious solution s that minimizes $\mathbb{E}[f(Y^s)]$

Stochastic Min Norm Load Balancing (StochNormLB)

Stochastic Min Norm Load Balancing (StochNormLB)

Input

J set of n independent stochastic jobs

m unrelated parallel machines $\{1, 2, \ldots, m\}$

X_{ij} nonnegative random variable denoting processing time of job j on machine i

(we assume X_{ij} and $X_{i'j'}$ are independent whenever $j \neq j'$)

$f : \mathbb{R}^m_{\geq 0} \to \mathbb{R}_{\geq 0}$ monotone symmetric norm

Stochastic Min Norm Load Balancing (StochNormLB)

Input

J set of n independent stochastic jobs

m unrelated parallel machines $\{1,2,\ldots,m\}$

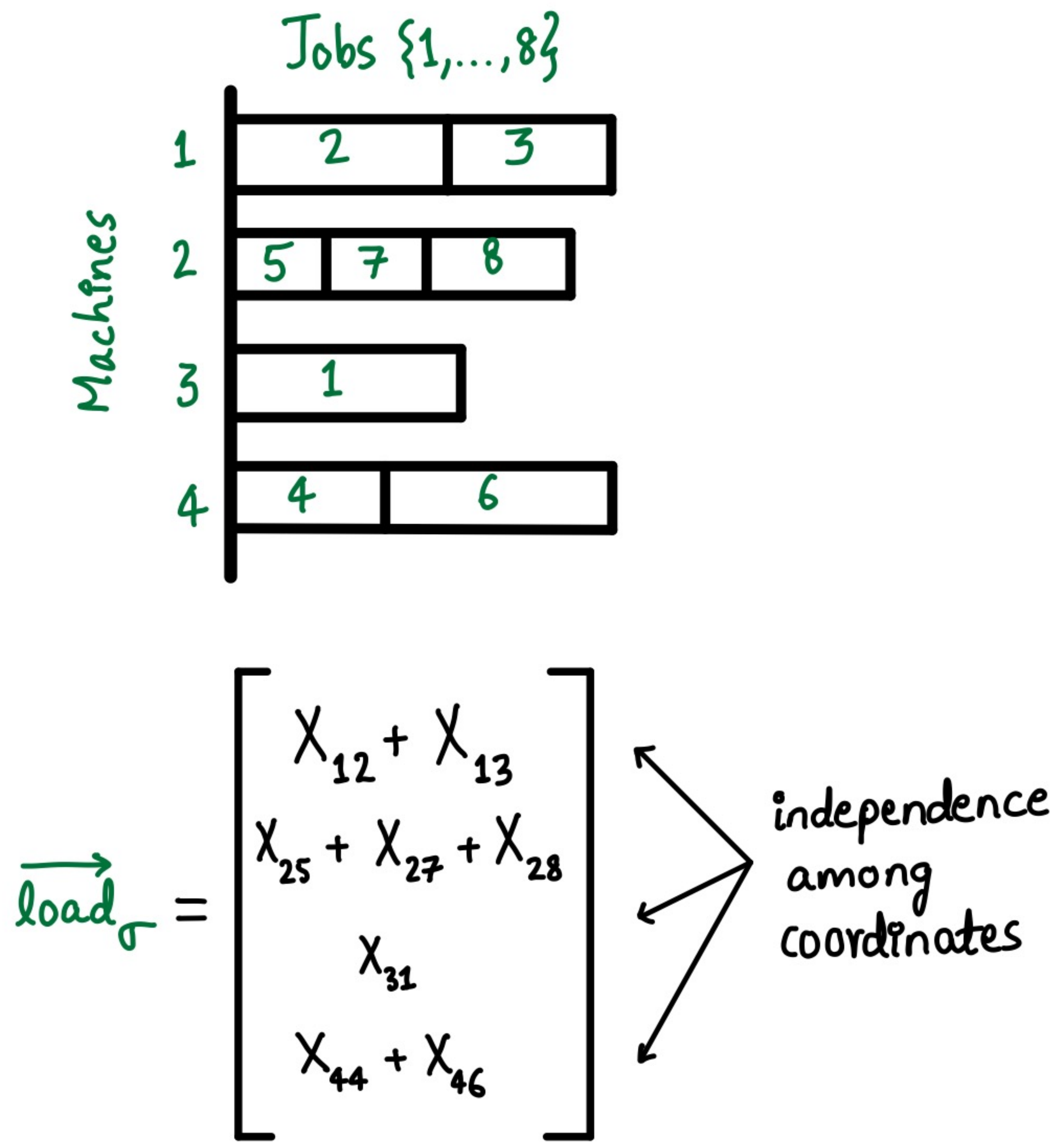
X_{ij} nonnegative random variable denoting processing time of job j on machine i

(we assume X_{ij} and $X_{i'j'}$ are independent whenever $j \neq j'$)

$f : \mathbb{R}^m_{\geq 0} \to \mathbb{R}_{\geq 0}$ monotone symmetric norm

Feasible Solutions

Assignment $\sigma : J \to \{1,\ldots,m\}$ of jobs to machines



Stochastic Min Norm Load Balancing (StochNormLB)

Input

J set of n independent stochastic jobs

m unrelated parallel machines $\{1,2,\ldots,m\}$

X_{ij} nonnegative random variable denoting processing time of job j on machine i

(we assume X_{ij} and $X_{i'j'}$ are independent whenever $j \neq j'$)

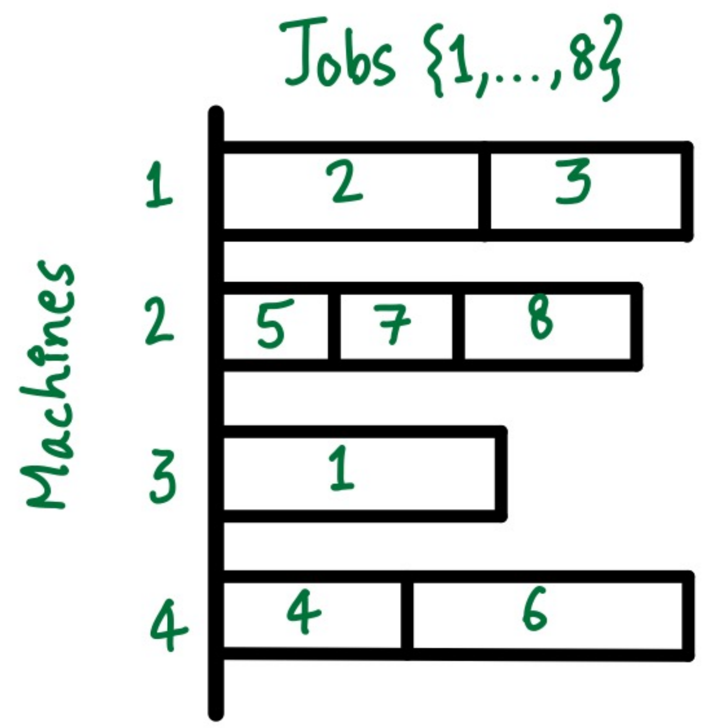
$f : \mathbb{R}^m_{\geq 0} \to \mathbb{R}_{\geq 0}$ monotone symmetric norm

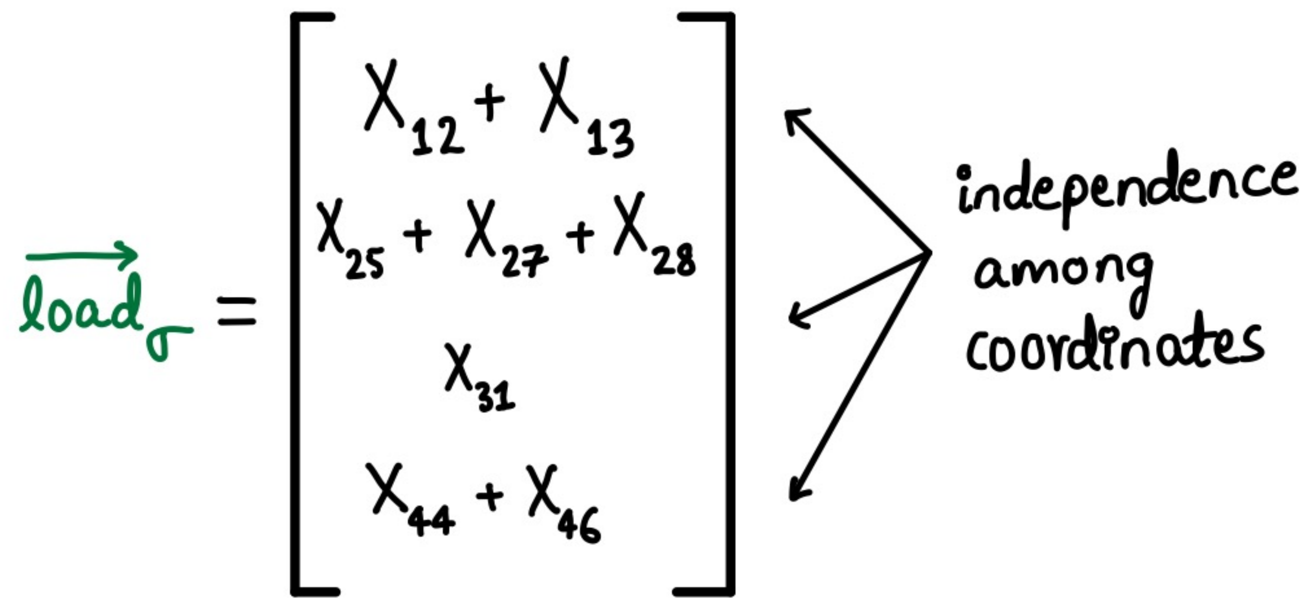
Feasible Solutions

Assignment $\sigma : J \to \{1,\ldots,m\}$ of jobs to machines

Goal

Find σ that minimizes $\mathbb{E}\left[f(\overrightarrow{\text{load}}_\sigma)\right]$





Stochastic Min Norm Spanning Tree (StochNormTree)

Stochastic Min Norm Spanning Tree (StochNormTree)

Input

$G = (V, E)$ undirected graph on n vertices

X_e nonnegative random variable denoting stochastic weight of edge $e \in E$

(we assume $\{X_e\}_{e \in E}$ are independent)

$f : \mathbb{R}^{n-1}_{\geq 0} \to \mathbb{R}_{\geq 0}$ monotone symmetric norm

Stochastic Min Norm Spanning Tree (StochNormTree)

Input

$G = (V, E)$ undirected graph on n vertices

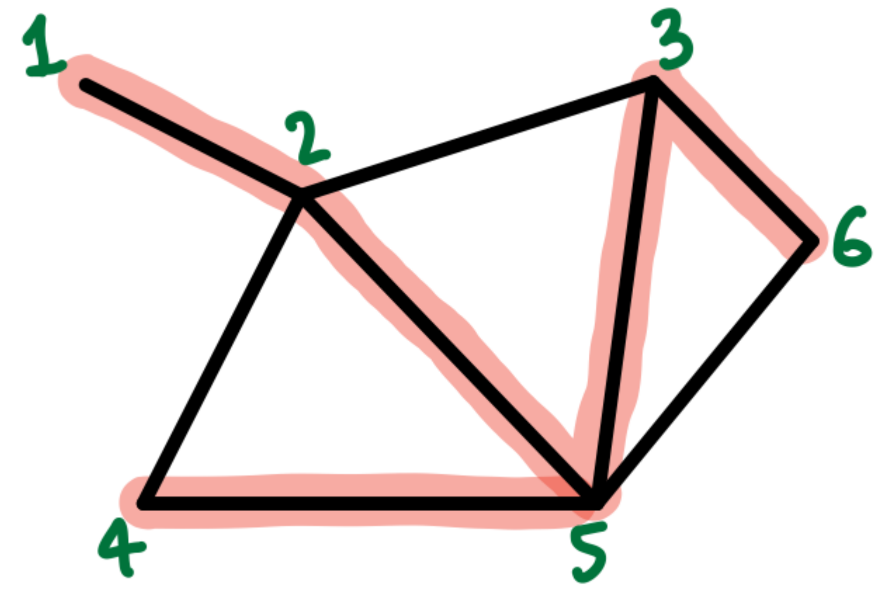
X_e nonnegative random variable denoting stochastic weight of edge $e \in E$

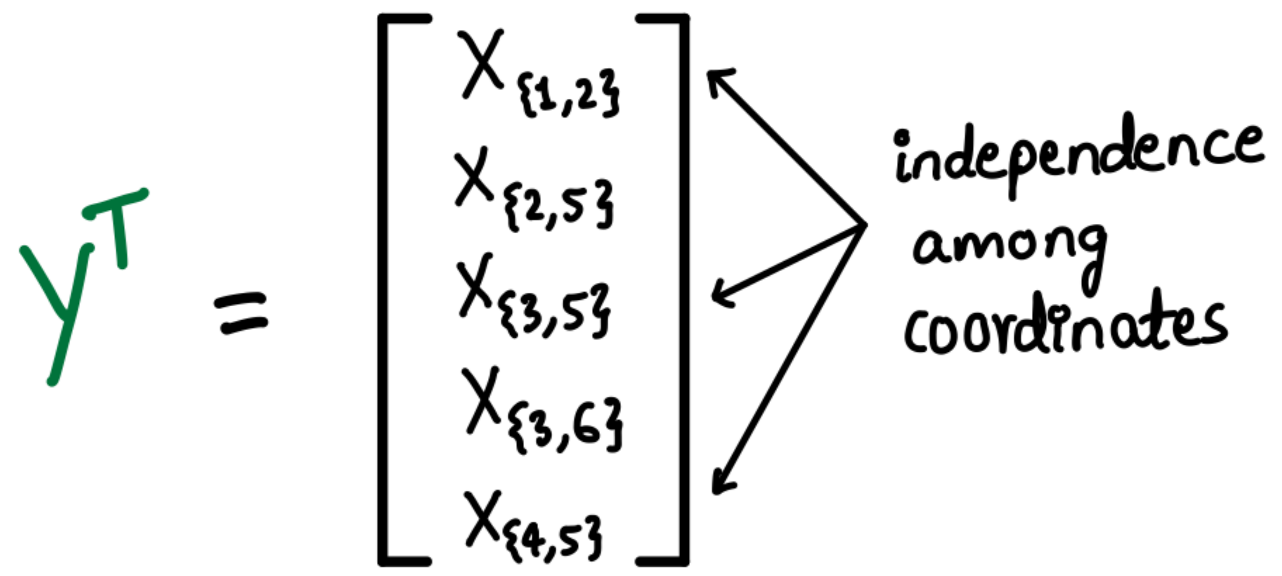
(we assume $\{X_e\}_{e \in E}$ are independent)

$f : \mathbb{R}^{n-1}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ monotone symmetric norm

Feasible Solutions

Spanning trees of G





Stochastic Min Norm Spanning Tree (StochNormTree)

Input

$G = (V, E)$ undirected graph on n vertices

X_e nonnegative random variable denoting stochastic weight of edge $e \in E$

(we assume $\{X_e\}_{e \in E}$ are independent)

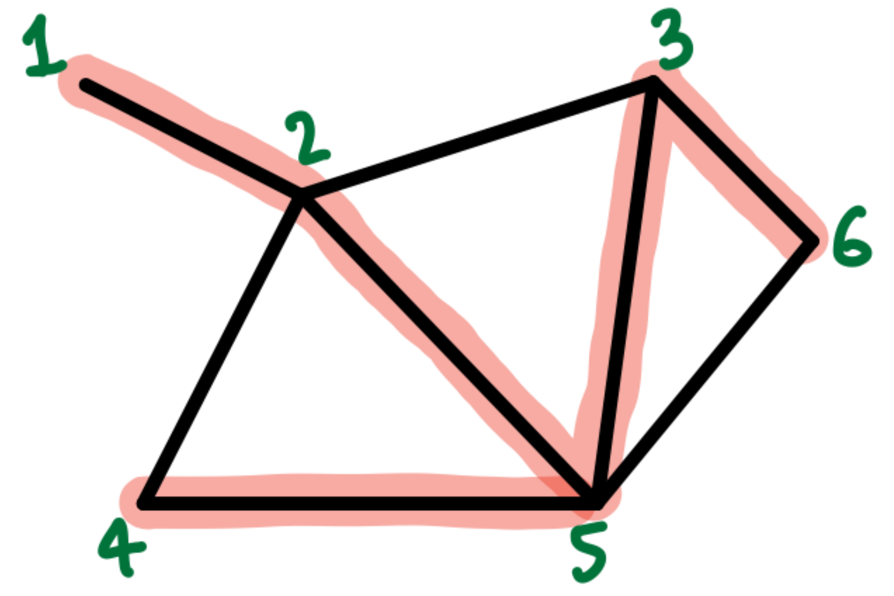
$f : \mathbb{R}_{\geq 0}^{n-1} \to \mathbb{R}_{\geq 0}$ monotone symmetric norm

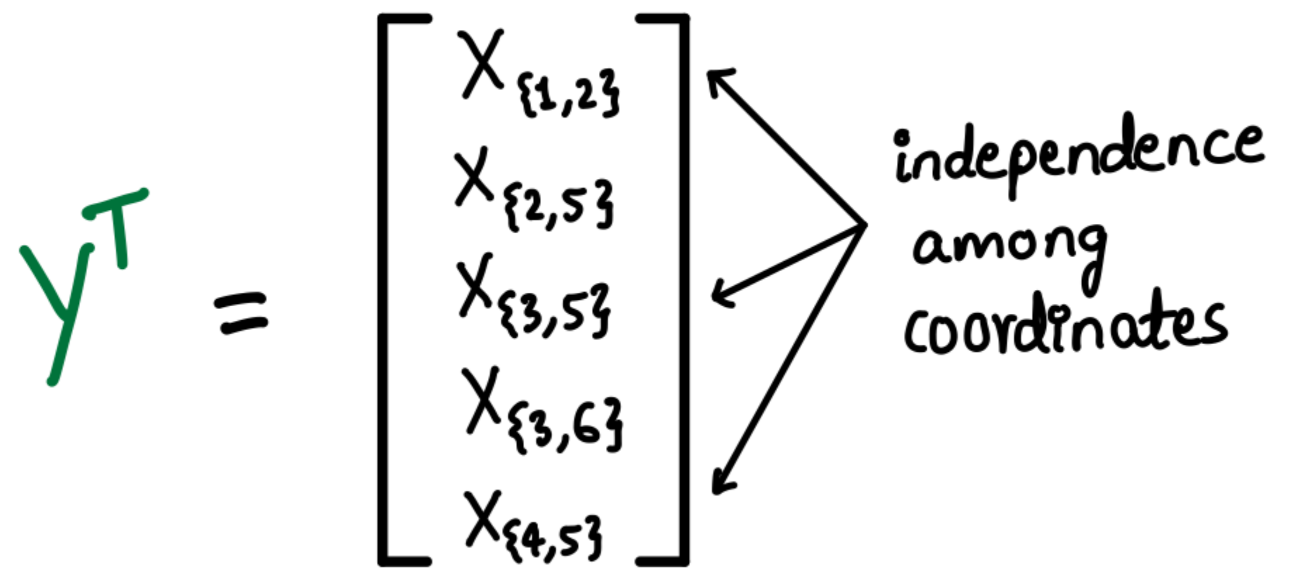
Feasible Solutions

Spanning trees of G

Goal

Find spanning tree $T = \{e_1, \ldots, e_{n-1}\}$ that minimizes $\mathbb{E}[f(Y^T)] = \mathbb{E}\left[f(X_{e_1}, \ldots, X_{e_{n-1}})\right]$





Some Remarks and Assumptions

Some Remarks and Assumptions

Goal: Find an oblivious solution s that minimizes $\mathbb{E}[f(Y^s)]$

Some Remarks and Assumptions

Goal: Find an oblivious solution s that minimizes $\mathbb{E}[f(Y^s)]$

▷ Three sources of complexity:

Feasibility for the combinatorial optimization problem

Stochasticity of costs

Controlling the norm of a cost vector

Some Remarks and Assumptions

Goal: Find an oblivious solution s that minimizes $\mathbb{E}[f(Y^s)]$

▷ Three sources of complexity:

Feasibility for the combinatorial optimization problem

Stochasticity of costs

Controlling the norm of a cost vector

▷ Assumptions:

Complete distributional information of r.v.s (job-size r.v.s and edge-weight r.v.s)

Can sample from these distributions

Can compute expected value, evaluate moment generating functions, or truncate above/below a threshold etc.

Significance of Our Work

Significance of Our Work

- Uncertainty is ubiquitous in real-world optimization problems
- A principled approach that works across multiple combinatorial-optimization settings

Significance of Our Work

- Uncertainty is ubiquitous in real-world optimization problems
- A principled approach that works across multiple combinatorial-optimization settings
- Working at the generality of monotone symmetric norms has several benefits:

Significance of Our Work

- Uncertainty is ubiquitous in real-world optimization problems
- A principled approach that works across multiple combinatorial-optimization settings
- Working at the generality of monotone symmetric norms has several benefits:
  - ▷ Contains popular objectives:

Significance of Our Work

- Uncertainty is ubiquitous in real-world optimization problems
- A principled approach that works across multiple combinatorial-optimization settings
- Working at the generality of monotone symmetric norms has several benefits:
  - ▷ Contains popular objectives:

min-max problem
$\max_i x_i$

Significance of Our Work

- Uncertainty is ubiquitous in real-world optimization problems
- A principled approach that works across multiple combinatorial-optimization settings
- Working at the generality of monotone symmetric norms has several benefits:
  - ▷ Contains popular objectives:

min-max problem
$\max_i x_i$

min-sum problem
$\sum_i x_i$

Significance of Our Work

- Uncertainty is ubiquitous in real-world optimization problems
- A principled approach that works across multiple combinatorial-optimization settings
- Working at the generality of monotone symmetric norms has several benefits:
  - ▷ Contains popular objectives:

min-max problem	ℓ_p norms	min-sum problem
$\max_i x_i$	$\left(\sum_i x_i^p\right)^{1/p}$	$\sum_i x_i$

Significance of Our Work

- Uncertainty is ubiquitous in real-world optimization problems
- A principled approach that works across multiple combinatorial-optimization settings
- Working at the generality of monotone symmetric norms has several benefits:
  - ▷ Contains popular objectives:

min-max problem
$\max_i x_i$

ℓ_p norms
$\left(\sum_i x_i^p\right)^{1/p}$

min-sum problem
$\sum_i x_i$

Top_ℓ norms, $\ell \in \{1,2,\ldots,m\}$
$\mathrm{Top}_\ell(x) =$ sum of ℓ largest entries of x

Significance of Our Work

- Uncertainty is ubiquitous in real-world optimization problems
- A principled approach that works across multiple combinatorial-optimization settings
- Working at the generality of monotone symmetric norms has several benefits:

▷ Contains popular objectives:

ℓ_p norms
$$\left(\sum_i x_i^p\right)^{1/p}$$

min-max problem
$$\max_i x_i$$

min-sum problem
$$\sum_i x_i$$

Top_ℓ norms, $\ell \in \{1,2,\ldots,m\}$
$\text{Top}_\ell(x)$ = sum of ℓ largest entries of x

▷ Unify algorithm-design principles for a wide class of objectives under one umbrella

Significance contd…

Significance contd...

- Increased modeling power due to closure properties under taking
  nonnegative linear combinations (f, f' mon. sym. norms $\implies \alpha f + \beta f'$ is a mon. sym. norm)
  and
  pointwise maximums (f, f' mon. sym. norms $\implies \max(f, f')$ is a mon. sym. norm)

Significance contd...

- ▷ Increased modeling power due to closure properties under taking nonnegative linear combinations (f, f' mon. sym. norms $\implies \alpha f + \beta f'$ is a mon. sym. norm) and pointwise maximums (f, f' mon. sym. norms $\implies \max(f, f')$ is a mon. sym. norm)

Feasibility problem with multiple norm-budget constraints:

Find $x \in P$ s.t. $f_r(x) \leq B_r \, \forall r = 1, \ldots, k$

Significance contd...

▷ Increased modeling power due to closure properties under taking

nonnegative linear combinations (f, f' mon. sym. norms $\implies \alpha f + \beta f'$ is a mon. sym. norm)

and

pointwise maximums (f, f' mon. sym. norms $\implies \max(f, f')$ is a mon. sym. norm)

Feasibility problem with multiple norm-budget constraints:

Find $x \in P$ s.t. $f_r(x) \leq B_r \forall r = 1,\ldots,k$

reduces to

Norm-minimization problem:

Is $\min\{g(x) : x \in P\} \leq 1$?

where $g(x) := \max_{r=1,\ldots,k} \frac{f_r(x)}{B_r}$

Summary of Our Contributions

Summary of Our Contributions

Approximation-Algorithms Framework for StochNormOpt

Summary of Our Contributions

Approximation-Algorithms Framework for StochNormOpt

Component #1: Work with simpler norms

Summary of Our Contributions

Approximation-Algorithms Framework for StochNormOpt

Component #1: Work with simpler norms

Find solution s such that
$$\mathbb{E}[f(Y^s)] = O(\alpha) \cdot \mathbb{E}[f(Y^{s^*})],$$
where s^* is an optimal solution

Summary of Our Contributions

Approximation-Algorithms Framework for StochNormOpt

Component #1: Work with simpler norms

Find solution s such that
$\mathbb{E}[f(Y^s)] = O(\alpha) \cdot \mathbb{E}[f(Y^{s^*})]$,
where s^* is an optimal solution

reduces to
(at an $O(1)$ loss
in approximation)

Summary of Our Contributions

Approximation-Algorithms Framework for StochNormOpt

Component #1: Work with simpler norms

Find solution s such that $\mathbb{E}[f(Y^s)] = O(\alpha) \cdot \mathbb{E}[f(Y^{s^*})]$, where s^* is an optimal solution	reduces to (at an $O(1)$ loss in approximation)	Find solution s such that $\mathbb{E}[\mathsf{Top}_\ell(Y^s)] \leq \alpha \cdot \mathbb{E}[\mathsf{Top}_\ell(Y^{s^*})]$ for all $\ell = 1,2,4,\ldots,2^{\lfloor \log_2 m \rfloor}$

Summary of Our Contributions

Approximation-Algorithms Framework for StochNormOpt

Component #1: Work with simpler norms

Find solution s such that $\mathbb{E}[f(Y^s)] = O(\alpha) \cdot \mathbb{E}[f(Y^{s^*})]$, where s^* is an optimal solution	reduces to (at an $O(1)$ loss in approximation)	Find solution s such that $\mathbb{E}[\mathsf{Top}_\ell(Y^s)] \leq \alpha \cdot \mathbb{E}[\mathsf{Top}_\ell(Y^{s^*})]$ for all $\ell = 1,2,4,\ldots,2^{\lfloor \log_2 m \rfloor}$

Component #2: Tools for handling stochastic optimization problem for a single Top_ℓ norm

Summary of Our Contributions

Approximation-Algorithms Framework for StochNormOpt

Component #1: Work with simpler norms

Find solution s such that $\mathbb{E}[f(Y^s)] = O(\alpha) \cdot \mathbb{E}[f(Y^{s^*})]$, where s^* is an optimal solution	reduces to (at an $O(1)$ loss in approximation)	Find solution s such that $\mathbb{E}[\mathsf{Top}_\ell(Y^s)] \le \alpha \cdot \mathbb{E}[\mathsf{Top}_\ell(Y^{s^*})]$ for all $\ell = 1,2,4,\ldots,2^{\lfloor \log_2 m \rfloor}$

Component #2: Tools for handling stochastic optimization problem for a single Top_ℓ norm

Our Contributions contd...

Our Contributions contd...

- Application of our framework to obtain approximation algorithms for
  stochastic min-norm load balancing and
  stochastic min-norm spanning tree

  (I will mention our main results towards the end of the talk)

Our Contributions contd...

- Application of our framework to obtain approximation algorithms for
  stochastic min-norm load balancing and
  stochastic min-norm spanning tree

  (I will mention our main results towards the end of the talk)

Most of this work has appeared at FOCS 2020, ICALP 2021, and SOSA 2022
[IS20] [IS21] [IS22]

How do we reason about $\mathbb{E}[f(Y)]$?

$f : \mathbb{R}^m_{\geq 0} \to \mathbb{R}_{\geq 0}$ an arbitrary monotone symmetric norm

Y an arbitrary product distribution on $\mathbb{R}^m_{\geq 0}$

Component #1: Reduction to Simpler Norms

Component #1: Reduction to Simpler Norms

Some results from the deterministic world…

Component #1: Reduction to Simpler Norms

Some results from the deterministic world…

▷ Classical result in majorization theory [HLP34]: for any $x, y \in \mathbb{R}^m_{\geq 0}$,

$\mathsf{Top}_\ell(x) \leq \mathsf{Top}_\ell(y)\, \forall \ell \in \{1, \ldots, m\} \implies f(x) \leq f(y)$ for any monotone symmetric norm f

Component #1: Reduction to Simpler Norms

Some results from the deterministic world...

- ▷ Classical result in majorization theory [HLP34]: for any $x, y \in \mathbb{R}^m_{\geq 0}$,

  $\mathsf{Top}_\ell(x) \leq \mathsf{Top}_\ell(y)\, \forall \ell \in \{1, \ldots, m\} \implies f(x) \leq f(y)$ for any monotone symmetric norm f

- ▷ By homogeneity of norms, working only with powers-of-2 ℓ leads to at most factor-2 loss in approximation

Component #1: Reduction to Simpler Norms

Some results from the deterministic world...

- ▷ Classical result in majorization theory [HLP34]: for any $x, y \in \mathbb{R}^m_{\geq 0}$,

  $\text{Top}_\ell(x) \leq \text{Top}_\ell(y)\ \forall \ell \in \{1,\ldots,m\} \implies f(x) \leq f(y)$ for any monotone symmetric norm f

- ▷ By homogeneity of norms, working only with powers-of-2 ℓ leads to at most factor-2 loss in approximation

- ▷ For any decreasing vector $w \in \mathbb{R}^m_{\geq 0}$, the w-ordered norm is defined as:

$$\|x\|_w := w^T x^{\downarrow} = \sum_{\ell \in [m]} w_\ell x^{\downarrow}_\ell = \sum_{\ell \in [m]} (w_\ell - w_{\ell+1})\text{Top}_\ell(x)$$

Component #1: Reduction to Simpler Norms

Some results from the deterministic world…

- ▷ Classical result in majorization theory [HLP34]: for any $x, y \in \mathbb{R}^m_{\geq 0}$,

  $\text{Top}_\ell(x) \leq \text{Top}_\ell(y)\ \forall \ell \in \{1,\ldots,m\} \implies f(x) \leq f(y)$ for any monotone symmetric norm f

- ▷ By homogeneity of norms, working only with powers-of-2 ℓ leads to at most factor-2 loss in approximation

- ▷ For any decreasing vector $w \in \mathbb{R}^m_{\geq 0}$, the w-ordered norm is defined as:

$$\|x\|_w := w^T x^\downarrow = \sum_{\ell \in [m]} w_\ell x^\downarrow_\ell = \sum_{\ell \in [m]} (w_\ell - w_{\ell+1}) \text{Top}_\ell(x)$$

Theorem [see CS19]: For any monotone symmetric norm $f : \mathbb{R}^m_{\geq 0} \to \mathbb{R}_{\geq 0}$,

$\exists$ a (possibly uncountable) collection $\mathscr{C}$ of ordered norms s.t.

$$f(x) = \sup_{w \in \mathscr{C}} w^T x^\downarrow \quad \text{for all } x$$

Component #1: Reduction to Simpler Norms

Some results from the deterministic world...

- ▷ Classical result in majorization theory [HLP34]: for any $x, y \in \mathbb{R}^m_{\geq 0}$,

  $\text{Top}_\ell(x) \leq \text{Top}_\ell(y) \, \forall \ell \in \{1, \ldots, m\} \implies f(x) \leq f(y)$ for any monotone symmetric norm f

- ▷ By homogeneity of norms, working only with powers-of-2 ℓ leads to at most factor-2 loss in approximation

- ▷ For any decreasing vector $w \in \mathbb{R}^m_{\geq 0}$, the w-ordered norm is defined as:

$$\|x\|_w := w^T x^\downarrow = \sum_{\ell \in [m]} w_\ell x^\downarrow_\ell = \sum_{\ell \in [m]} (w_\ell - w_{\ell+1}) \text{Top}_\ell(x)$$

Theorem [see CS19]: For any monotone symmetric norm $f : \mathbb{R}^m_{\geq 0} \to \mathbb{R}_{\geq 0}$,

$\exists$ a (possibly uncountable) collection $\mathscr{C}$ of ordered norms s.t.

$$f(x) = \sup_{w \in \mathscr{C}} w^T x^\downarrow \quad \text{for all } x$$

- ▷ Does the above extend to the stochastic setting?

Component #1: Reduction to Simpler Norms

Some results from the deterministic world...

- ▷ Classical result in majorization theory [HLP34]: for any $x, y \in \mathbb{R}^m_{\geq 0}$,

  $\mathsf{Top}_\ell(x) \leq \mathsf{Top}_\ell(y)\ \forall \ell \in \{1,\ldots,m\} \implies f(x) \leq f(y)$ for any monotone symmetric norm f

- ▷ By homogeneity of norms, working only with powers-of-2 ℓ leads to at most factor-2 loss in approximation

- ▷ For any decreasing vector $w \in \mathbb{R}^m_{\geq 0}$, the w-ordered norm is defined as:

$$\|x\|_w := w^T x^{\downarrow} = \sum_{\ell \in [m]} w_\ell x^{\downarrow}_\ell = \sum_{\ell \in [m]} (w_\ell - w_{\ell+1})\mathsf{Top}_\ell(x)$$

Theorem [see CS19]: For any monotone symmetric norm $f : \mathbb{R}^m_{\geq 0} \to \mathbb{R}_{\geq 0}$,

$\exists$ a (possibly uncountable) collection $\mathscr{C}$ of ordered norms s.t.

$$f(x) = \sup_{w \in \mathscr{C}} w^T x^{\downarrow} \quad \text{for all } x$$

- ▷ Does the above extend to the stochastic setting? YES!! approximately..

Approximate Stochastic Majorization

Approximate Stochastic Majorization

Theorem [IS20]

For product distributions Y and W, and any monotone symmetric norm f,

$$\mathbb{E}[\mathsf{Top}_\ell(Y)] \leq \alpha \cdot \mathbb{E}[\mathsf{Top}_\ell(W)] \; \forall \ell \in [m] \implies \mathbb{E}[f(Y)] \leq O(\alpha) \cdot \mathbb{E}[f(W)]$$

Approximate Stochastic Majorization

Theorem [IS20]

For product distributions Y and W, and any monotone symmetric norm f,

$$\mathbb{E}[\mathrm{Top}_\ell(Y)] \leq \alpha \cdot \mathbb{E}[\mathrm{Top}_\ell(W)]\ \forall \ell \in [m] \implies \mathbb{E}[f(Y)] \leq O(\alpha) \cdot \mathbb{E}[f(W)]$$

Consequence of:

Main Theorem [IS20]: We have $f(\mathbb{E}[Y^{\downarrow}]) \leq \mathbb{E}[f(Y)] \leq O(1) \cdot f(\mathbb{E}[Y^{\downarrow}])$

Approximate Stochastic Majorization

Theorem [IS20]

For product distributions Y and W, and any monotone symmetric norm f,

$$\mathbb{E}[\mathrm{Top}_\ell(Y)] \leq \alpha \cdot \mathbb{E}[\mathrm{Top}_\ell(W)] \; \forall \ell \in [m] \implies \mathbb{E}[f(Y)] \leq O(\alpha) \cdot \mathbb{E}[f(W)]$$

Consequence of:

Main Theorem [IS20]: We have $f(\mathbb{E}[Y^\downarrow]) \leq \mathbb{E}[f(Y)] \leq O(1) \cdot f(\mathbb{E}[Y^\downarrow])$

▷ Lower bound follows from symmetry and convexity: $\mathbb{E}[f(Y)] = \mathbb{E}[f(Y^\downarrow)] \geq f(\mathbb{E}[Y^\downarrow])$

Approximate Stochastic Majorization

Theorem [IS20]

For product distributions Y and W, and any monotone symmetric norm f,

$$\mathbb{E}[\text{Top}_\ell(Y)] \leq \alpha \cdot \mathbb{E}[\text{Top}_\ell(W)] \, \forall \ell \in [m] \implies \mathbb{E}[f(Y)] \leq O(\alpha) \cdot \mathbb{E}[f(W)]$$

Consequence of:

Main Theorem [IS20]: We have $f(\mathbb{E}[Y^{\downarrow}]) \leq \mathbb{E}[f(Y)] \leq O(1) \cdot f(\mathbb{E}[Y^{\downarrow}])$

▷ Lower bound follows from symmetry and convexity: $\mathbb{E}[f(Y)] = \mathbb{E}[f(Y^{\downarrow})] \geq f(\mathbb{E}[Y^{\downarrow}])$

▷ Note $\mathbb{E}[Y^{\downarrow}_\ell] =$ expected value of the ℓ^{th} maximum in Y, and $\mathbb{E}[\text{Top}_\ell(Y)] = \text{Top}_\ell\big(\mathbb{E}[Y^{\downarrow}]\big)$

Approximate Stochastic Majorization

Theorem [IS20]

For product distributions Y and W, and any monotone symmetric norm f,

$$\mathbb{E}[\mathrm{Top}_\ell(Y)] \le \alpha \cdot \mathbb{E}[\mathrm{Top}_\ell(W)]\ \forall \ell \in [m] \implies \mathbb{E}[f(Y)] \le O(\alpha) \cdot \mathbb{E}[f(W)]$$

Consequence of:

Main Theorem [IS20]: We have $f(\mathbb{E}[Y^{\downarrow}]) \le \mathbb{E}[f(Y)] \le O(1) \cdot f(\mathbb{E}[Y^{\downarrow}])$

▷ Lower bound follows from symmetry and convexity: $\mathbb{E}[f(Y)] = \mathbb{E}[f(Y^{\downarrow})] \ge f(\mathbb{E}[Y^{\downarrow}])$

▷ Note $\mathbb{E}[Y^{\downarrow}_\ell] =$ expected value of the ℓ^{th} maximum in Y, and $\mathbb{E}[\mathrm{Top}_\ell(Y)] = \mathrm{Top}_\ell\big(\mathbb{E}[Y^{\downarrow}]\big)$

▷ Proof of approximate stochastic majorization using main theorem:

$$\mathbb{E}[f(Y)] = O(1) \cdot f(\mathbb{E}[Y^{\downarrow}]) \le O(1) \cdot \alpha \cdot f(\mathbb{E}[W^{\downarrow}]) \le O(1) \cdot \alpha \cdot \mathbb{E}[f(W)]$$

Component #2: Stochastic Top_l Optimization for a single ℓ

Component #2: Stochastic Top_l Optimization for a single ℓ

- Non-separable nature of Top_ℓ norms makes controlling $\mathbb{E}[\mathsf{Top}_l(Y)]$ challenging

Component #2: Stochastic Top_l Optimization for a single ℓ

- Non-separable nature of Top_ℓ norms makes controlling $\mathbb{E}[\mathsf{Top}_l(Y)]$ challenging

$\mathbb{E}[\mathsf{Top}_1(Y)]$ is usually controlled by deriving probability bounds on the upper tail of $Y_i \forall i \in [m]$

Component #2: Stochastic Top_l Optimization for a single ℓ

- Non-separable nature of Top_ℓ norms makes controlling $\mathbb{E}[\mathsf{Top}_l(Y)]$ challenging

$\mathbb{E}[\mathsf{Top}_1(Y)]$ is usually controlled by deriving probability bounds on the upper tail of $Y_i \, \forall i \in [m]$

- We give two simple separable proxy functions for $\mathbb{E}[\mathsf{Top}_l(Y)]$ that lead to an $O(1)$ loss in approximation

First Proxy Function for Stochastic Top_l Optimization

Based on exceptional variables....

First Proxy Function for Stochastic Top_l Optimization

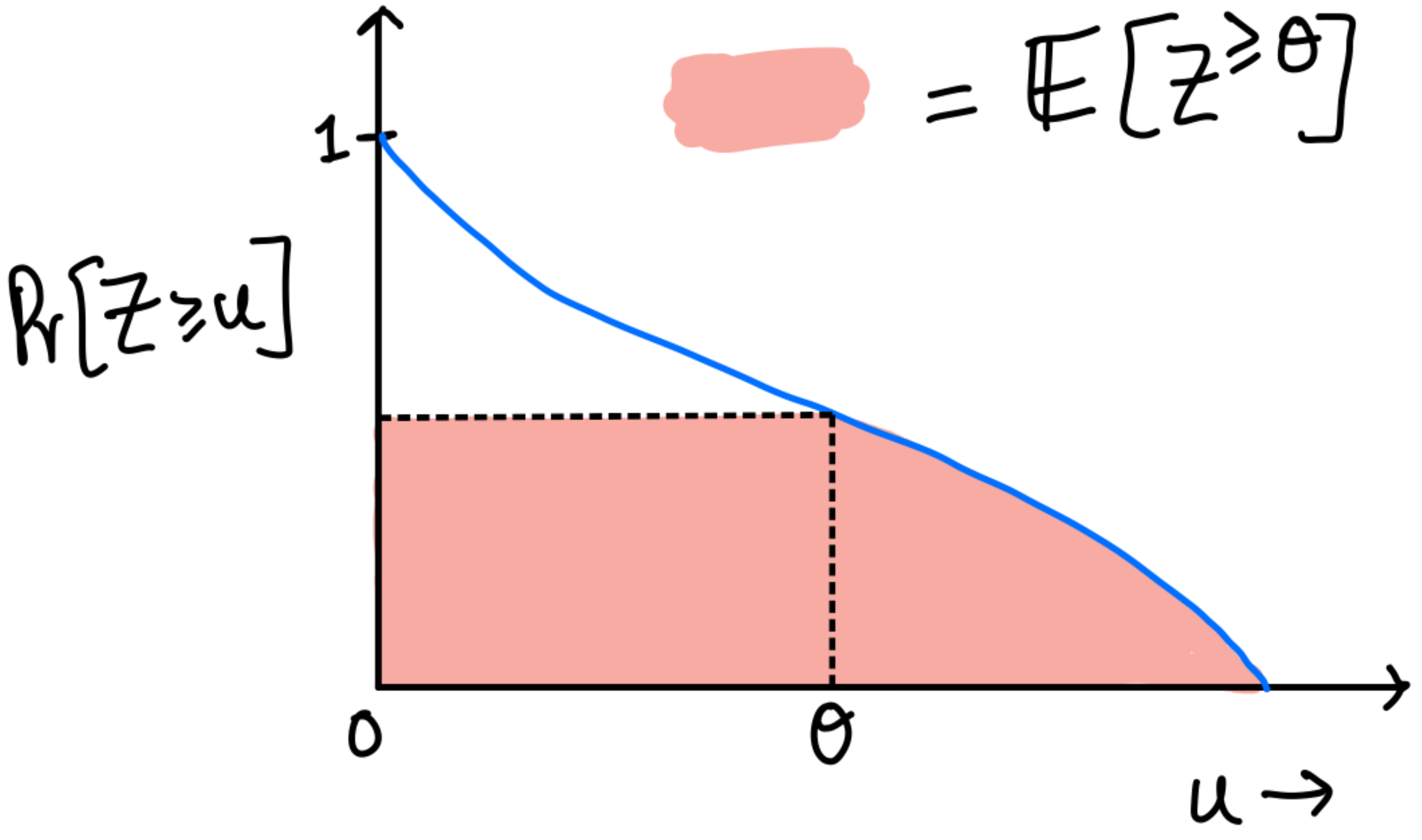
Based on exceptional variables....

For nonnegative r.v. Z and scalar $\theta \geq 0$, define *exceptional variable* $Z^{\geq\theta} := \begin{cases} Z & \text{if } Z \geq \theta \\ 0 & \text{otherwise} \end{cases}$

First Proxy Function for Stochastic Top_l Optimization

Based on exceptional variables....

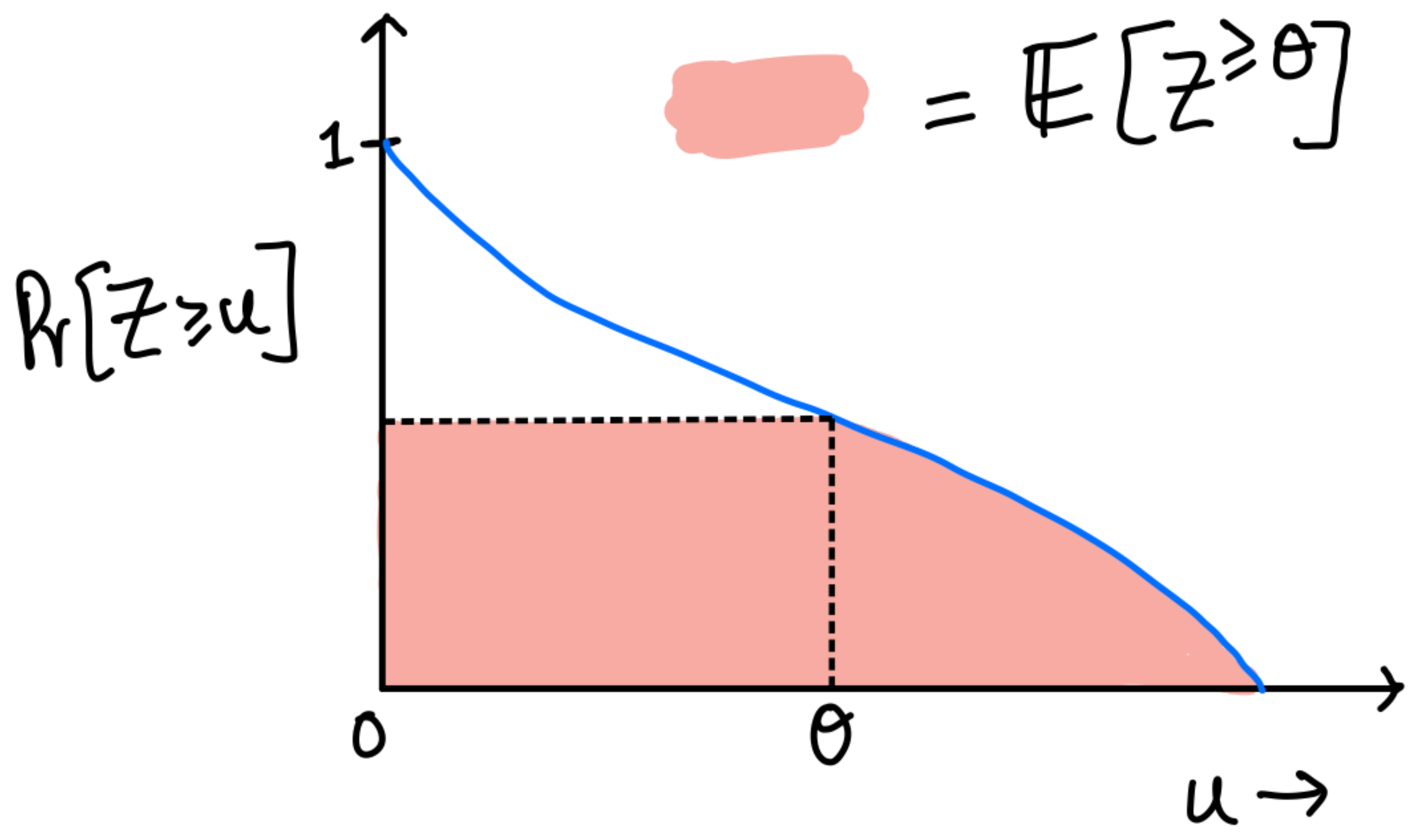
For nonnegative r.v. Z and scalar $\theta \geq 0$, define *exceptional variable* $Z^{\geq \theta} := \begin{cases} Z & \textit{if } Z \geq \theta \\ 0 & \textit{otherwise} \end{cases}$



First Proxy Function for Stochastic Top_l Optimization

Based on exceptional variables....

For nonnegative r.v. Z and scalar $\theta \geq 0$, define *exceptional variable* $Z^{\geq\theta} := \begin{cases} Z & \text{if } Z \geq \theta \\ 0 & \text{otherwise} \end{cases}$



Theorem [IS20]: For product distribution Y,

$$\sum_{i\in[m]} \mathbb{E}[Y_i^{\geq\theta}] > \ell\theta \implies \mathbb{E}[\text{Top}_\ell(Y)] > \ell\theta/2$$

$$\sum_{i\in[m]} \mathbb{E}[Y_i^{\geq\theta}] \leq \ell\theta \implies \mathbb{E}[\text{Top}_\ell(Y)] \leq 2\ell\theta$$

Second Proxy Function for Stochastic Top_l Optimization

Second Proxy Function for Stochastic Top_l Optimization

Based on a "fractile" viewpoint....

Define $\tau_\ell(Y) :=$ smallest θ s.t. $\sum_{i \in [m]} \mathbf{Pr}[Y_i > \theta] < \ell$

for deterministic $y \in \mathbb{R}^m_{\geq 0}$, $\tau_\ell(y) = y^\downarrow_\ell$

in general, $\mathbb{E}[Y^\downarrow_\ell] \geq \tau_\ell(Y)/2$

Second Proxy Function for Stochastic Top_l Optimization

Based on a "fractile" viewpoint....

Define $\tau_\ell(Y)$:= smallest θ s.t. $\sum_{i \in [m]} \mathbf{Pr}[Y_i > \theta] < \ell$

for deterministic $y \in \mathbb{R}^m_{\geq 0}$, $\tau_\ell(y) = y^\downarrow_\ell$

in general, $\mathbb{E}[Y^\downarrow_\ell] \geq \tau_\ell(Y)/2$

Define $\gamma_\ell(Y) := \ell \tau_\ell + \sum_{i \in [m]} \mathbb{E}[\max(0, Y_i - \tau_\ell)]$

for deterministic $y \in \mathbb{R}^m_{\geq 0}$, $\gamma_\ell(y) = \text{Top}_\ell(y)$

Second Proxy Function for Stochastic Top_l Optimization

Based on a "fractile" viewpoint....

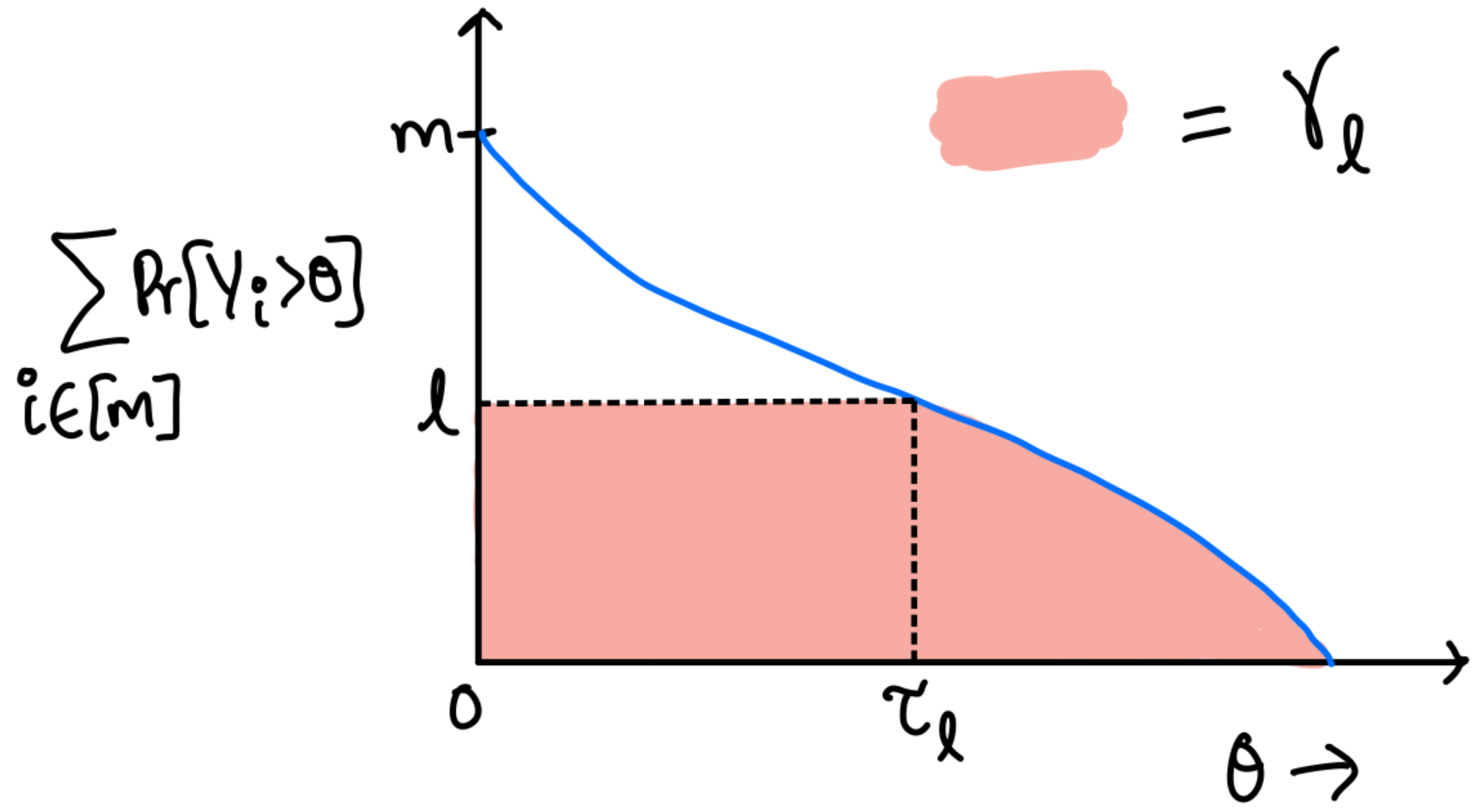
Define $\tau_\ell(Y) :=$ smallest θ s.t. $\sum_{i \in [m]} \mathbf{Pr}[Y_i > \theta] < \ell$

for deterministic $y \in \mathbb{R}^m_{\geq 0}$, $\tau_\ell(y) = y^\downarrow_\ell$

in general, $\mathbb{E}[Y^\downarrow_\ell] \geq \tau_\ell(Y)/2$

Define $\gamma_\ell(Y) := \ell\tau_\ell + \sum_{i \in [m]} \mathbb{E}[\max(0, Y_i - \tau_\ell)]$

for deterministic $y \in \mathbb{R}^m_{\geq 0}$, $\gamma_\ell(y) = \text{Top}_\ell(y)$



Second Proxy Function for Stochastic Top_l Optimization

Based on a "fractile" viewpoint....

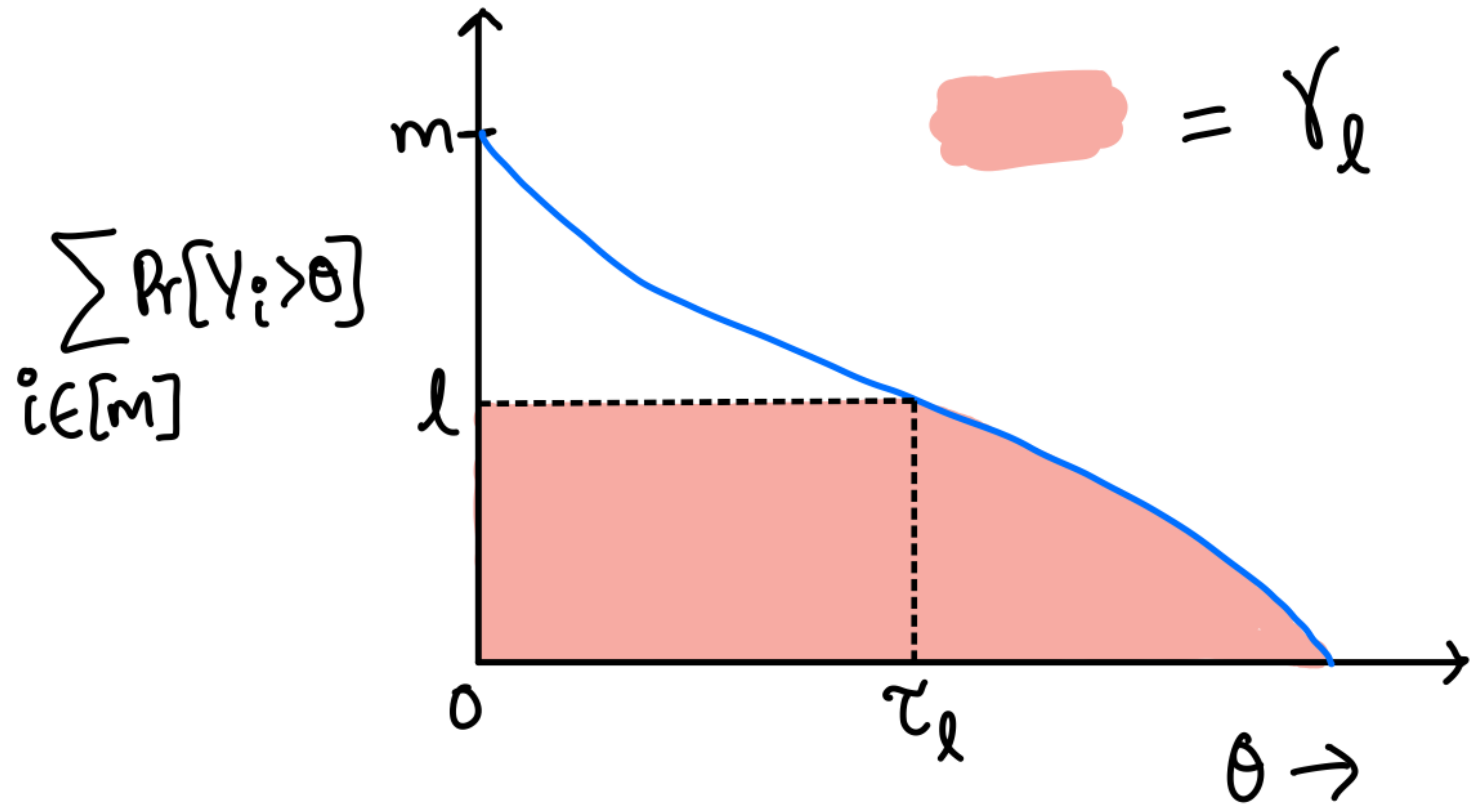
Define $\tau_\ell(Y) :=$ smallest θ s.t. $\sum_{i \in [m]} \mathbf{Pr}[Y_i > \theta] < \ell$

for deterministic $y \in \mathbb{R}^m_{\geq 0}$, $\tau_\ell(y) = y^{\downarrow}_\ell$

in general, $\mathbb{E}[Y^{\downarrow}_\ell] \geq \tau_\ell(Y)/2$

Define $\gamma_\ell(Y) := \ell \tau_\ell + \sum_{i \in [m]} \mathbb{E}[\max(0, Y_i - \tau_\ell)]$

for deterministic $y \in \mathbb{R}^m_{\geq 0}$, $\gamma_\ell(y) = \text{Top}_\ell(y)$



Theorem [IS20]: For product distribution Y,

$$\frac{\gamma_\ell}{2} \leq \mathbb{E}[\text{Top}_\ell(Y)] \leq \gamma_\ell(Y)$$

i.e., $\ell\theta + \sum_{i \in [m]} \mathbb{E}[\max(0, Y_i - \theta)] \approx \mathbb{E}[\text{Top}_\ell(Y)]$ when $\theta \approx \tau_\ell$

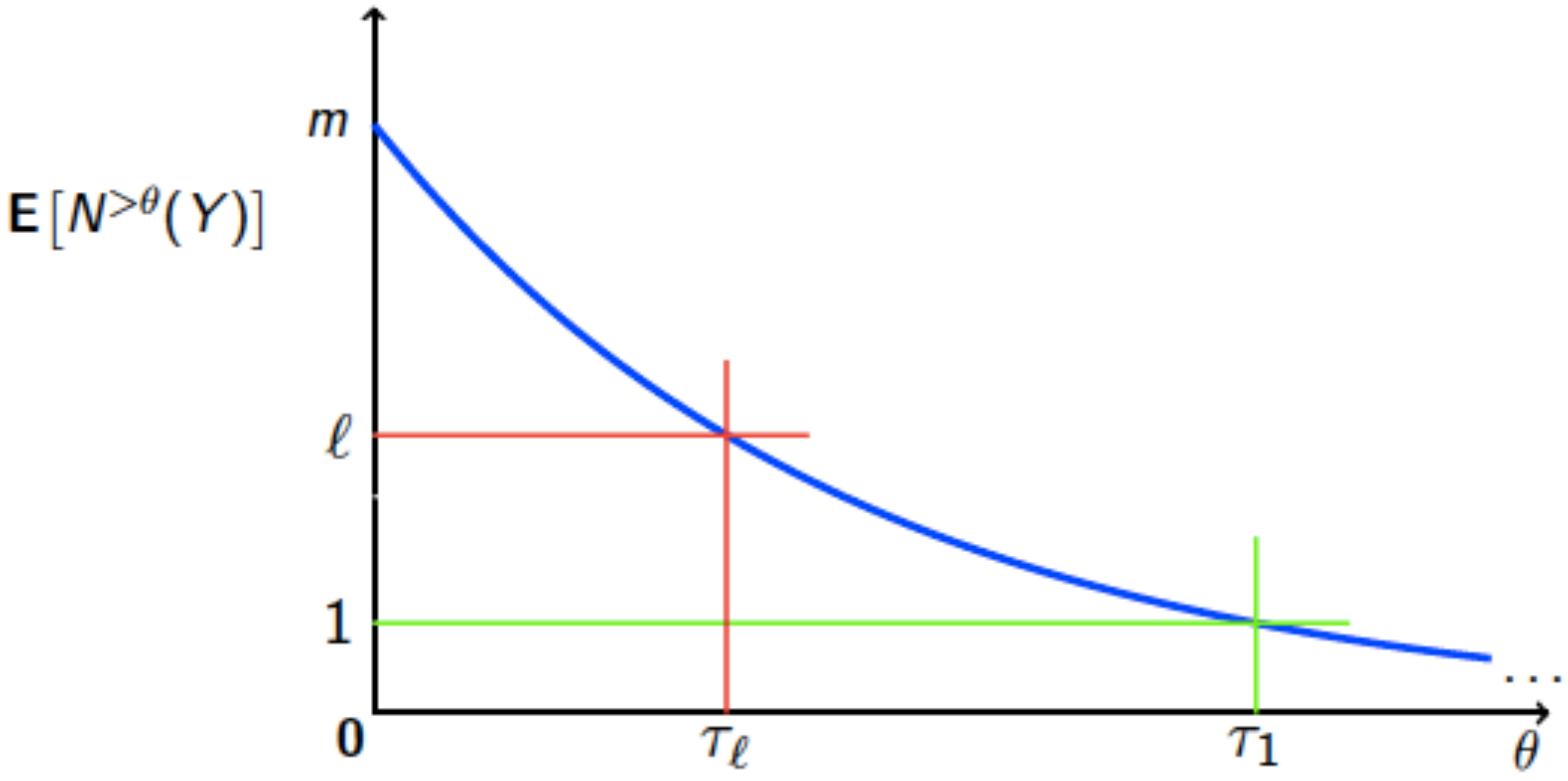
Pictorial Proof of $\mathbb{E}[f(Y)] \leq O(1) \cdot f(\mathbb{E}[Y^{\downarrow}])$

Pictorial Proof of $\mathbb{E}[f(Y)] \leq O(1) \cdot f(\mathbb{E}[Y^{\downarrow}])$

▷ Consider integer r.v. $N^{>\theta}(Y)$:= # of coordinates of Y that are $> \theta$

Pictorial Proof of $\mathbb{E}[f(Y)] \le O(1) \cdot f(\mathbb{E}[Y^{\downarrow}])$

- ▷ Consider integer r.v. $N^{>\theta}(Y) :=$ # of coordinates of Y that are $> \theta$
- ▷ Plot *expected histogram curve* $\mathbb{E}[N^{>\theta}(Y)] = \sum_{i \in [m]} \Pr[Y_i > \theta]$ vs. θ



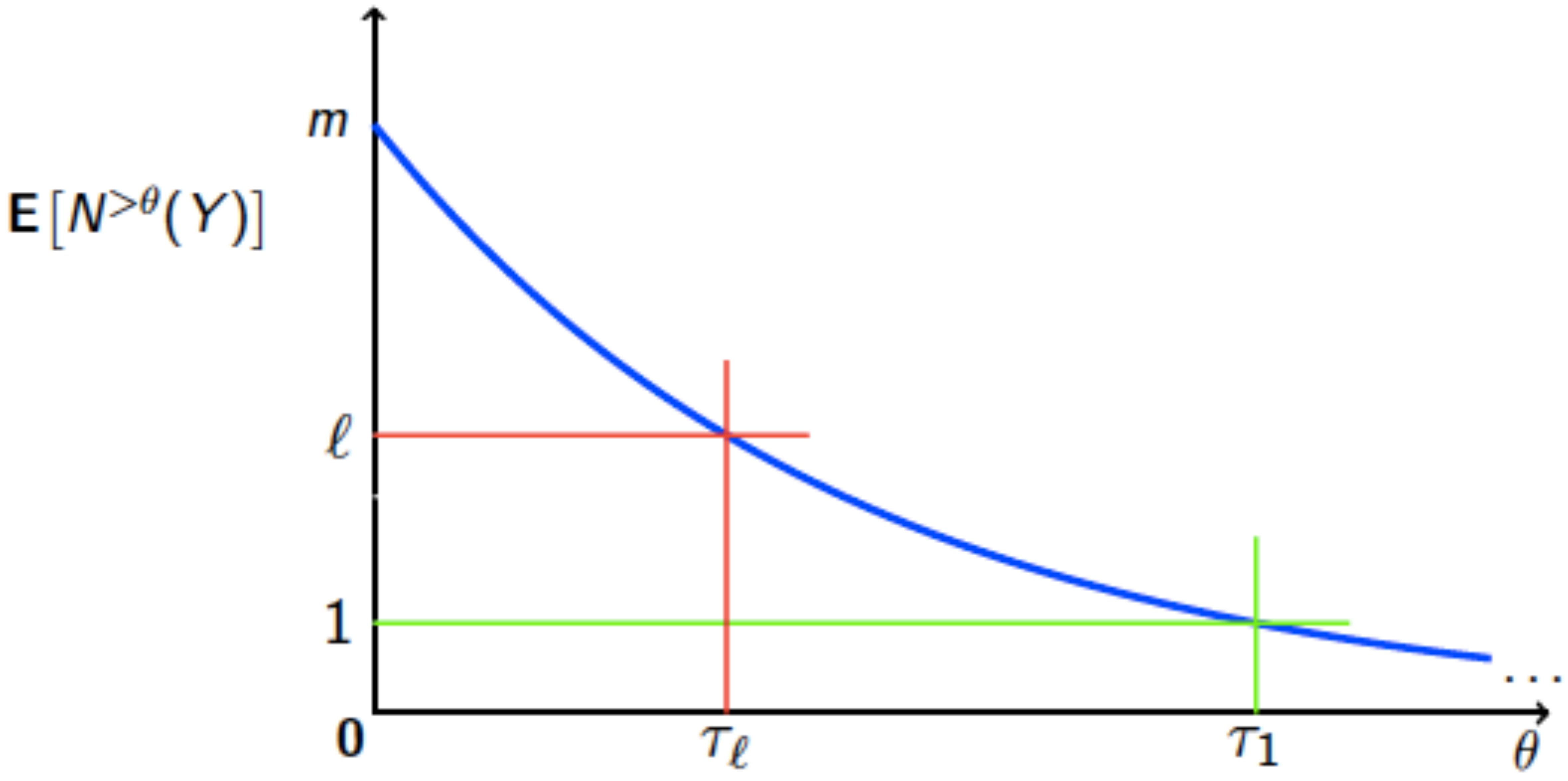
Expected Histogram Curve

Pictorial Proof of $\mathbb{E}[f(Y)] \leq O(1) \cdot f(\mathbb{E}[Y^{\downarrow}])$

▷ Consider integer r.v. $N^{>\theta}(Y) :=$ # of coordinates of Y that are $> \theta$

▷ Plot *expected histogram curve* $\mathbb{E}[N^{>\theta}(Y)] = \sum_{i \in [m]} \Pr[Y_i > \theta]$ vs. θ

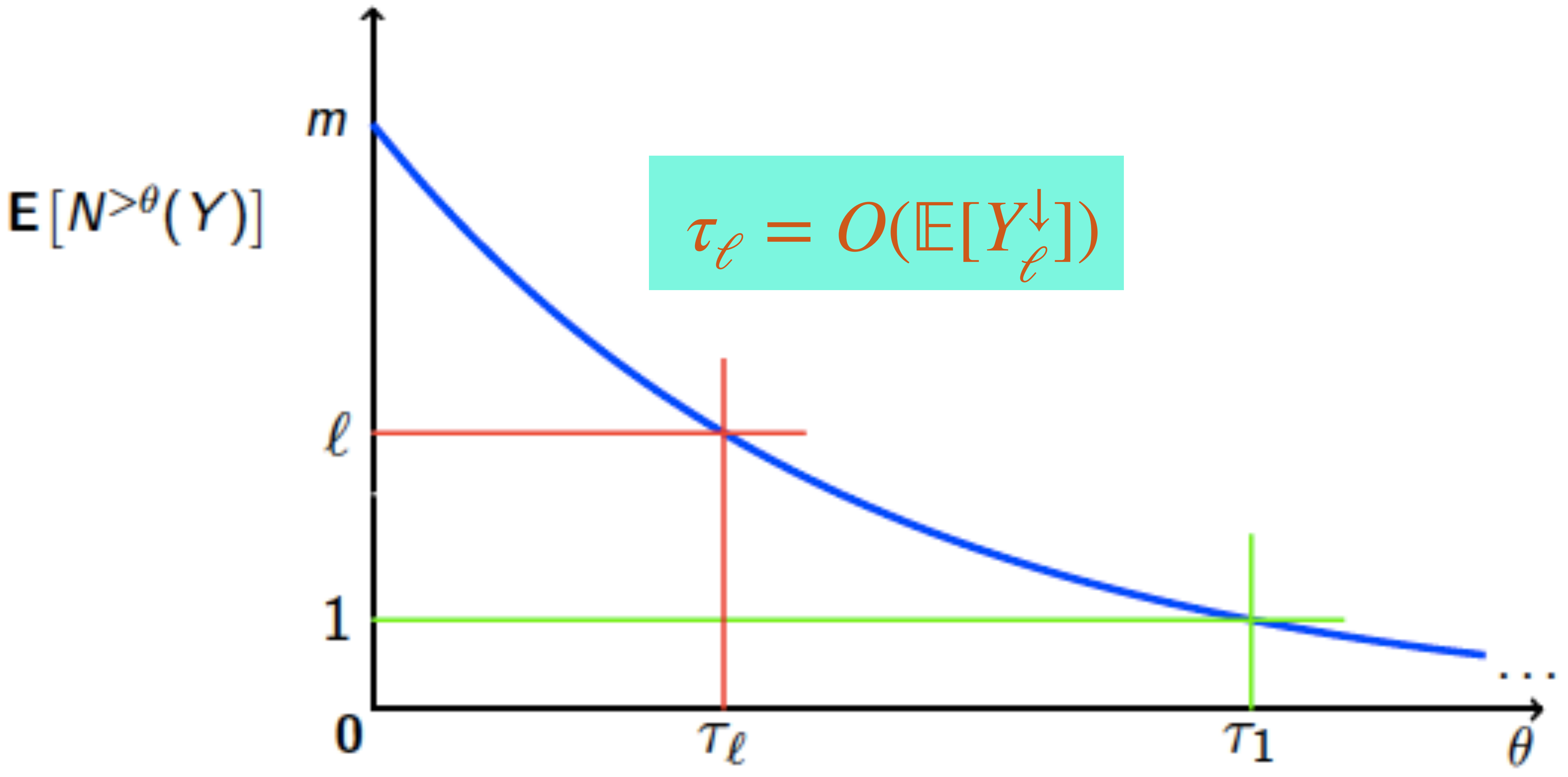
▷ Define τ_ℓ such that $\mathbb{E}[N^{>\tau_\ell}(Y)] = \ell$ holds



Expected Histogram Curve

Pictorial Proof of $\mathbb{E}[f(Y)] \leq O(1) \cdot f(\mathbb{E}[Y^{\downarrow}])$

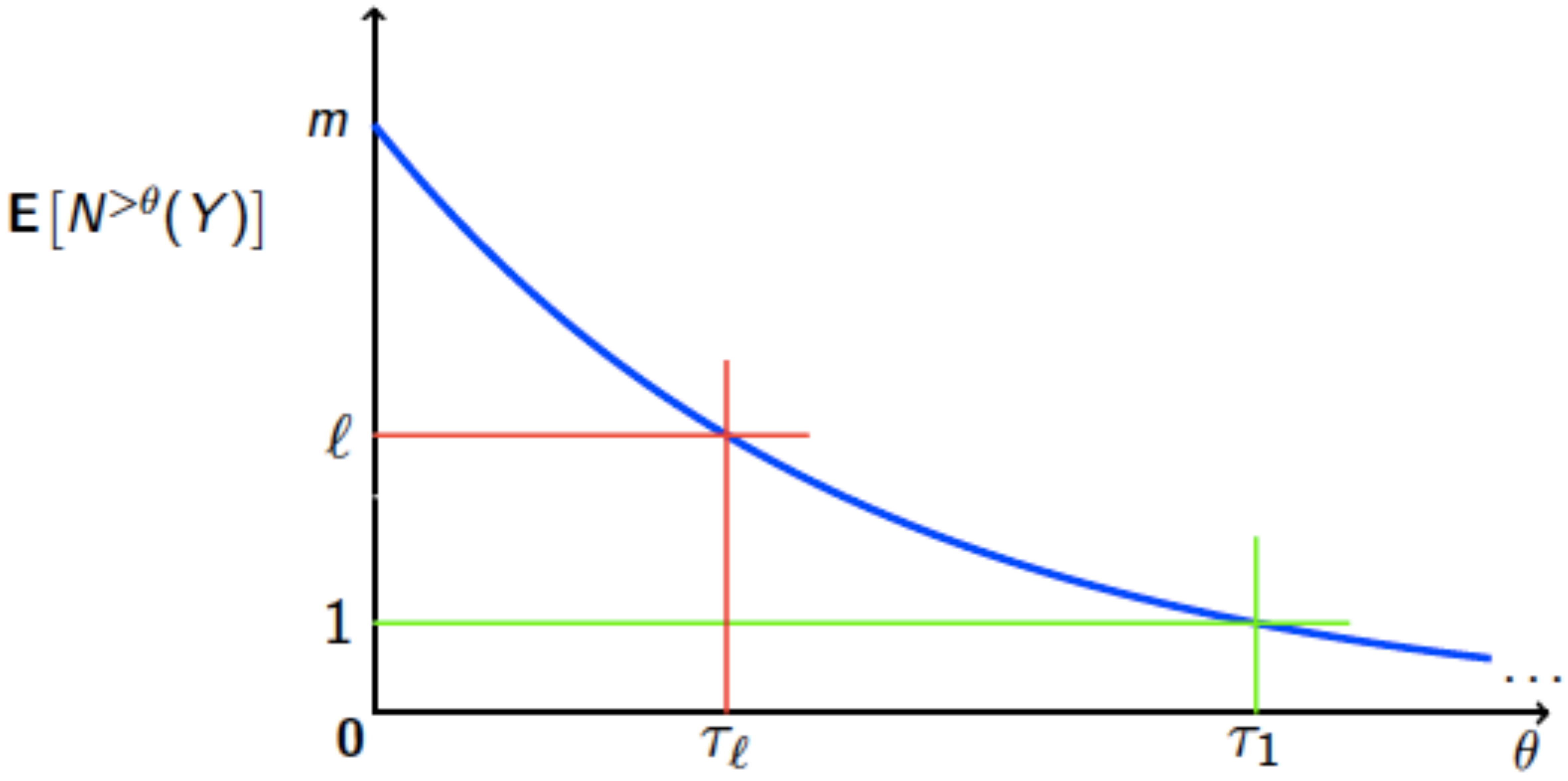
- ▷ Consider integer r.v. $N^{>\theta}(Y) :=$ # of coordinates of Y that are $> \theta$
- ▷ Plot *expected histogram curve* $\mathbb{E}[N^{>\theta}(Y)] = \sum_{i \in [m]} \Pr[Y_i > \theta]$ vs. θ
- ▷ Define τ_ℓ such that $\mathbb{E}[N^{>\tau_\ell}(Y)] = \ell$ holds



Expected Histogram Curve

Pictorial Proof of $\mathbb{E}[f(Y)] \leq O(1) \cdot f(\mathbb{E}[Y^{\downarrow}])$

- ▷ Consider integer r.v. $N^{>\theta}(Y) :=$ # of coordinates of Y that are $> \theta$
- ▷ Plot *expected histogram curve* $\mathbb{E}[N^{>\theta}(Y)] = \sum_{i \in [m]} \Pr[Y_i > \theta]$ vs. θ
- ▷ Define τ_ℓ such that $\mathbb{E}[N^{>\tau_\ell}(Y)] = \ell$ holds



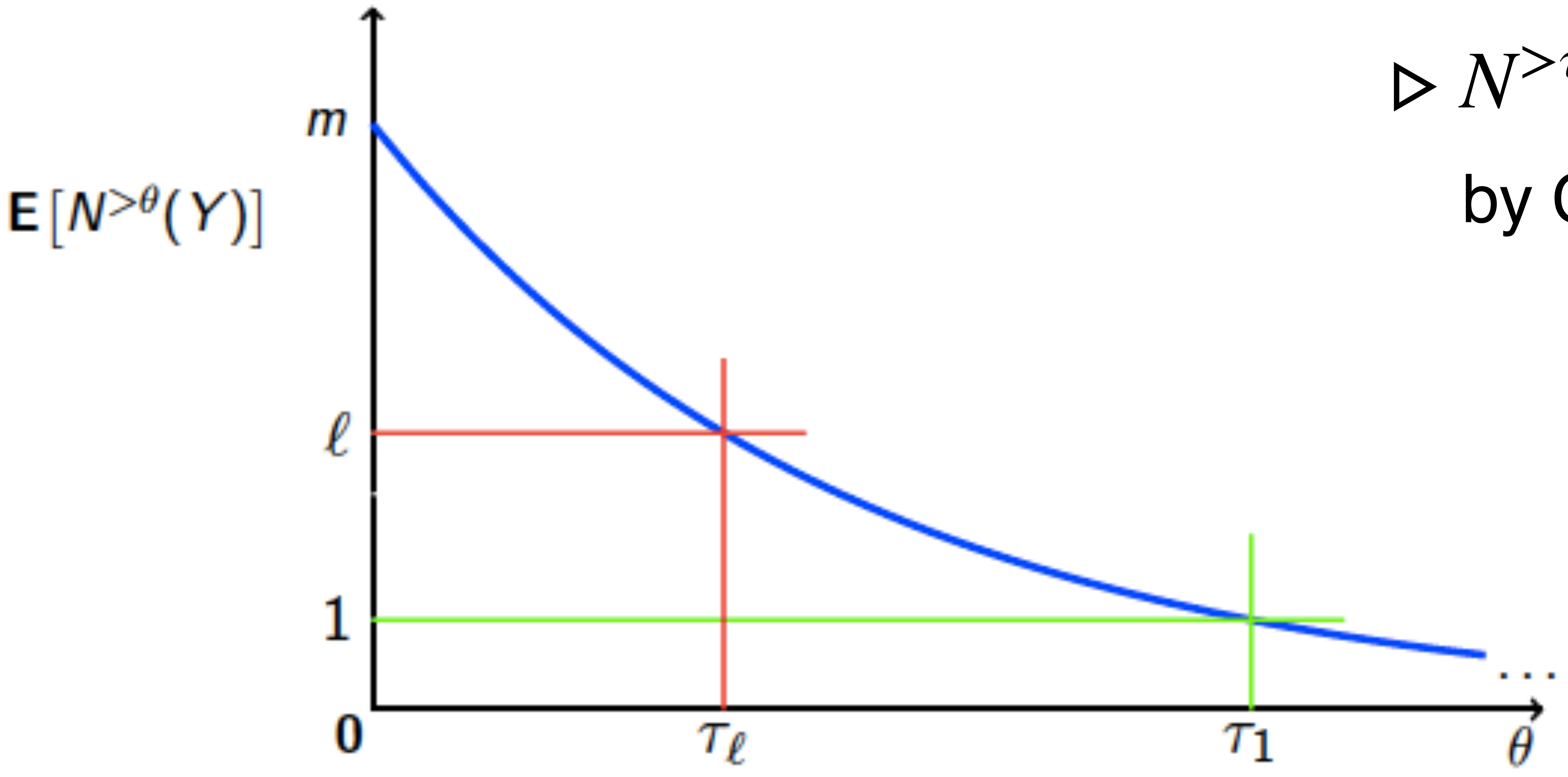
Expected Histogram Curve

Pictorial Proof of $\mathbb{E}[f(Y)] \le O(1) \cdot f(\mathbb{E}[Y^{\downarrow}])$

▷ Consider integer r.v. $N^{>\theta}(Y) :=$ # of coordinates of Y that are $> \theta$

▷ Plot *expected histogram curve* $\mathbb{E}[N^{>\theta}(Y)] = \sum_{i\in[m]} \Pr[Y_i > \theta]$ vs. θ

▷ Define τ_ℓ such that $\mathbb{E}[N^{>\tau_\ell}(Y)] = \ell$ holds

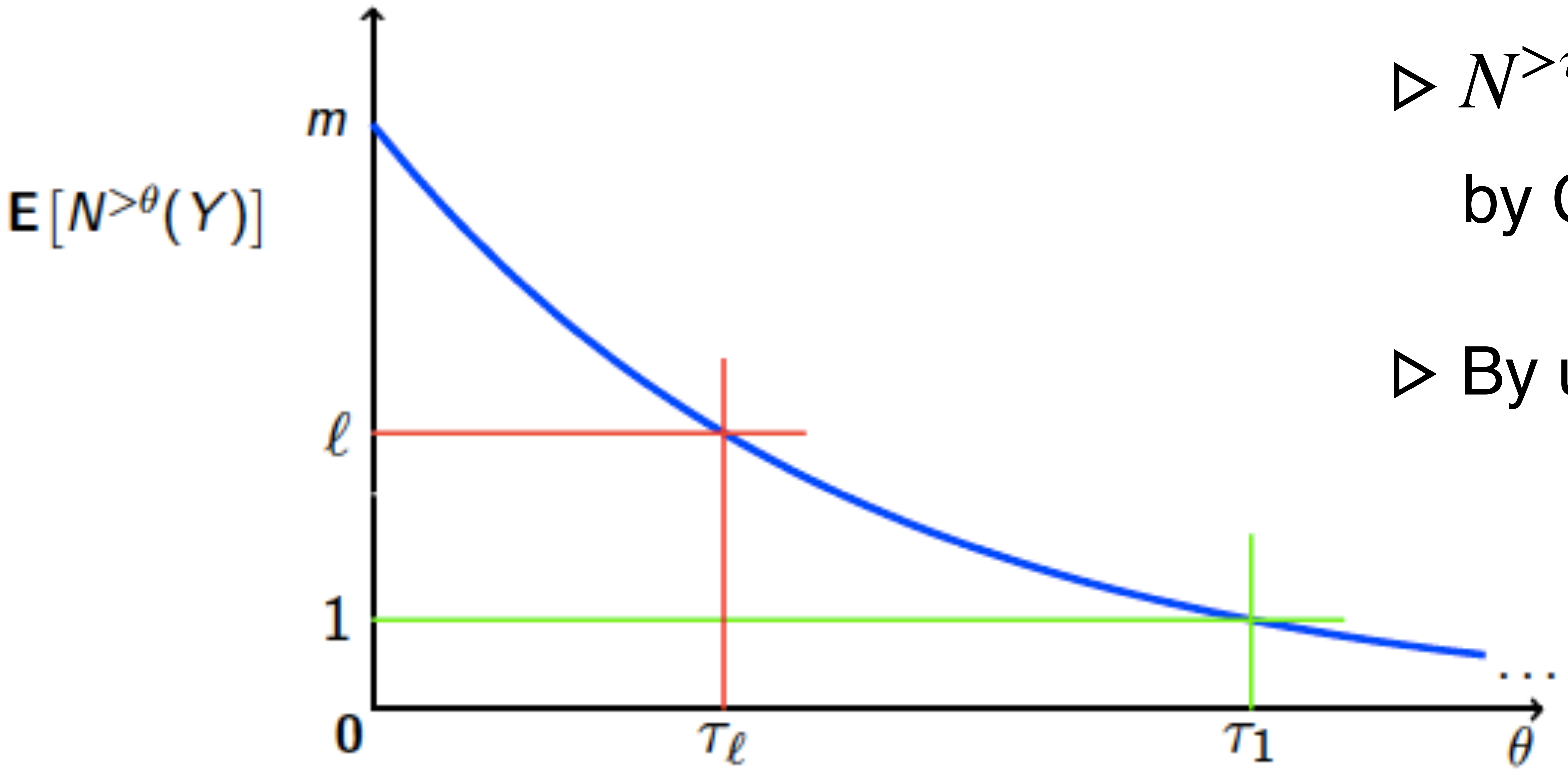


Expected Histogram Curve

▷ $N^{>\tau_\ell}(Y)$ is a sum of independent Bernoullis with mean ℓ, so by Chernoff, $\Pr[Y^{\downarrow}_{\alpha\ell} > \tau_\ell] = O(\exp(-\alpha\ell))$ for large α

Pictorial Proof of $\mathbb{E}[f(Y)] \le O(1) \cdot f(\mathbb{E}[Y^{\downarrow}])$

- ▷ Consider integer r.v. $N^{>\theta}(Y) :=$ # of coordinates of Y that are $> \theta$
- ▷ Plot *expected histogram curve* $\mathbb{E}[N^{>\theta}(Y)] = \sum_{i\in[m]} \Pr[Y_i > \theta]$ vs. θ
- ▷ Define τ_ℓ such that $\mathbb{E}[N^{>\tau_\ell}(Y)] = \ell$ holds



Expected Histogram Curve

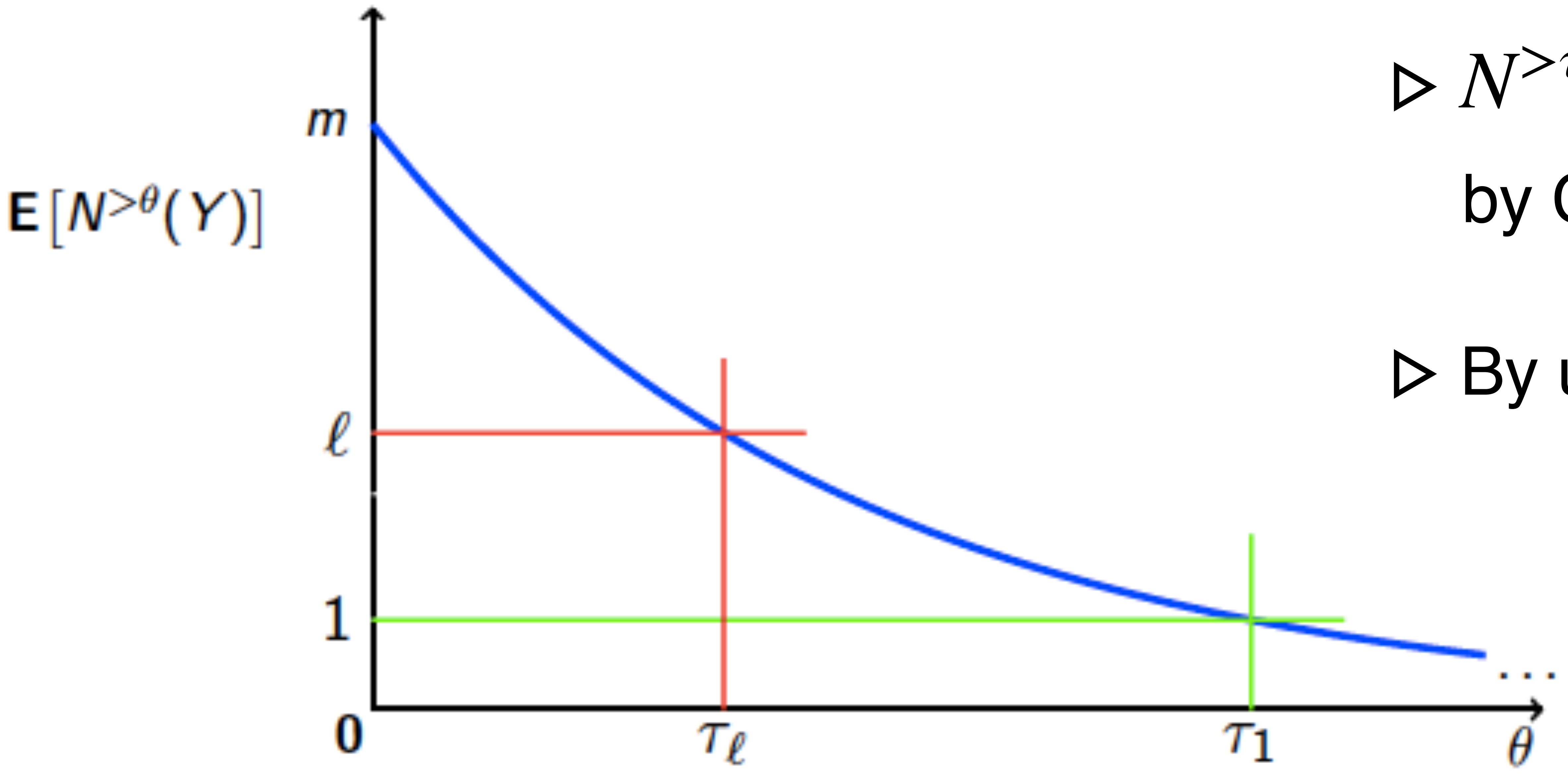
- ▷ $N^{>\tau_\ell}(Y)$ is a sum of independent Bernoullis with mean ℓ, so by Chernoff, $\Pr[Y^{\downarrow}_{\alpha\ell} > \tau_\ell] = O(\exp(-\alpha\ell))$ for large α
- ▷ By union-bound, $\Pr\left[Y^{\downarrow}_{\alpha\ell} \le \tau_\ell \,\forall \ell \in [m]\right] \ge 1 - O(\exp(-\alpha))$

Pictorial Proof of $\mathbb{E}[f(Y)] \le O(1) \cdot f(\mathbb{E}[Y^{\downarrow}])$

▷ Consider integer r.v. $N^{>\theta}(Y) :=$ # of coordinates of Y that are $> \theta$

▷ Plot *expected histogram curve* $\mathbb{E}[N^{>\theta}(Y)] = \sum_{i\in[m]} \Pr[Y_i > \theta]$ vs. θ

▷ Define τ_ℓ such that $\mathbb{E}[N^{>\tau_\ell}(Y)] = \ell$ holds



Expected Histogram Curve

▷ $N^{>\tau_\ell}(Y)$ is a sum of independent Bernoullis with mean ℓ, so by Chernoff, $\Pr[Y^{\downarrow}_{\alpha\ell} > \tau_\ell] = O(\exp(-\alpha\ell))$ for large α

▷ By union-bound, $\Pr\left[Y^{\downarrow}_{\alpha\ell} \le \tau_\ell \, \forall \ell \in [m]\right] \ge 1 - O(\exp(-\alpha))$

▷ By properties of f,

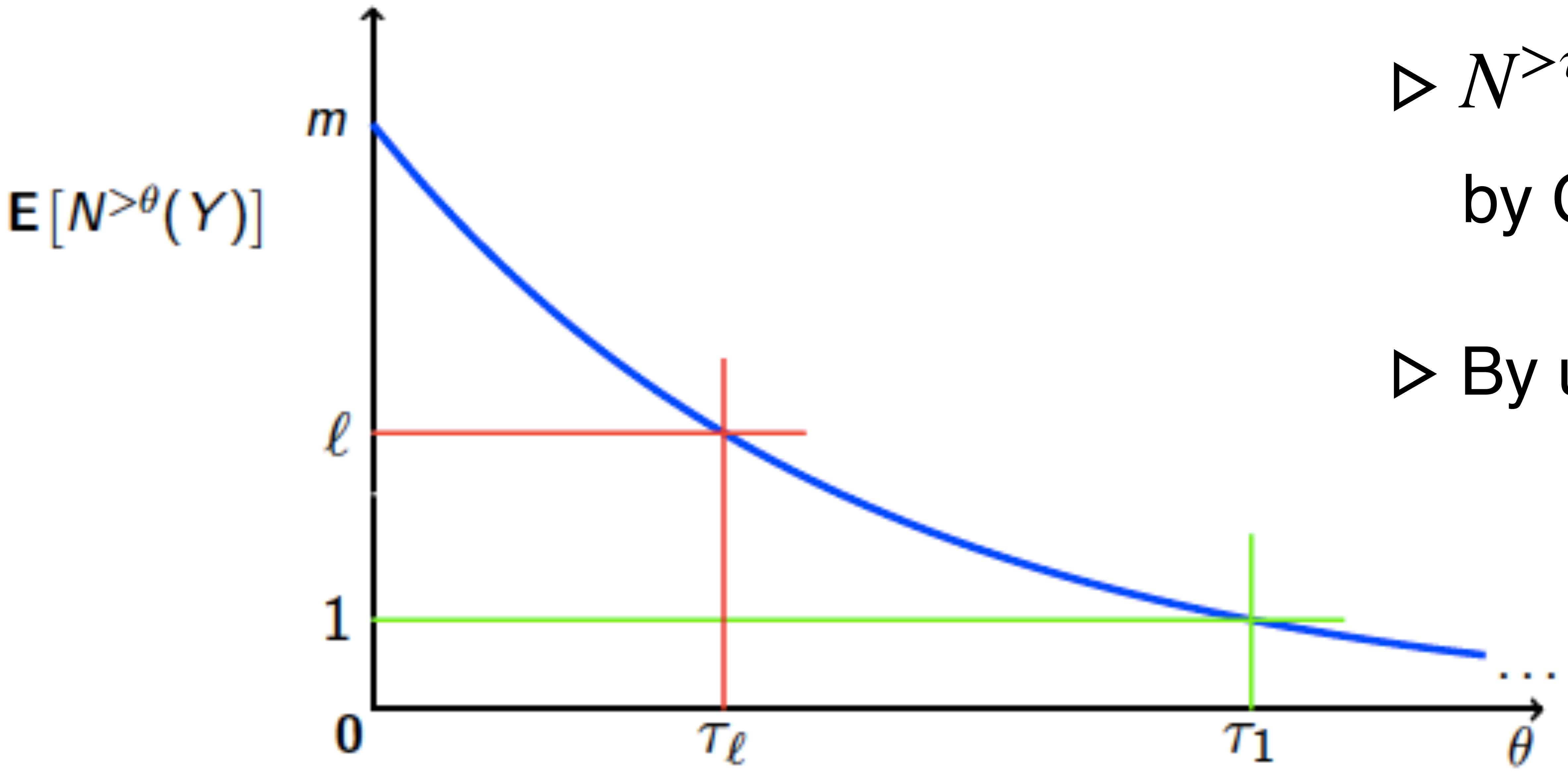
$\Pr[\, f(Y) > \alpha \cdot f(\tau_1, \tau_2, \ldots, \tau_m) \,] \le O(\exp(-\alpha))$ ■

Pictorial Proof of $\mathbb{E}[f(Y)] \le O(1) \cdot f(\mathbb{E}[Y^{\downarrow}])$

▷ Consider integer r.v. $N^{>\theta}(Y)$:= # of coordinates of Y that are $> \theta$

▷ Plot *expected histogram curve* $\mathbb{E}[N^{>\theta}(Y)] = \sum_{i \in [m]} \Pr[Y_i > \theta]$ vs. θ

▷ Define τ_ℓ such that $\mathbb{E}[N^{>\tau_\ell}(Y)] = \ell$ holds



Expected Histogram Curve

▷ $N^{>\tau_\ell}(Y)$ is a sum of independent Bernoullis with mean ℓ, so by Chernoff, $\Pr[Y^{\downarrow}_{\alpha\ell} > \tau_\ell] = O(\exp(-\alpha\ell))$ for large α

▷ By union-bound, $\Pr\left[Y^{\downarrow}_{\alpha\ell} \le \tau_\ell \,\forall \ell \in [m]\right] \ge 1 - O(\exp(-\alpha))$

▷ By properties of f,

$\Pr[\, f(Y) > \alpha \cdot f(\tau_1, \tau_2, \ldots, \tau_m)\,] \le O(\exp(-\alpha))$ ■

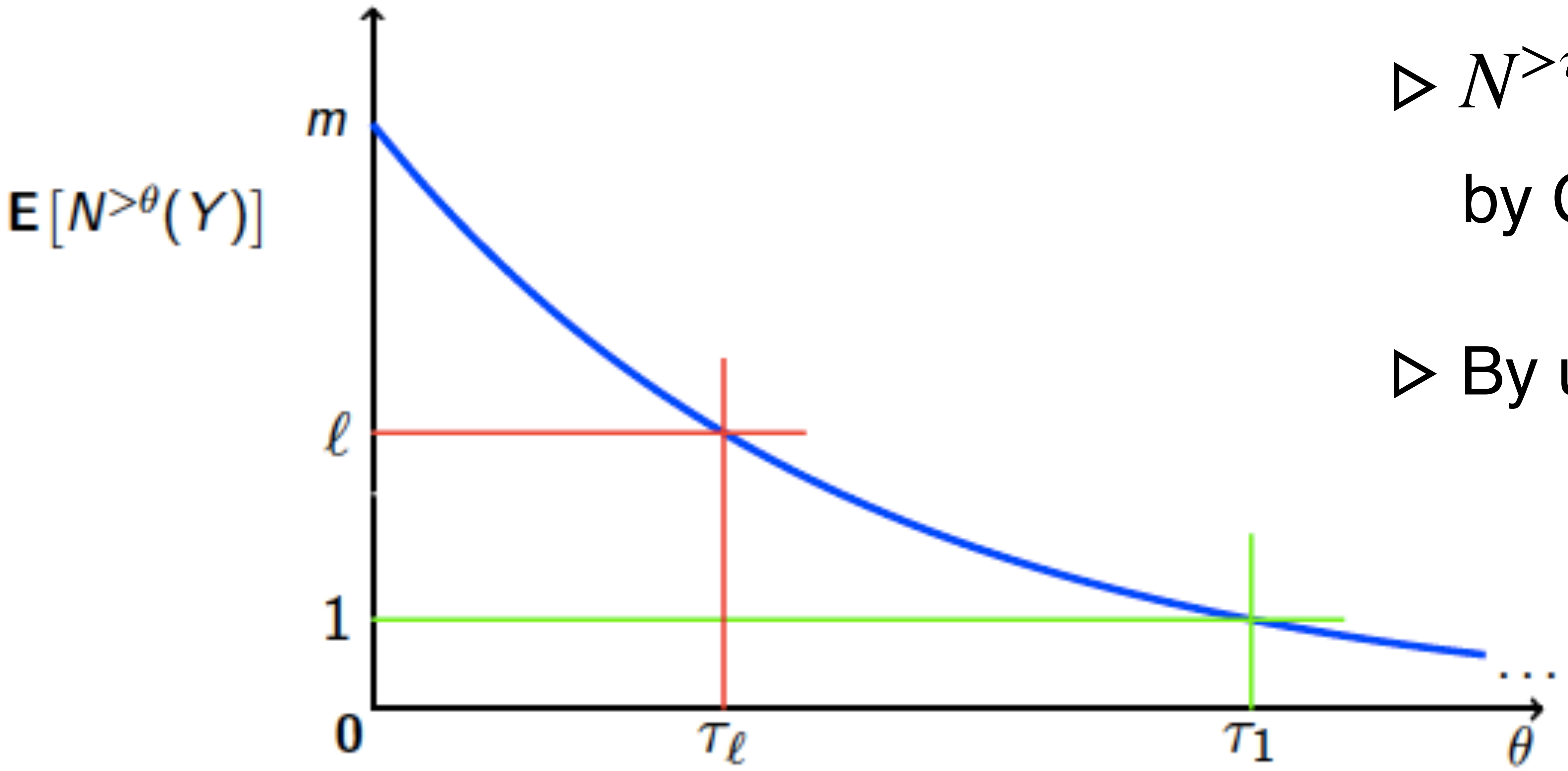
Theorem: $\mathbb{E}[f(Y)] = \Theta\left(f\left(\tau_1 + \sum_{i \in [m]} \mathbb{E}[\max(0, Y_i - \tau_1)], \tau_2, \ldots, \tau_m\right)\right)$

Pictorial Proof of $\mathbb{E}[f(Y)] \le O(1)\cdot f(\mathbb{E}[Y^{\downarrow}])$

▷ Consider integer r.v. $N^{>\theta}(Y)$:= # of coordinates of Y that are $> \theta$

▷ Plot *expected histogram curve* $\mathbb{E}[N^{>\theta}(Y)] = \sum_{i\in[m]} \Pr[Y_i > \theta]$ vs. θ

▷ Define τ_ℓ such that $\mathbb{E}[N^{>\tau_\ell}(Y)] = \ell$ holds



Expected Histogram Curve

▷ $N^{>\tau_\ell}(Y)$ is a sum of independent Bernoullis with mean ℓ, so by Chernoff, $\Pr[Y^{\downarrow}_{\alpha\ell} > \tau_\ell] = O(\exp(-\alpha\ell))$ for large α

▷ By union-bound, $\Pr\left[Y^{\downarrow}_{\alpha\ell} \le \tau_\ell \,\forall \ell \in [m]\right] \ge 1 - O(\exp(-\alpha))$

▷ By properties of f,

$\Pr[\, f(Y) > \alpha \cdot f(\tau_1, \tau_2, \ldots, \tau_m)\,] \le O(\exp(-\alpha))$ ■

Theorem: $\mathbb{E}[f(Y)] = \Theta\Big(f\big(\tau_1 + \sum_{i\in[m]} \mathbb{E}[\max(0, Y_i - \tau_1)], \tau_2, \ldots, \tau_m\big)\Big)$

Applications

Approximation Strategy for StochNormTree

Approximation Strategy for StochNormTree

Given: $G = (V, E)$, stochastic weights $\{X_e\}_{e \in E}$, norm f

Find: Spanning tree $T = \{e_1, \ldots, e_{n-1}\}$ minimizing $\mathbb{E}[f(Y^T)] = \mathbb{E}\left[f(X_{e_1}, \ldots, X_{e_{n-1}})\right]$

Approximation Strategy for StochNormTree

Given: $G = (V, E)$, stochastic weights $\{X_e\}_{e \in E}$, norm f

Find: Spanning tree $T = \{e_1, \ldots, e_{n-1}\}$ minimizing $\mathbb{E}[\, f(Y^T)] = \mathbb{E}\left[f(X_{e_1}, \ldots, X_{e_{n-1}})\right]$

Recall ... $\mathbb{E}[f(Y)] = \Theta\left(f\left(\tau_1 + \sum_{e \in Y} \mathbb{E}[\max(0, Y_e - \tau_1)], \tau_2, \ldots, \tau_{n-1}\right)\right)$

Approximation Strategy for StochNormTree

Given: $G = (V, E)$, stochastic weights $\{X_e\}_{e \in E}$, norm f

Find: Spanning tree $T = \{e_1, \ldots, e_{n-1}\}$ minimizing $\mathbb{E}[\, f(Y^T)] = \mathbb{E}\left[f(X_{e_1}, \ldots, X_{e_{n-1}})\right]$

Recall ... $$\mathbb{E}[f(Y)] = \Theta\left(f\left(\tau_1 + \sum_{e \in Y} \mathbb{E}[\max(0, Y_e - \tau_1)], \tau_2, \ldots, \tau_{n-1}\right)\right)$$

Guess the right τ_ℓ-statistics

W := cost vector of optimal tree

Guess $t_\ell \approx \tau_\ell(W)$ for powers-of-2 ℓs

Approximation Strategy for StochNormTree

Given: $G = (V, E)$, stochastic weights $\{X_e\}_{e \in E}$, norm f

Find: Spanning tree $T = \{e_1, \dots, e_{n-1}\}$ minimizing $\mathbb{E}[\, f(Y^T)] = \mathbb{E}\left[f(X_{e_1}, \dots, X_{e_{n-1}})\right]$

Recall ... $\mathbb{E}[f(Y)] = \Theta\Big(f\big(\tau_1 + \sum_{e \in Y} \mathbb{E}[\max(0, Y_e - \tau_1)], \tau_2, \dots, \tau_{n-1}\big)\Big)$

Guess the right τ_ℓ-statistics

W := cost vector of optimal tree

Guess $t_\ell \approx \tau_\ell(W)$ for powers-of-2 ℓs

→

Solve an LP

Consider an optimal fractional spanning tree for:

$$\min \sum_{e \in E} \mathbb{E}[\max(0, X_e - t_1)] \cdot z_e$$

$$\text{s.t.} \quad \sum_{e \in E} \Pr[X_e > t_\ell] \cdot z_e \le \ell \quad \forall \ell = 1,2,4,\dots$$

$z \in$ spanning tree polytope

Approximation Strategy for StochNormTree

Given: $G = (V, E)$, stochastic weights $\{X_e\}_{e \in E}$, norm f

Find: Spanning tree $T = \{e_1, \ldots, e_{n-1}\}$ minimizing $\mathbb{E}[f(Y^T)] = \mathbb{E}\left[f(X_{e_1}, \ldots, X_{e_{n-1}})\right]$

Recall … $$\mathbb{E}[f(Y)] = \Theta\left(f\left(\tau_1 + \sum_{e \in Y} \mathbb{E}[\max(0, Y_e - \tau_1)], \tau_2, \ldots, \tau_{n-1}\right)\right)$$

Guess the right τ_ℓ-statistics

$W :=$ cost vector of optimal tree

Guess $t_\ell \approx \tau_\ell(W)$ for powers-of-2 ℓs

→

Solve an LP

Consider an optimal fractional spanning tree for:

$$\begin{aligned} \min \quad & \sum_{e \in E} \mathbb{E}[\max(0, X_e - t_1)] \cdot z_e \\ \text{s.t.} \quad & \sum_{e \in E} \Pr[X_e > t_\ell] \cdot z_e \le \ell \quad \forall \ell = 1,2,4,\ldots \\ & z \in \text{spanning tree polytope} \end{aligned}$$

Note: Column sums in the LP are bounded by $O(1)$ after normalising each RHS to 1

Approximation Strategy for StochNormTree

Given: $G = (V, E)$, stochastic weights $\{X_e\}_{e \in E}$, norm f

Find: Spanning tree $T = \{e_1, \ldots, e_{n-1}\}$ minimizing $\mathbb{E}[f(Y^T)] = \mathbb{E}\left[f(X_{e_1}, \ldots, X_{e_{n-1}})\right]$

Recall ... $\mathbb{E}[f(Y)] = \Theta\left(f\left(\tau_1 + \sum_{e \in Y} \mathbb{E}[\max(0, Y_e - \tau_1)], \tau_2, \ldots, \tau_{n-1}\right)\right)$

Guess the right τ_ℓ-statistics

$W :=$ cost vector of optimal tree

Guess $t_\ell \approx \tau_\ell(W)$ for powers-of-2 ℓs

Solve an LP

Consider an optimal fractional spanning tree for:

$$\min \sum_{e \in E} \mathbb{E}[\max(0, X_e - t_1)] \cdot z_e$$

$$\text{s.t.} \quad \sum_{e \in E} \Pr[X_e > t_\ell] \cdot z_e \leq \ell \quad \forall \ell = 1,2,4,\ldots$$

$z \in$ spanning tree polytope

Note: Column sums in the LP are bounded by $O(1)$ after normalising each RHS to 1

Rounding

Using [LOSZ20]'s iterative rounding machinery, we obtain tree T s.t. τ_ℓ statistics of Y^T and W are "comparable"

Our Approximation Guarantees for StochNormTree

Our Approximation Guarantees for StochNormTree

Theorem [IS20]: $O(1)$-approximation algorithm with
arbitrary monotone symmetric norms and
arbitrary edge-weight distributions.

Improved $(2+\varepsilon)$-approximation for Top_ℓ norms

Our Approximation Guarantees for StochNormTree

Theorem [IS20]: $O(1)$-approximation algorithm with
arbitrary monotone symmetric norms and
arbitrary edge-weight distributions.

Improved $(2+\varepsilon)$-approximation for Top_ℓ norms

▷ Same approximation strategy extends to stochastic min-norm versions of:

Matroid Base

Bounded-Degree Spanning Tree (with $O(1)$ violation in degree constraints)

Traveling Salesman Problem

Approximation Algorithms for StochNormLB

Approximation Algorithms for StochNormLB

Given: m unrelated machines, stochastic job-sizes $\{X_{ij}\}_{i\in[m],j\in J}$, norm f

Find: Assignment $\sigma : J \to [m]$ minimizing $\mathbb{E}\left[f(\overrightarrow{\text{load}_\sigma}) \right]$

Approximation Algorithms for StochNormLB

Given: m unrelated machines, stochastic job-sizes $\{X_{ij}\}_{i \in [m], j \in J}$, norm f

Find: Assignment $\sigma : J \to [m]$ minimizing $\mathbb{E}\left[f(\overrightarrow{\text{load}_\sigma}) \right]$

Using our framework, we obtain:

Approximation Algorithms for StochNormLB

Given: m unrelated machines, stochastic job-sizes $\{X_{ij}\}_{i\in[m],j\in J}$, norm f

Find: Assignment $\sigma : J \to [m]$ minimizing $\mathbb{E}\left[f(\overrightarrow{\text{load}_\sigma})\right]$

Using our framework, we obtain:

	Weighted Bernoullis $X_{ij} = \begin{cases} s_{ij} \text{ w.p. } q_{ij} \\ 0 \text{ otherwise} \end{cases}$	Arbitrary Job-Size Distributions
Top_ℓ norms	[redacted]	
Arbitrary Monotone Symmetric Norms		

Approximation Algorithms for StochNormLB

Given: m unrelated machines, stochastic job-sizes $\{X_{ij}\}_{i \in [m], j \in J}$, norm f

Find: Assignment $\sigma : J \to [m]$ minimizing $\mathbb{E}\left[f(\overrightarrow{\text{load}_\sigma}) \right]$

Using our framework, we obtain:

	Weighted Bernoullis $X_{ij} = \begin{cases} s_{ij} \text{ w.p. } q_{ij} \\ 0 \text{ otherwise} \end{cases}$	Arbitrary Job-Size Distributions
Top_ℓ norms	[redacted]	$O(1)$-approximation
Arbitrary Monotone Symmetric Norms		

Approximation Algorithms for StochNormLB

Given: m unrelated machines, stochastic job-sizes $\{X_{ij}\}_{i\in[m], j\in J}$, norm f

Find: Assignment $\sigma : J \to [m]$ minimizing $\mathbb{E}\left[f(\overrightarrow{\text{load}_\sigma})\right]$

Using our framework, we obtain:

	Weighted Bernoullis $X_{ij} = \begin{cases} s_{ij} \text{ w.p. } q_{ij} \\ 0 \text{ otherwise} \end{cases}$	Arbitrary Job-Size Distributions
Top$_\ell$ norms	[redacted]	$O(1)$-approximation Substantial generalization of [KRT00,GKNS18]'s O(1)-approximation for the $\ell = 1$ case
Arbitrary Monotone Symmetric Norms		

Approximation Algorithms for StochNormLB

Given: m unrelated machines, stochastic job-sizes $\{X_{ij}\}_{i \in [m], j \in J}$, norm f

Find: Assignment $\sigma : J \to [m]$ minimizing $\mathbb{E}\left[f(\overrightarrow{\text{load}_\sigma})\right]$

Using our framework, we obtain:

	Weighted Bernoullis $X_{ij} = \begin{cases} s_{ij} \text{ w.p. } q_{ij} \\ 0 \text{ otherwise} \end{cases}$	Arbitrary Job-Size Distributions
Top_ℓ norms	[redacted]	$O(1)$-approximation Substantial generalization of [KRT00,GKNS18]'s O(1)-approximation for the $\ell = 1$ case Incomparable to [Molinaro19]'s $O(1)$-approximation for ℓ_p norms
Arbitrary Monotone Symmetric Norms		

Approximation Algorithms for StochNormLB

Given: m unrelated machines, stochastic job-sizes $\{X_{ij}\}_{i \in [m], j \in J}$, norm f

Find: Assignment $\sigma : J \to [m]$ minimizing $\mathbb{E}\left[f(\overrightarrow{\text{load}_\sigma}) \right]$

Using our framework, we obtain:

	Weighted Bernoullis $X_{ij} = \begin{cases} s_{ij} \text{ w.p. } q_{ij} \\ 0 \text{ otherwise} \end{cases}$	Arbitrary Job-Size Distributions
Top$_\ell$ norms	[redacted]	$O(1)$-approximation Substantial generalization of [KRT00,GKNS18]'s O(1)-approximation for the $\ell = 1$ case Incomparable to [Molinaro19]'s $O(1)$-approximation for ℓ_p norms
Arbitrary Monotone Symmetric Norms	$O(1)$-approximation	

Approximation Algorithms for StochNormLB

Given: m unrelated machines, stochastic job-sizes $\{X_{ij}\}_{i\in[m],j\in J}$, norm f

Find: Assignment $\sigma : J \to [m]$ minimizing $\mathbb{E}\left[f(\overrightarrow{\text{load}_\sigma})\right]$

Using our framework, we obtain:

	Weighted Bernoullis $X_{ij} = \begin{cases} s_{ij} \text{ w.p. } q_{ij} \\ 0 \text{ otherwise} \end{cases}$	Arbitrary Job-Size Distributions
Top$_\ell$ norms	[redacted]	$O(1)$-approximation Substantial generalization of [KRT00,GKNS18]'s O(1)-approximation for the $\ell = 1$ case Incomparable to [Molinaro19]'s $O(1)$-approximation for ℓ_p norms
Arbitrary Monotone Symmetric Norms	$O(1)$-approximation Strictly generalizes [CS19]'s O(1)-approximation for the deterministic setting	

Approximation Algorithms for StochNormLB

Given: m unrelated machines, stochastic job-sizes $\{X_{ij}\}_{i \in [m], j \in J}$, norm f

Find: Assignment $\sigma : J \to [m]$ minimizing $\mathbb{E}\left[f(\overrightarrow{\text{load}_\sigma}) \right]$

Using our framework, we obtain:

	Weighted Bernoullis $X_{ij} = \begin{cases} s_{ij} \text{ w.p. } q_{ij} \\ 0 \text{ otherwise} \end{cases}$	Arbitrary Job-Size Distributions
Top$_\ell$ norms	[redacted]	$O(1)$-approximation Substantial generalization of [KRT00,GKNS18]'s O(1)-approximation for the $\ell = 1$ case Incomparable to [Molinaro19]'s $O(1)$-approximation for ℓ_p norms
Arbitrary Monotone Symmetric Norms	$O(1)$-approximation Strictly generalizes [CS19]'s O(1)-approximation for the deterministic setting	$O\left(\frac{\log \log m}{\log \log \log m}\right)$-approximation

StochNormOpt with Poisson Random Variables

StochNormOpt with Poisson Random Variables

Suppose entries of cost vector Y are Poisson variables

StochNormOpt with Poisson Random Variables

Suppose entries of cost vector Y are Poisson variables

for $\lambda \in \mathbb{R}_{>0}$, Pois($\lambda$) takes size $k \in \mathbb{Z}_{\geq 0}$ w.p. $e^{-\lambda}\lambda^k/k!$

StochNormOpt with Poisson Random Variables

Suppose entries of cost vector Y are Poisson variables

for $\lambda \in \mathbb{R}_{>0}$, Pois($\lambda$) takes size $k \in \mathbb{Z}_{\geq 0}$ w.p. $e^{-\lambda}\lambda^k/k!$

Using Schur-convexity of $g(y) := \mathbb{E}[f(\text{Pois}(y_1), \ldots, \text{Pois}(y_m))]$ over $\mathbb{R}^m_{\geq 0}$,

StochNormOpt with Poisson Random Variables

Suppose entries of cost vector Y are Poisson variables

for $\lambda \in \mathbb{R}_{>0}$, Pois($\lambda$) takes size $k \in \mathbb{Z}_{\geq 0}$ w.p. $e^{-\lambda}\lambda^k/k!$

Using Schur-convexity of $g(y) := \mathbb{E}[f(\text{Pois}(y_1), \ldots, \text{Pois}(y_m))]$ over $\mathbb{R}^m_{\geq 0}$,

PoisNormOpt essentially reduces to Deterministic Min-Norm Opt (possibly with a different norm)

StochNormOpt with Poisson Random Variables

Suppose entries of cost vector Y are Poisson variables

for $\lambda \in \mathbb{R}_{>0}$, Pois($\lambda$) takes size $k \in \mathbb{Z}_{\geq 0}$ w.p. $e^{-\lambda}\lambda^k/k!$

Using Schur-convexity of $g(y) := \mathbb{E}[f(\text{Pois}(y_1), \ldots, \text{Pois}(y_m))]$ over $\mathbb{R}^m_{\geq 0}$,

PoisNormOpt	essentially reduces to	Deterministic Min-Norm Opt (possibly with a different norm)

▷ In PoisNormLB, size of job j on m/c i is distributed as Pois(λ_{ij})

Load on m/c i is $\text{Pois}\left(\sum_{\text{job } j \mapsto i} \lambda_{ij} \right)$

In [IS21] we give: $(2+\varepsilon)$-approx for unrelated machines, PTAS for identical machines

StochNormOpt with Poisson Random Variables

Suppose entries of cost vector Y are Poisson variables

for $\lambda \in \mathbb{R}_{>0}$, Pois($\lambda$) takes size $k \in \mathbb{Z}_{\geq 0}$ w.p. $e^{-\lambda}\lambda^k/k!$

Using Schur-convexity of $g(y) := \mathbb{E}[f(\text{Pois}(y_1), \ldots, \text{Pois}(y_m))]$ over $\mathbb{R}^m_{\geq 0}$,

PoisNormOpt	essentially reduces to	Deterministic Min-Norm Opt (possibly with a different norm)

▷ In PoisNormLB, size of job j on m/c i is distributed as Pois(λ_{ij})

Load on m/c i is Pois$\left(\sum_{\text{job } j \mapsto i} \lambda_{ij}\right)$

In [IS21] we give: $(2+\varepsilon)$-approx for unrelated machines, PTAS for identical machines

Substantial generalization of [DKLN20]'s PTAS for ℓ_∞ case

StochNormOpt with Poisson Random Variables

Suppose entries of cost vector Y are Poisson variables

for $\lambda \in \mathbb{R}_{>0}$, Pois($\lambda$) takes size $k \in \mathbb{Z}_{\geq 0}$ w.p. $e^{-\lambda}\lambda^k/k!$

Using Schur-convexity of $g(y) := \mathbb{E}[f(\text{Pois}(y_1), \ldots, \text{Pois}(y_m))]$ over $\mathbb{R}^m_{\geq 0}$,

PoisNormOpt	essentially reduces to	Deterministic Min-Norm Opt (possibly with a different norm)

▷ In PoisNormLB, size of job j on m/c i is distributed as Pois(λ_{ij})

Load on m/c i is $\text{Pois}\left(\sum_{\text{job } j \mapsto i} \lambda_{ij}\right)$

In [IS21] we give: $(2+\varepsilon)$-approx for unrelated machines, PTAS for identical machines

▷ In PoisNormTree, weight of edge e is Pois(λ_e)

PoisNormTree can be solved exactly because deterministic version can be solved exactly

Conclusions

Summary

Summary

▷ We show $\mathbb{E}[f(Y)] = \Theta\Big(f(\mathbb{E}[Y^{\downarrow}])\Big)$ for product distributions Y.

Summary

- ▷ We show $\mathbb{E}[f(Y)] = \Theta\big(f(\mathbb{E}[Y^{\downarrow}])\big)$ for product distributions Y.

- ▷ We give a framework for stochastic f-norm optimization based on a reduction to simultaneous stochastic Top_{ℓ}-norm optimization for ℓ's that are powers-of-2.

Summary

- ▷ We show $\mathbb{E}[f(Y)] = \Theta\big(f(\mathbb{E}[Y^{\downarrow}])\big)$ for product distributions Y.
- ▷ We give a framework for stochastic f-norm optimization based on a reduction to simultaneous stochastic Top_ℓ-norm optimization for ℓ's that are powers-of-2.
- ▷ We give two $\Theta(1)$-approximate proxy functions for $\mathbb{E}[\mathsf{Top}_\ell(Y)]$ that are separable, linear, and simple.

Summary

- ▷ We show $\mathbb{E}[f(Y)] = \Theta\big(f(\mathbb{E}[Y^{\downarrow}])\big)$ for product distributions Y.
- ▷ We give a framework for stochastic f-norm optimization based on a reduction to simultaneous stochastic Top_ℓ-norm optimization for ℓ's that are powers-of-2.
- ▷ We give two $\Theta(1)$-approximate proxy functions for $\mathbb{E}[\mathsf{Top}_\ell(Y)]$ that are separable, linear, and simple.
- ▷ Using our framework, we obtain approximation algorithms for stochastic min-norm optimization problems arising from load balancing and spanning tree applications.

Open Problems

Open Problems

- Is there an $O(1)$-approximation algorithm for StochNormLB with arbitrary monotone symmetric norms and arbitrary job-size distributions? (even the setting with identical machines is open)

Open Problems

▷ Is there an $O(1)$-approximation algorithm for StochNormLB with arbitrary monotone symmetric norms and arbitrary job-size distributions?
(even the setting with identical machines is open)

▷ Tighter bounds on the gap between $\mathbb{E}[f(Y)]$ and $f(\mathbb{E}[Y^{\downarrow}])$
(current best bound is 7.634)

Open Problems

- ▷ Is there an $O(1)$-approximation algorithm for StochNormLB with arbitrary monotone symmetric norms and arbitrary job-size distributions?
  (even the setting with identical machines is open)

- ▷ Tighter bounds on the gap between $\mathbb{E}[f(Y)]$ and $f(\mathbb{E}[Y^{\downarrow}])$
  (current best bound is 7.634)

- ▷ Other natural settings of StochNormOpt:
  bipartite perfect matchings
  k-clustering

Thank You

More general models of StochNormOpt

- ▷ StochNormOpt when the cost vector Y **does not** follow a product distribution:

  stochastic load balancing with correlated jobs

  stochastic unsplittable flows

- ▷ Generalizations of StochNormOpt that allow:

  Probing

  Multi-stage decisions

  Adaptive solutions

Gap between $\mathbb{E}[f(Y)]$ and $f(\mathbb{E}[Y^{\downarrow}])$

- ▷ Consider product distribution Y over $\mathbb{R}^m_{\geq 0}$ where each Y_i is a Bernoulli with activation probability $1/m$
- ▷ Suppose $m \to \infty$
- ▷ $\mathbb{E}[\mathsf{Top}_1(Y)] = 1 - \frac{1}{e}$ and $\mathbb{E}[\mathsf{Top}_m(Y)] = 1$
- ▷ Consider monotone symmetric norm f given by

$$f(y) := \max\left(\frac{e}{e-1}\mathsf{Top}_1(y),\ \mathsf{Top}_m(y)\right)$$

- ▷ $f(\mathbb{E}[Y^{\downarrow}]) = 1$ and $\mathbb{E}[f(Y)] = 1 + \frac{1}{e(e-1)} \approx 1.21$

Prior Work on Stochastic Bin Packing

[KRT00]

For weighted Bernoulli items

ALG $O(1/\varepsilon)$-apx. w/ $(1+\varepsilon)$-size bins and overflow prob. p

ALG $O(1/\varepsilon)$-apx. w/ size-1 bins and overflow p, OPT uses size-1 bins and overflow $p^{1+\varepsilon}$

ALG uses $O\left(\sqrt{\frac{\log 1/p}{\log\log 1/p}}\right)B^* + O(\log 1/p)$ size-1 bins and overflow is p

For general distributions, incur a multiplicative $O(\log n)$ loss.

Prior Work on Stochastic Knapsack

[KRT00] Items with high-low sizes

ALG $O(\log 1/p)$-approx.

ALG $O(1/\varepsilon)$-approx. w/ violation in knapsack capacity or overflow probability

[De17]

Bernoullis: (nearly) FPTAS by relaxing overflow probability

Items with shared constant-size support: quasi-FPTAS by relaxing overflow probability

Hypercontractive r.v.s: PTAS that relaxes both capacity and overflow probability

Prior Work on Stochastic Unsplittable Flow

[GK17] Collect value v_j for successfully routed stochastic flows S_j w/ mean μ_j.

Flow paths are chosen adaptively.

No-bottleneck assumption, i.e., $\mathrm{supp}(S_j) \subseteq [0,1]$ and edge capacity are at least 1

Single-sink case: $O\left(\min\left(\log k, \log \frac{\max_j v_j/\mu_j}{\min_j v_j/\mu_j}\right)\right)$-approximation

Trees: (non-adaptive) $O(1)$-approximation

DAGs: $O(\sqrt{n \log k})$-approximation, where k is # source-sink pairs

General graphs: approximation quality depends on max degree, max expansion, and $O(\log^2 n)$