Scenario Min-Max Optimization and the Risk of Empirical Costs

Algo Carè

with S. Garatti[a] and M.C. Campi[b]

[a]Politecnico di Milano [b]University of Brescia

Scenario Min-Max Optimization and the Risk of Empirical Costs

Algo Carè

University of Brescia, Italy

with S. Garatti[a] and M.C. Campi[b]

[a]Politecnico di Milano [b]University of Brescia

Scenario Min-Max Optimization
and
the Risk of Empirical Costs

Algo Carè

University of Brescia, Italy

University of Melbourne, Australia

with S. Garatti[a] and M.C. Campi[b]

[a]Politecnico di Milano [b]University of Brescia

Scenario Min-Max Optimization
and
the Risk of Empirical Costs

Algo Carè

University of Brescia, Italy

University of Melbourne, Australia

MTA SZTAKI, Budapest, Hungary

with S. Garatti[a] and M.C. Campi[b]

[a]Politecnico di Milano [b]University of Brescia

Convex problem

$$\min_{\theta} \ell(\theta)$$

Convex problem

$$\min_{\theta} \ell(\theta)$$

θ design parameter(s)

Convex problem

$$\min_{\theta} \ell(\theta, \delta)$$

θ design parameter(s)

Convex problem

$$\min_{\theta} \ell(\theta, \delta)$$

θ design parameter(s)

δ uncertain parameter

Convex problem

$$\min_{\theta} \ell(\theta, \delta)$$

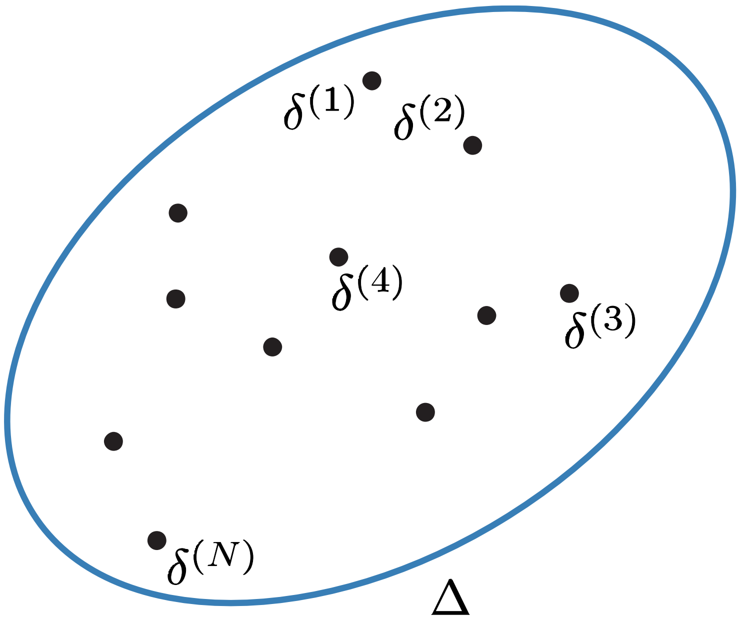
θ design parameter(s)

δ uncertain parameter

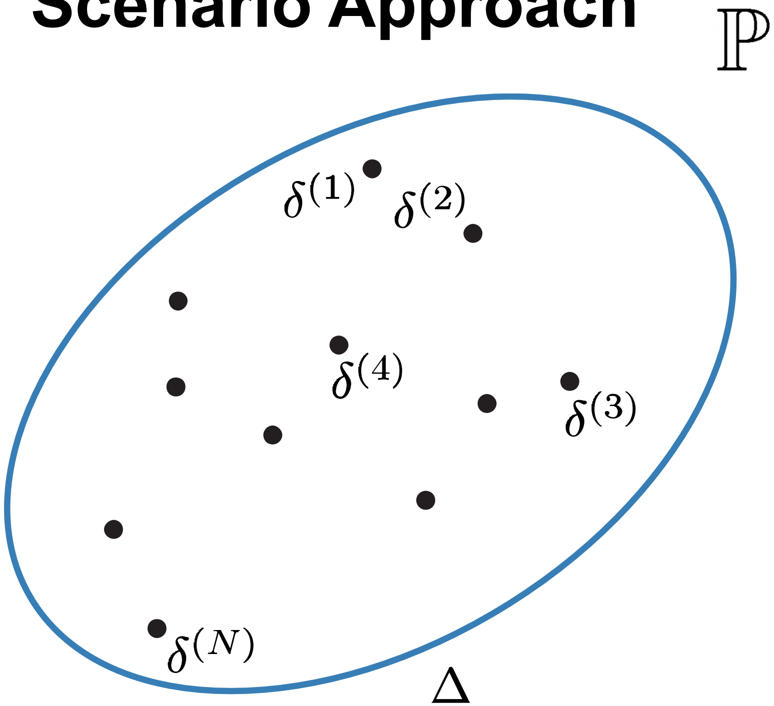
Uncertain problem!

The Scenario Approach

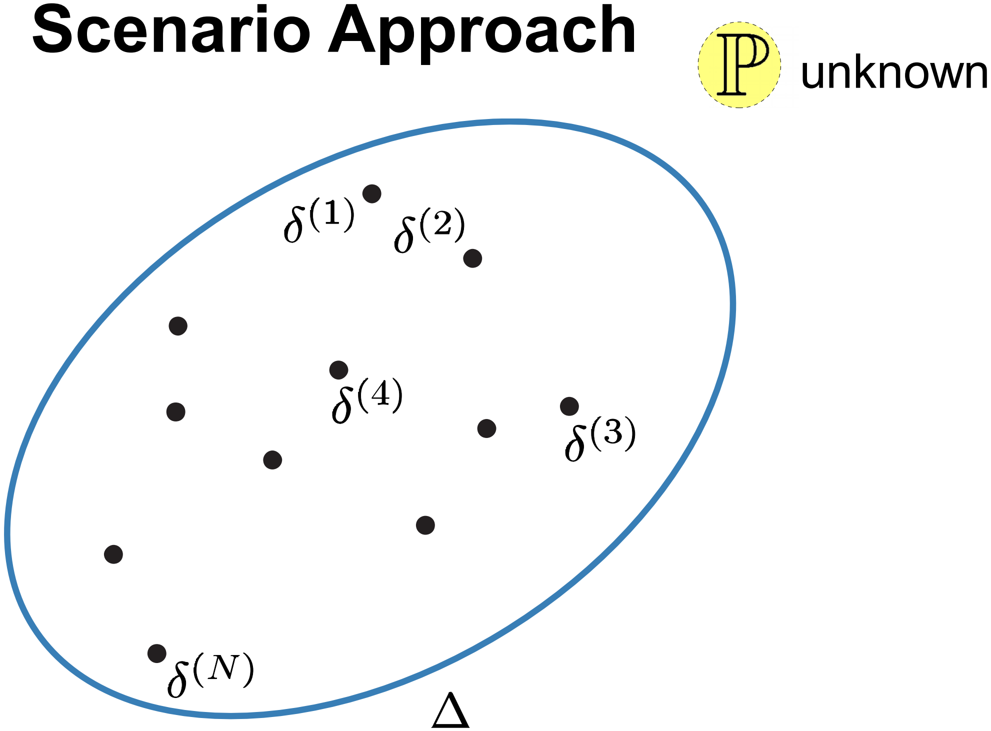
Scenario Approach



Scenario Approach

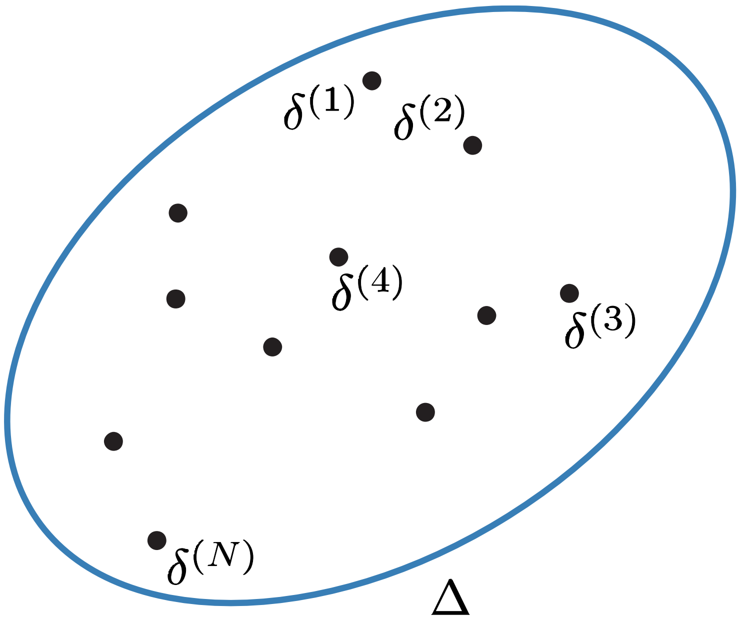


Scenario Approach



Scenario Approach

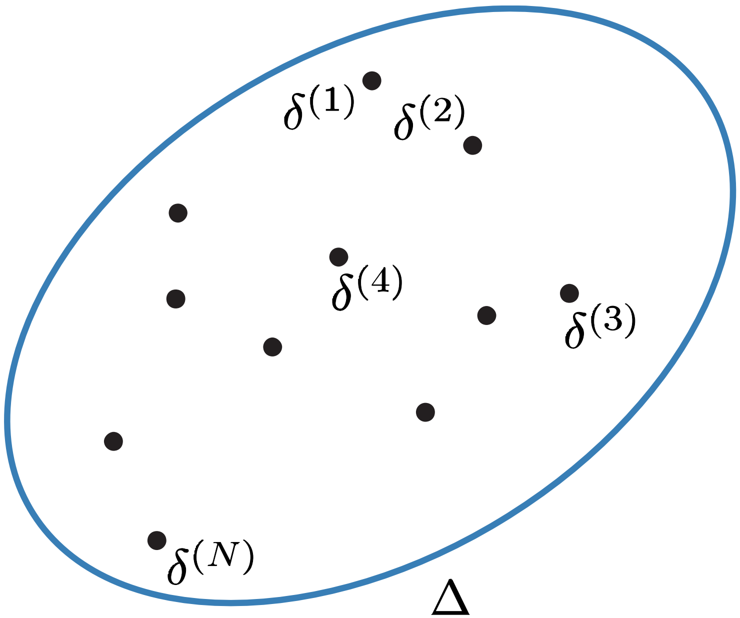
$\mathbb{P}$ unknown



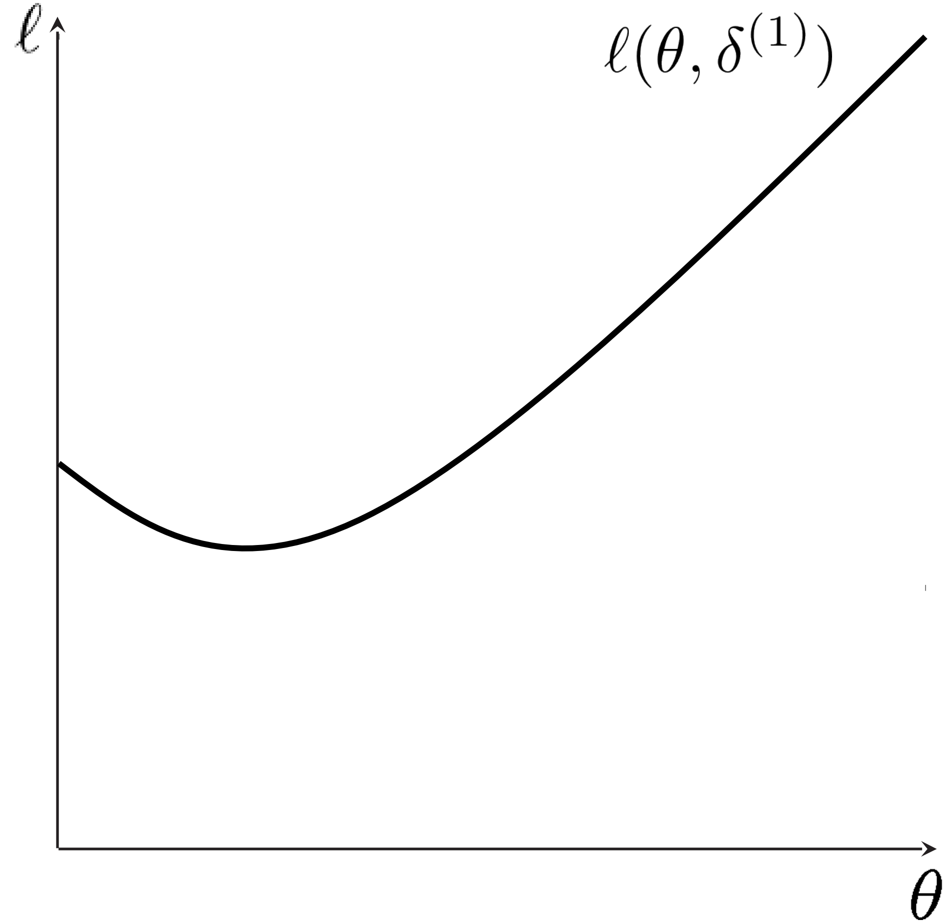
$$\delta^{(1)} \longrightarrow \ell(\theta, \delta^{(1)})$$

Scenario Approach

$\mathbb{P}$ unknown

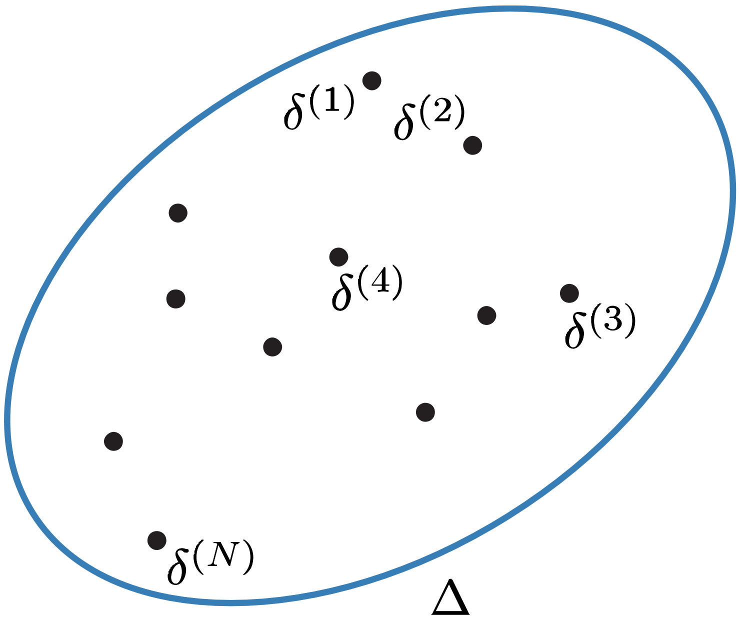


$$\delta^{(1)} \longrightarrow \ell(\theta, \delta^{(1)})$$



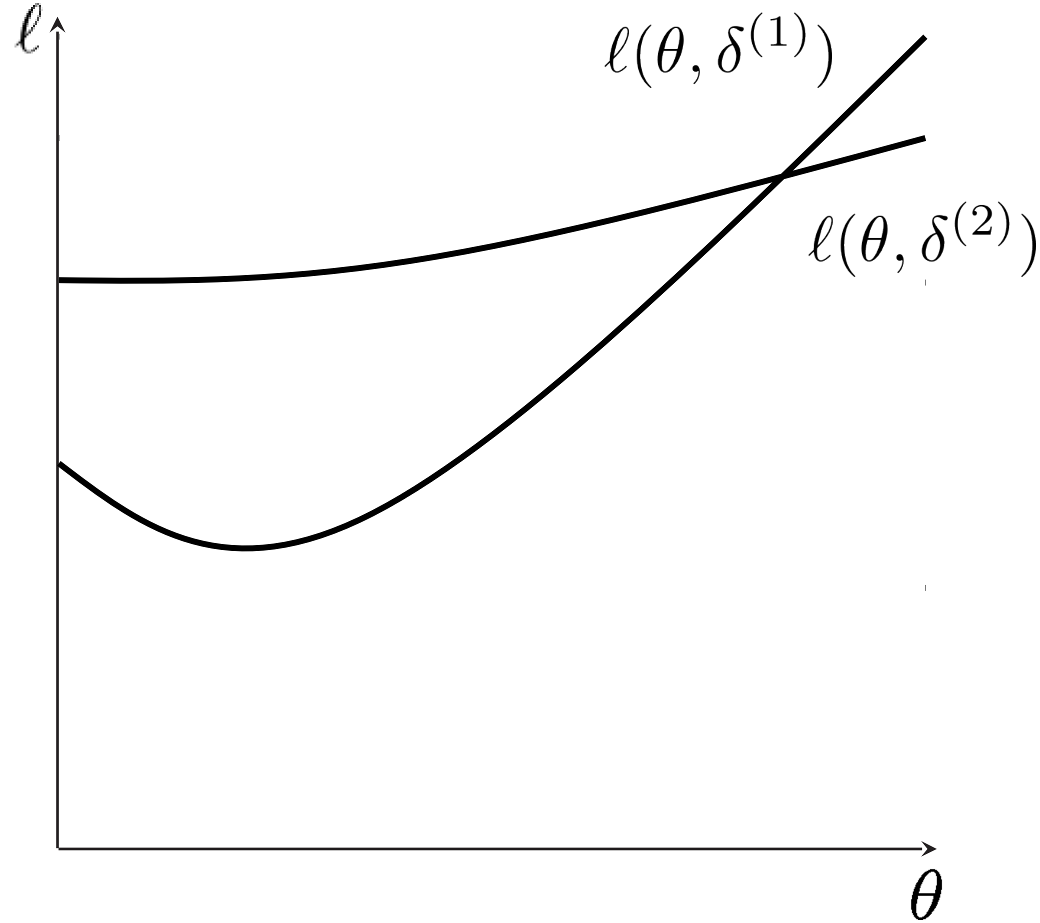
Scenario Approach

$\mathbb{P}$ unknown



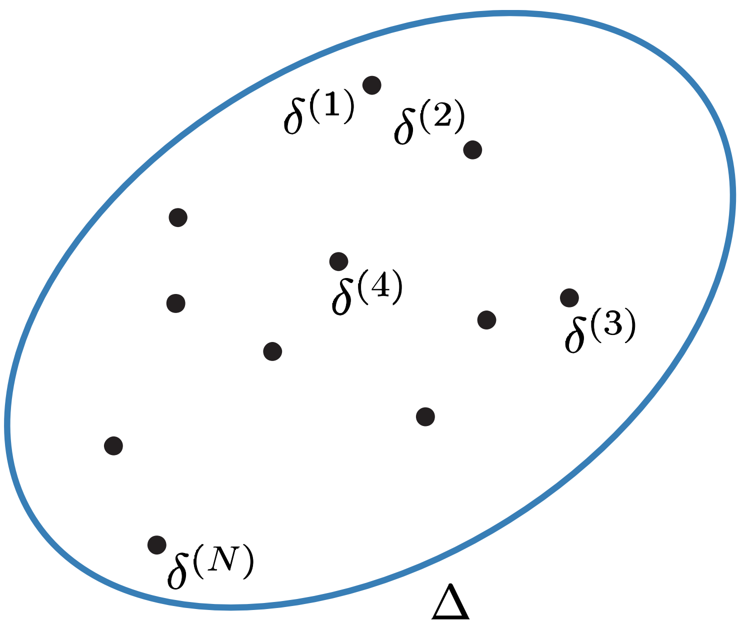
$$\delta^{(1)} \rightarrow \ell(\theta, \delta^{(1)})$$

$$\delta^{(2)} \rightarrow \ell(\theta, \delta^{(2)})$$



Scenario Approach

$\mathbb{P}$ unknown

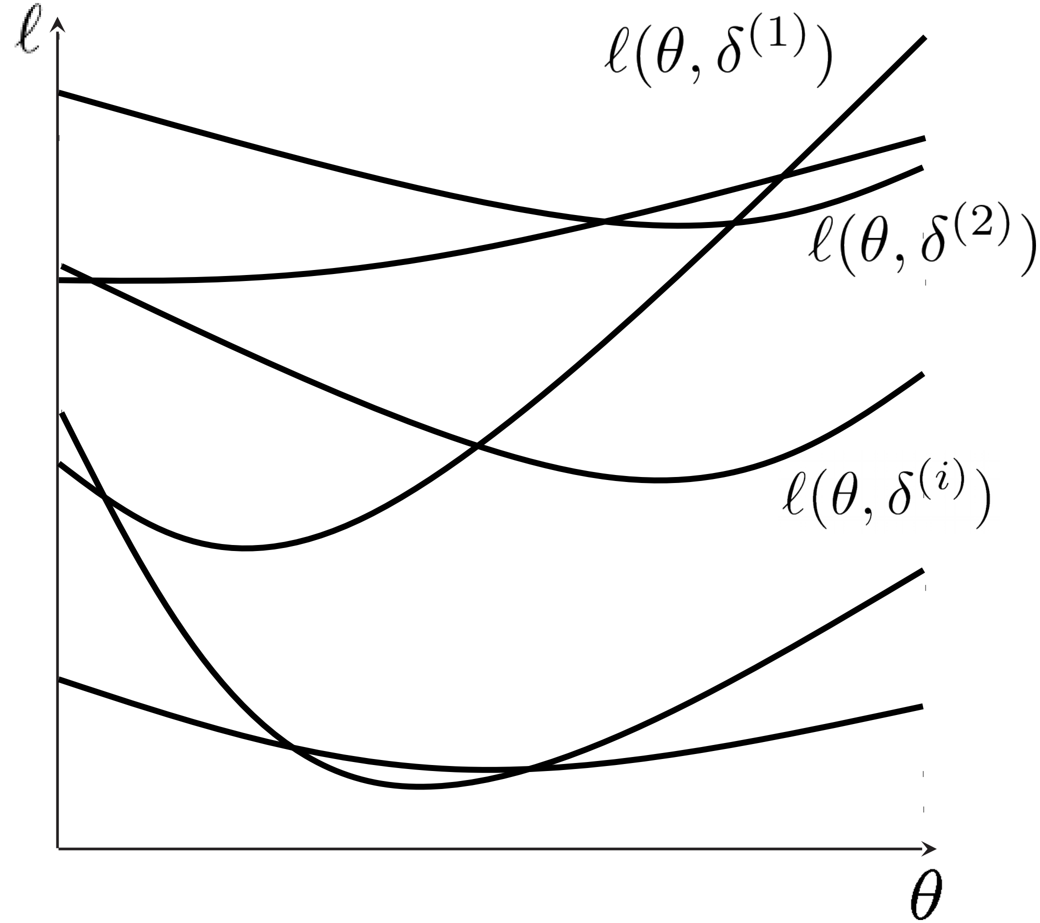


$$\delta^{(1)} \longrightarrow \ell(\theta, \delta^{(1)})$$

$$\delta^{(2)} \longrightarrow \ell(\theta, \delta^{(2)})$$

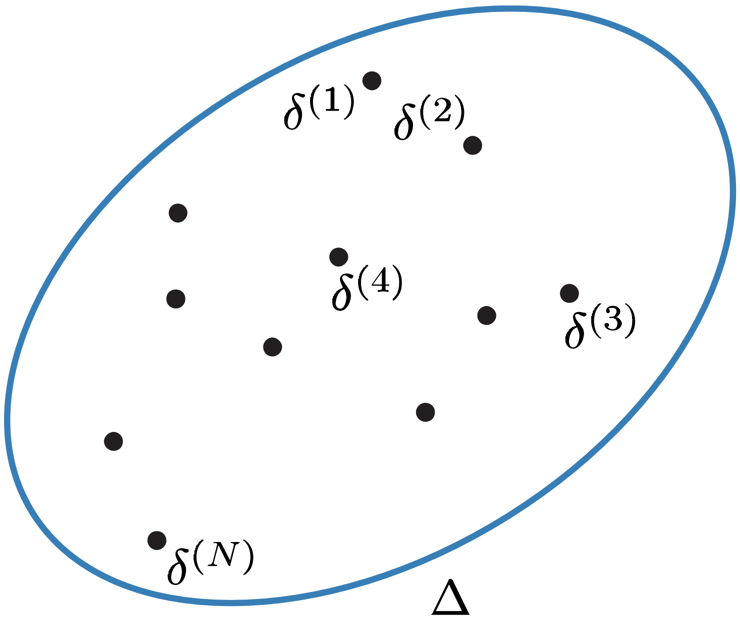
$$\vdots$$

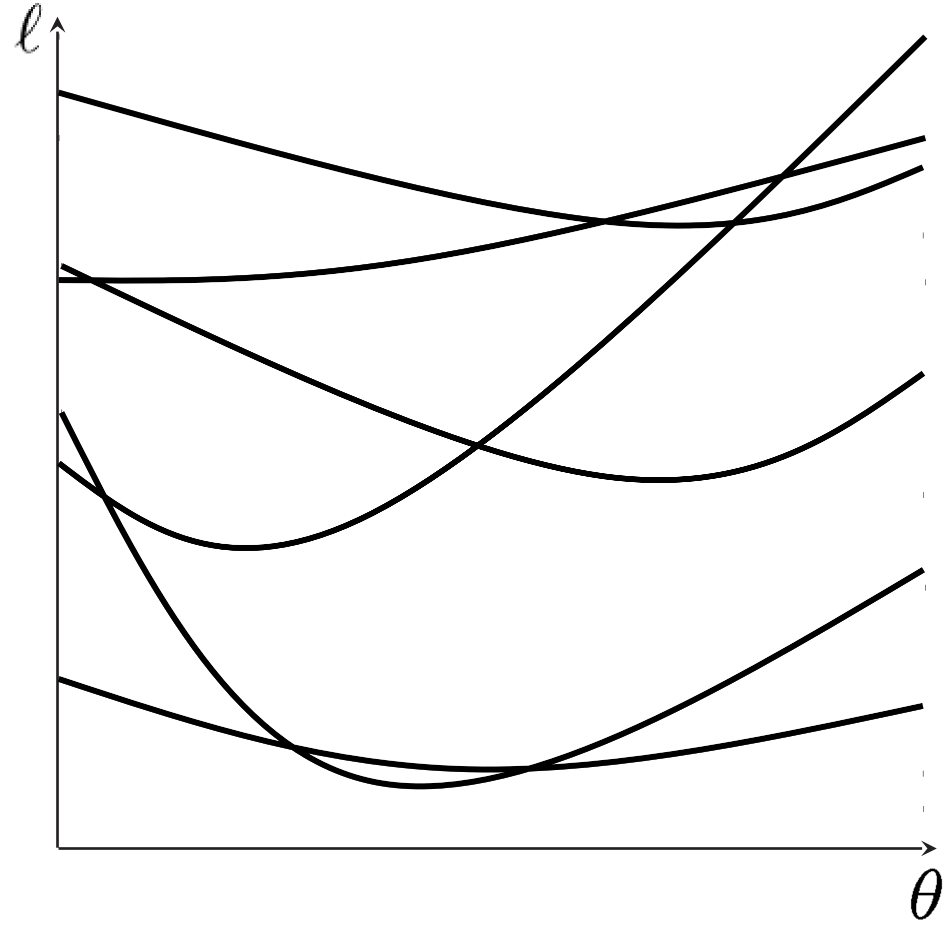
$$\delta^{(N)} \longrightarrow \ell(\theta, \delta^{(N)})$$



Scenario Approach

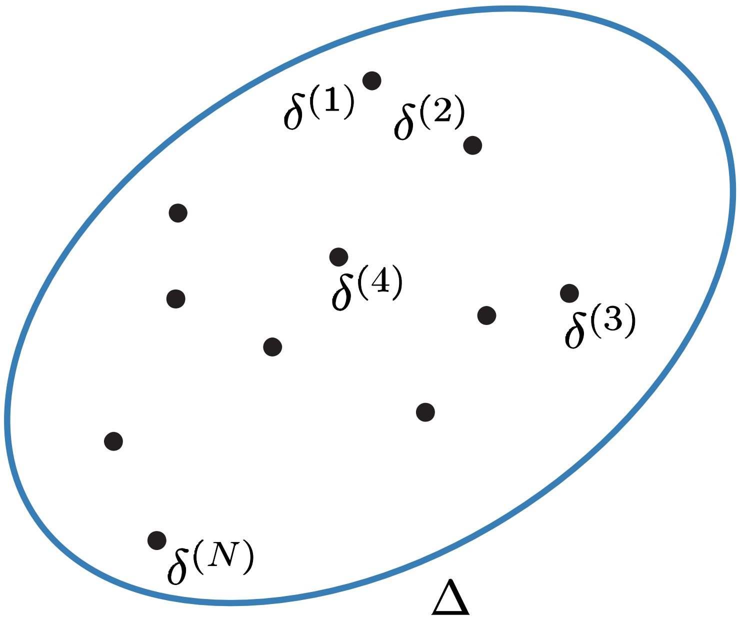
$\mathbb{P}$ unknown



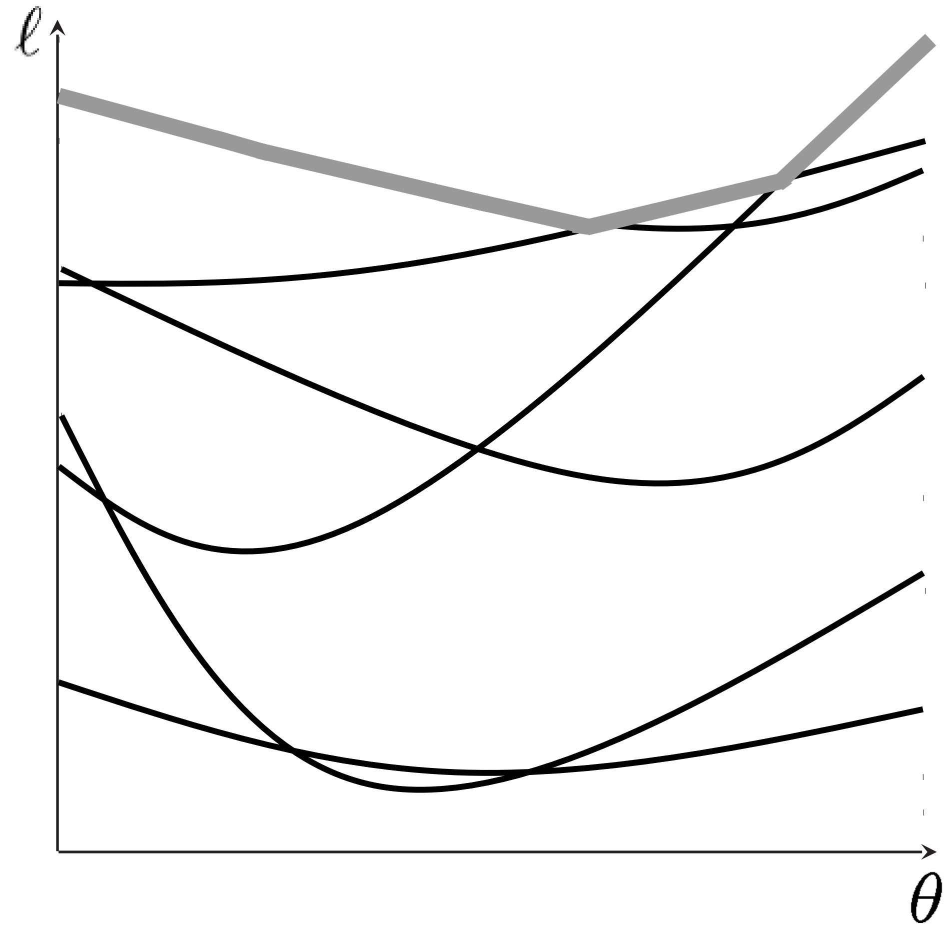


Scenario Approach

$\mathbb{P}$ unknown

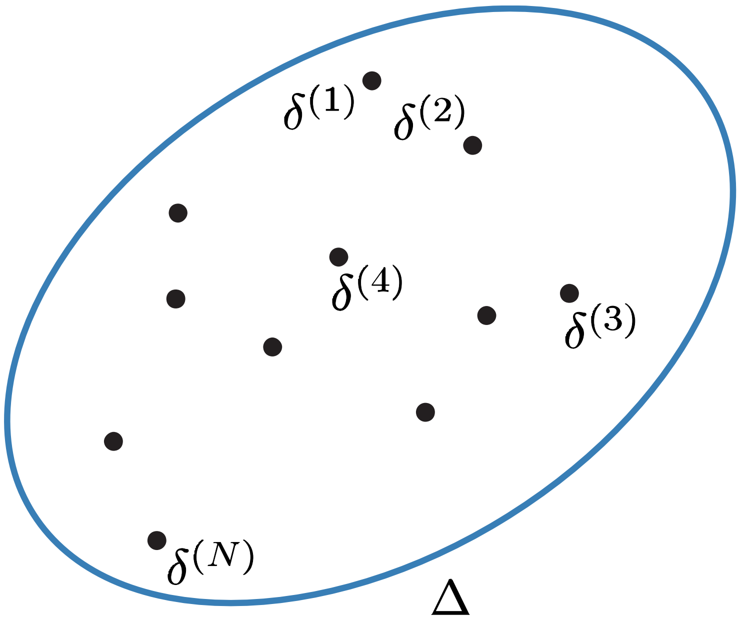


$$\max_{i=1,\ldots,N} \ell(\theta, \delta^{(i)})$$

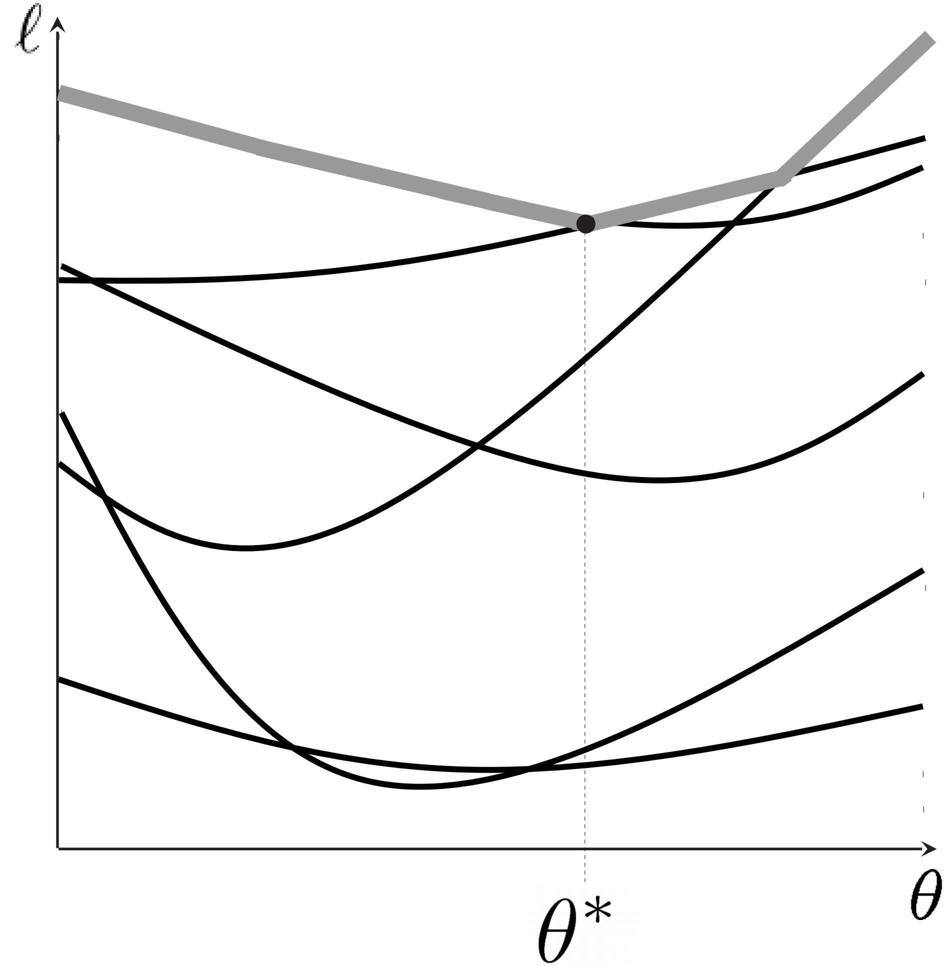


Scenario Approach

$\mathbb{P}$ unknown

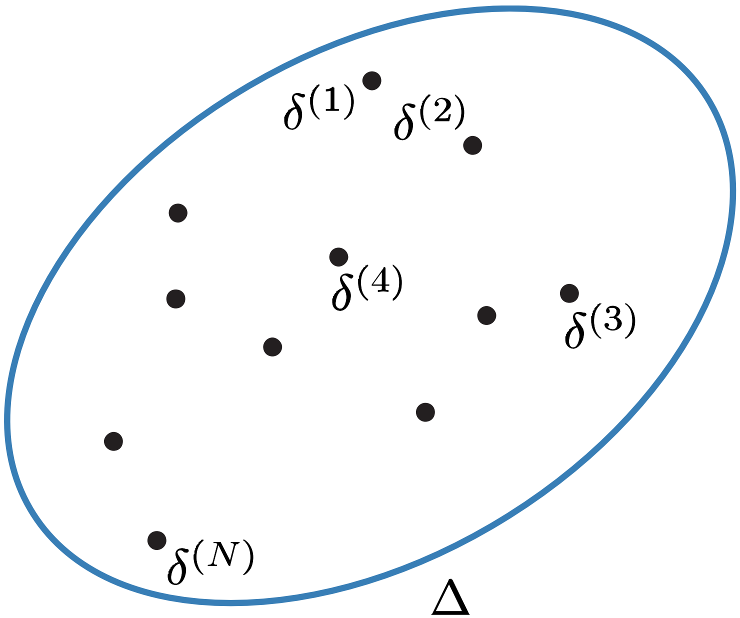


$$\min_{\theta} \max_{i=1,\ldots,N} \ell(\theta, \delta^{(i)})$$

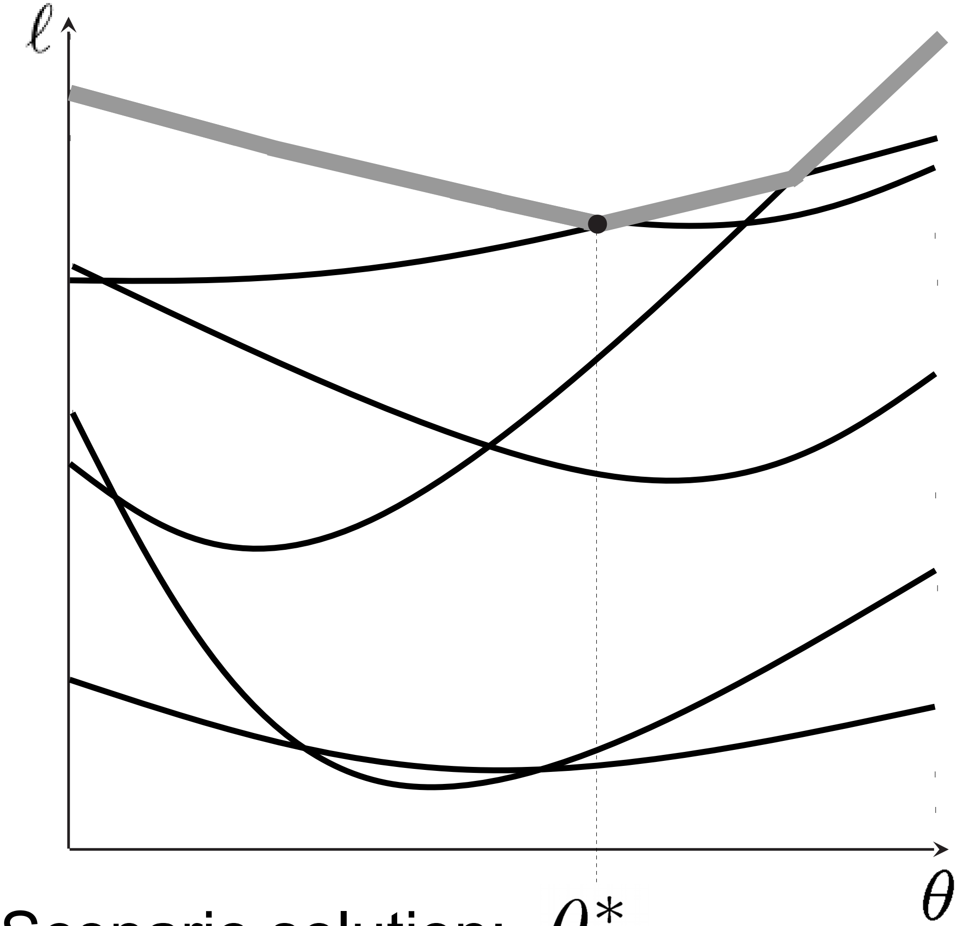


Scenario Approach

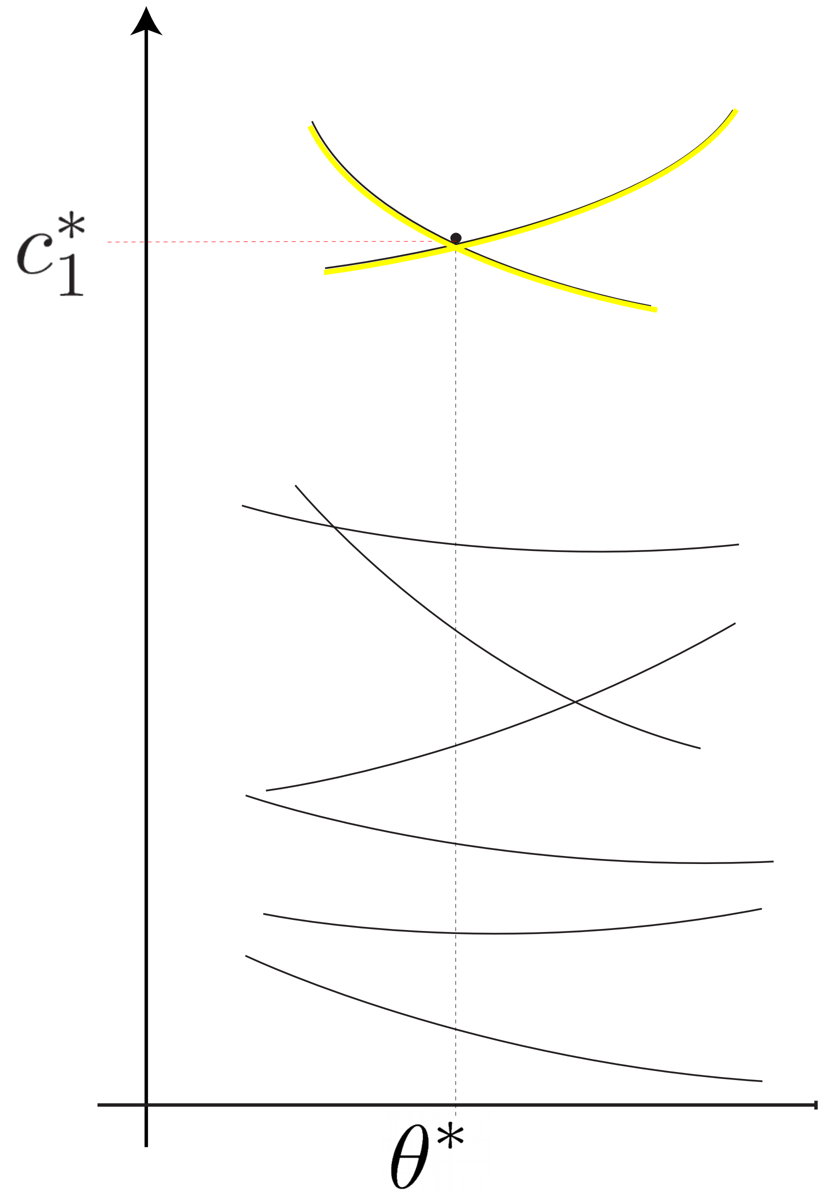
$\mathbb{P}$ unknown



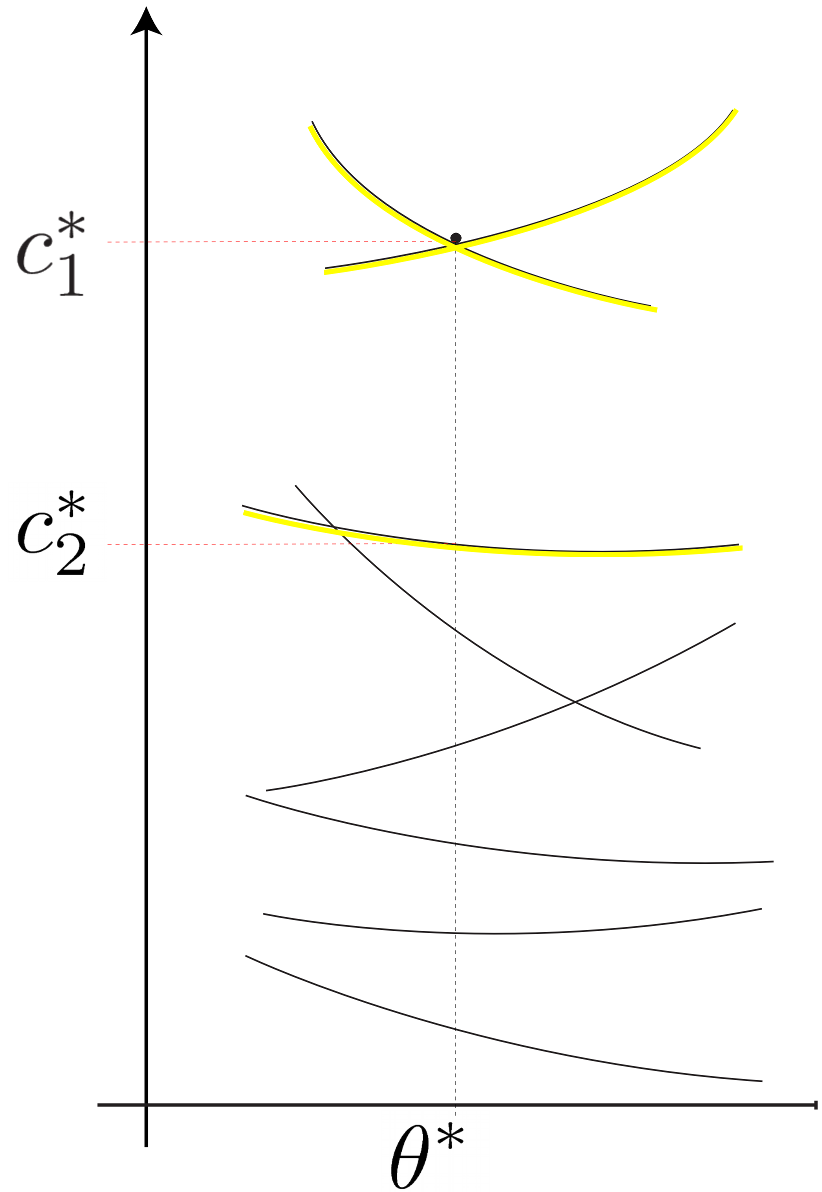
$$\min_{\theta} \max_{i=1,\ldots,N} \ell(\theta, \delta^{(i)})$$



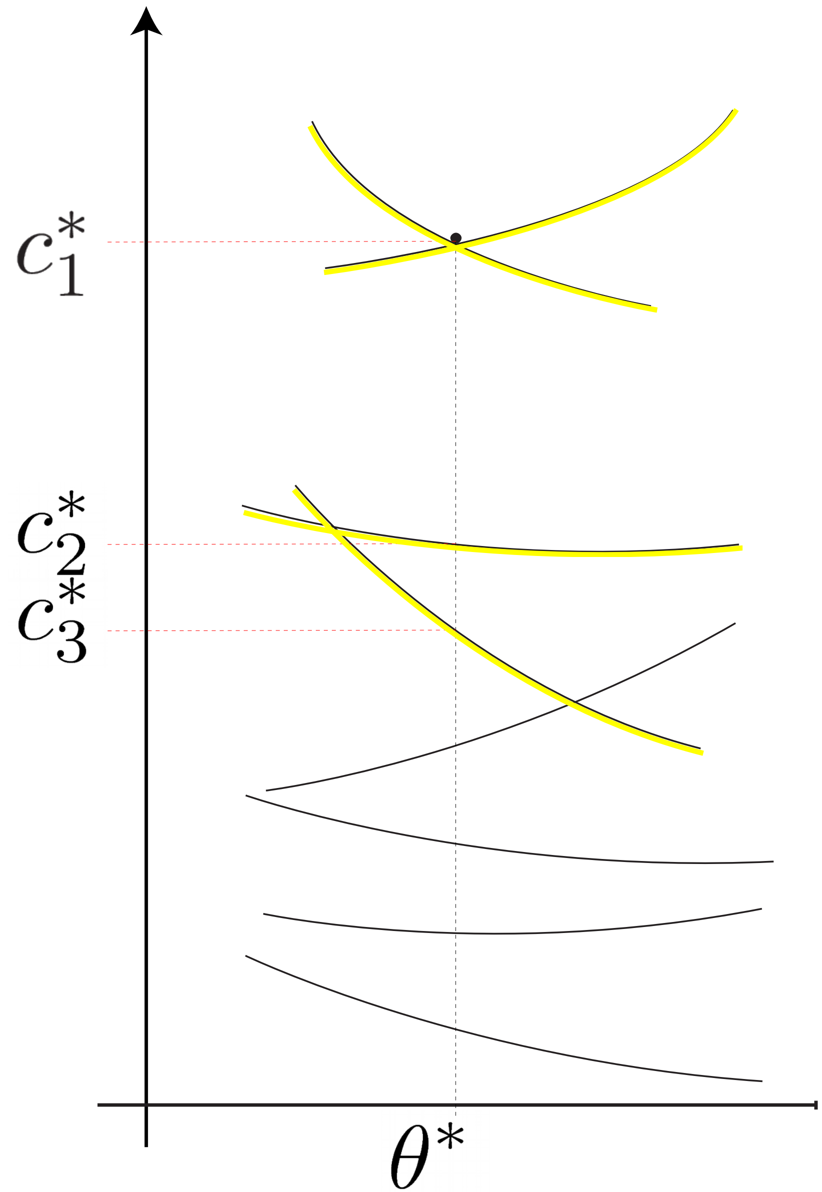
Scenario solution: θ^*



c_1^*
θ^*



c_1^*
c_2^*
θ^*

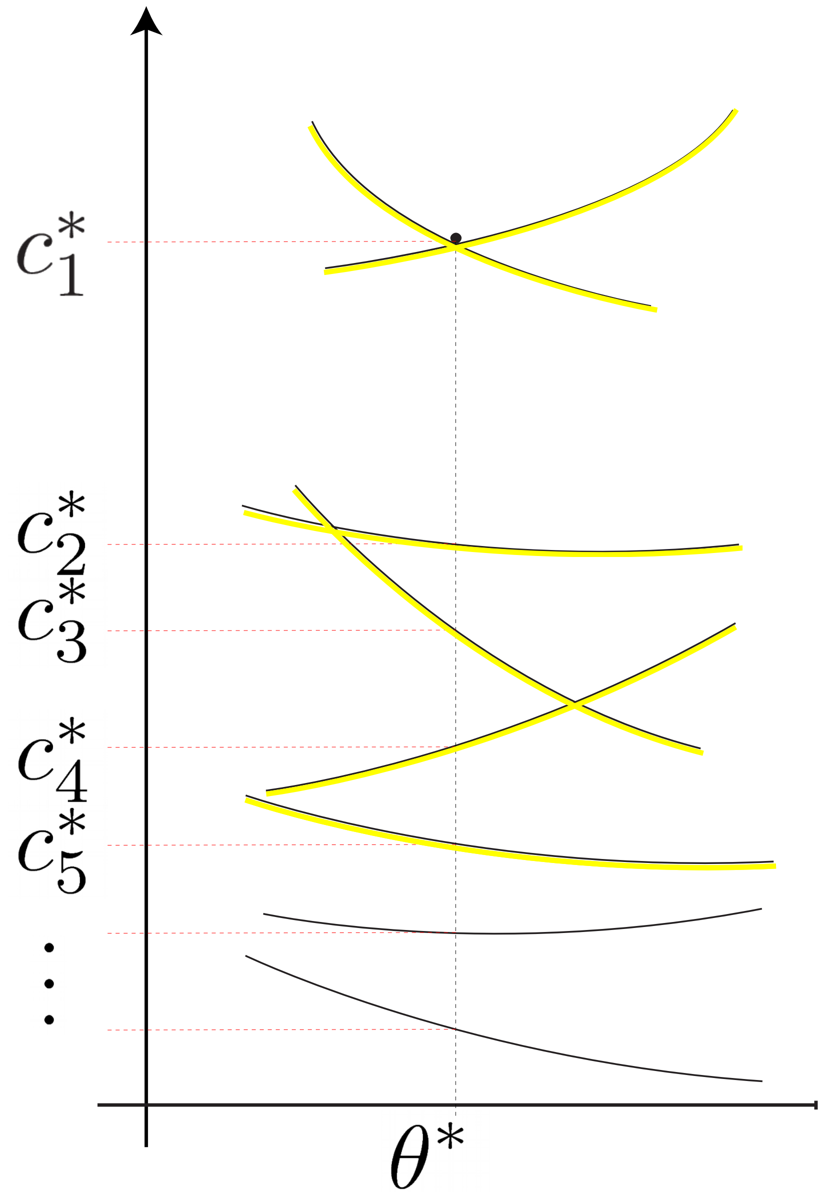

c_1^*
c_2^*
c_3^*
θ^*

Ordered values of

$$\ell(\theta^*, \delta^{(1)})$$

$$\vdots$$

$$\ell(\theta^*, \delta^{(N)})$$



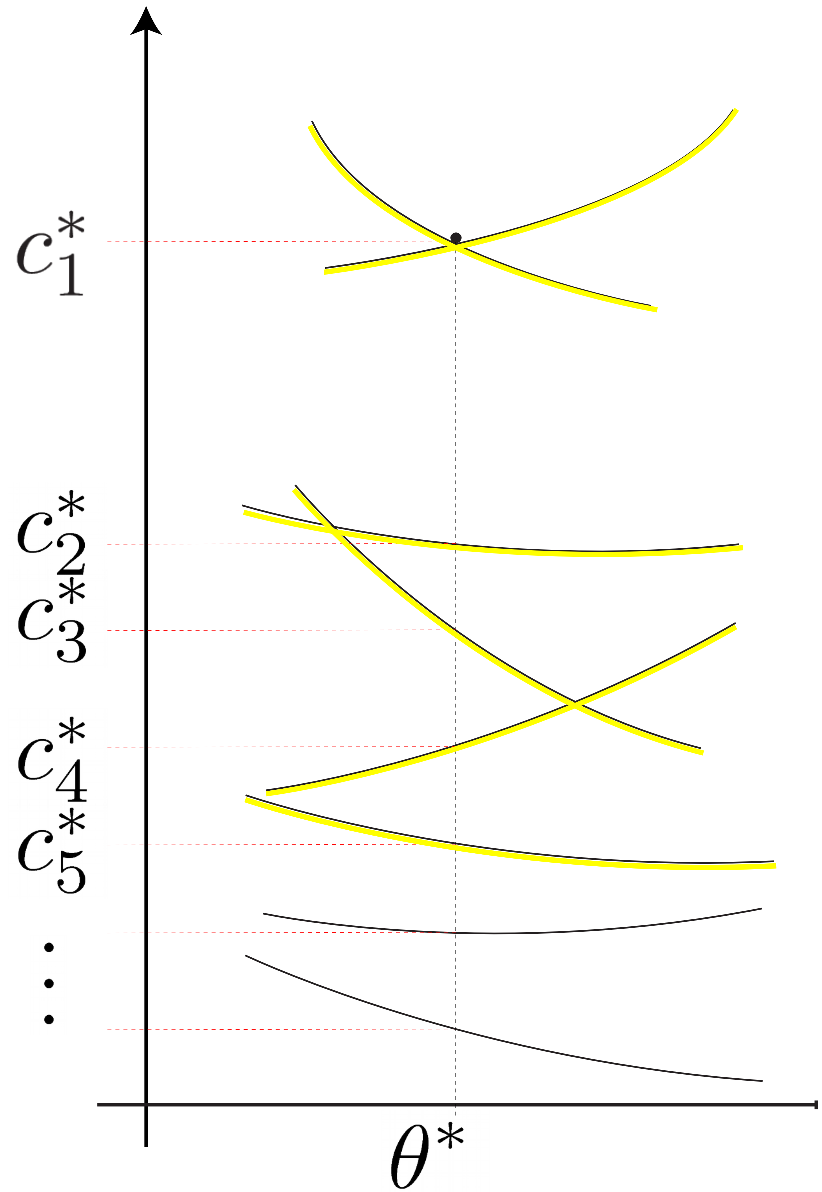
Empirical distribution of $\ell(\theta^*, \delta)$

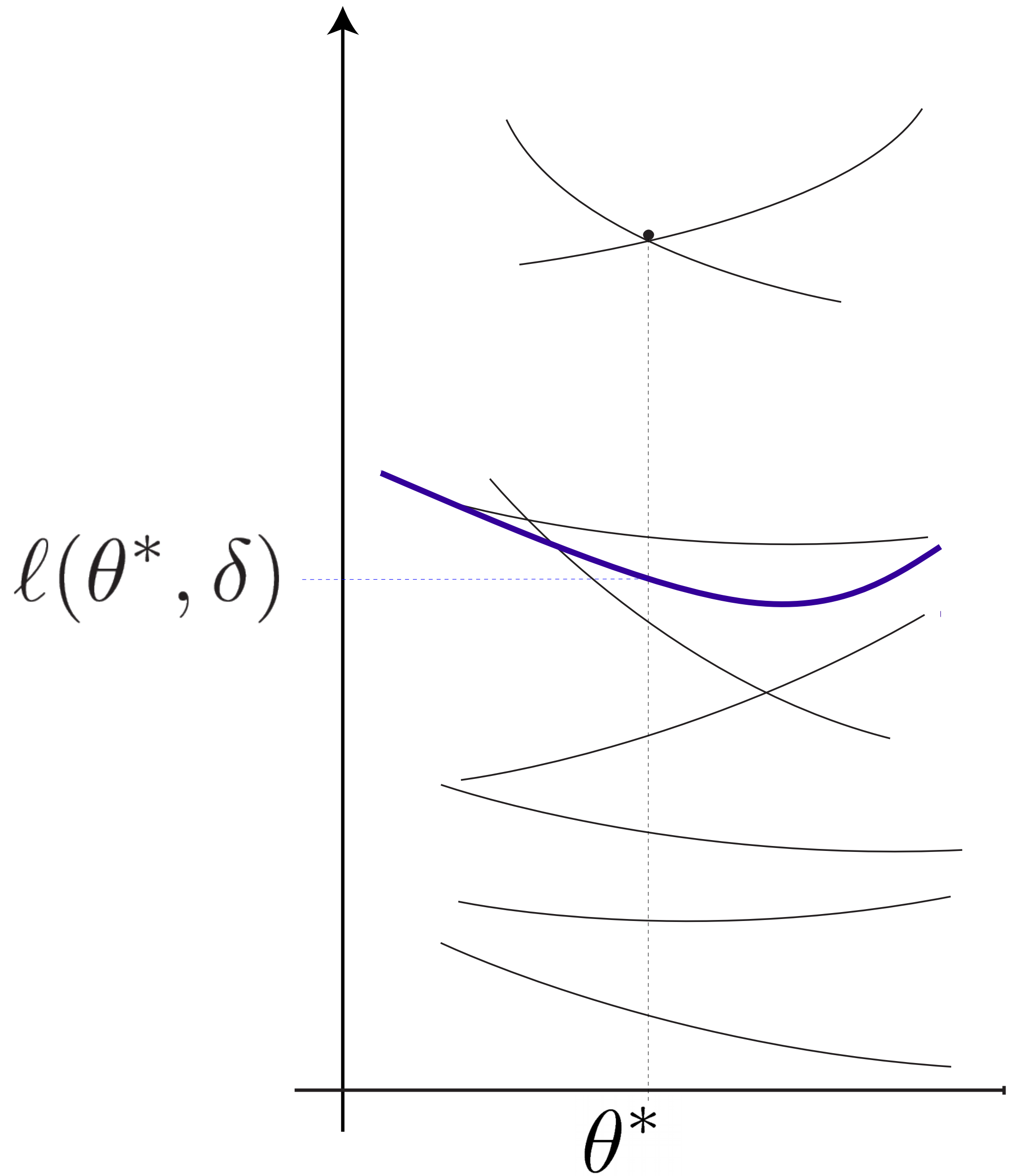
Ordered values of

$$\ell(\theta^*, \delta^{(1)})$$

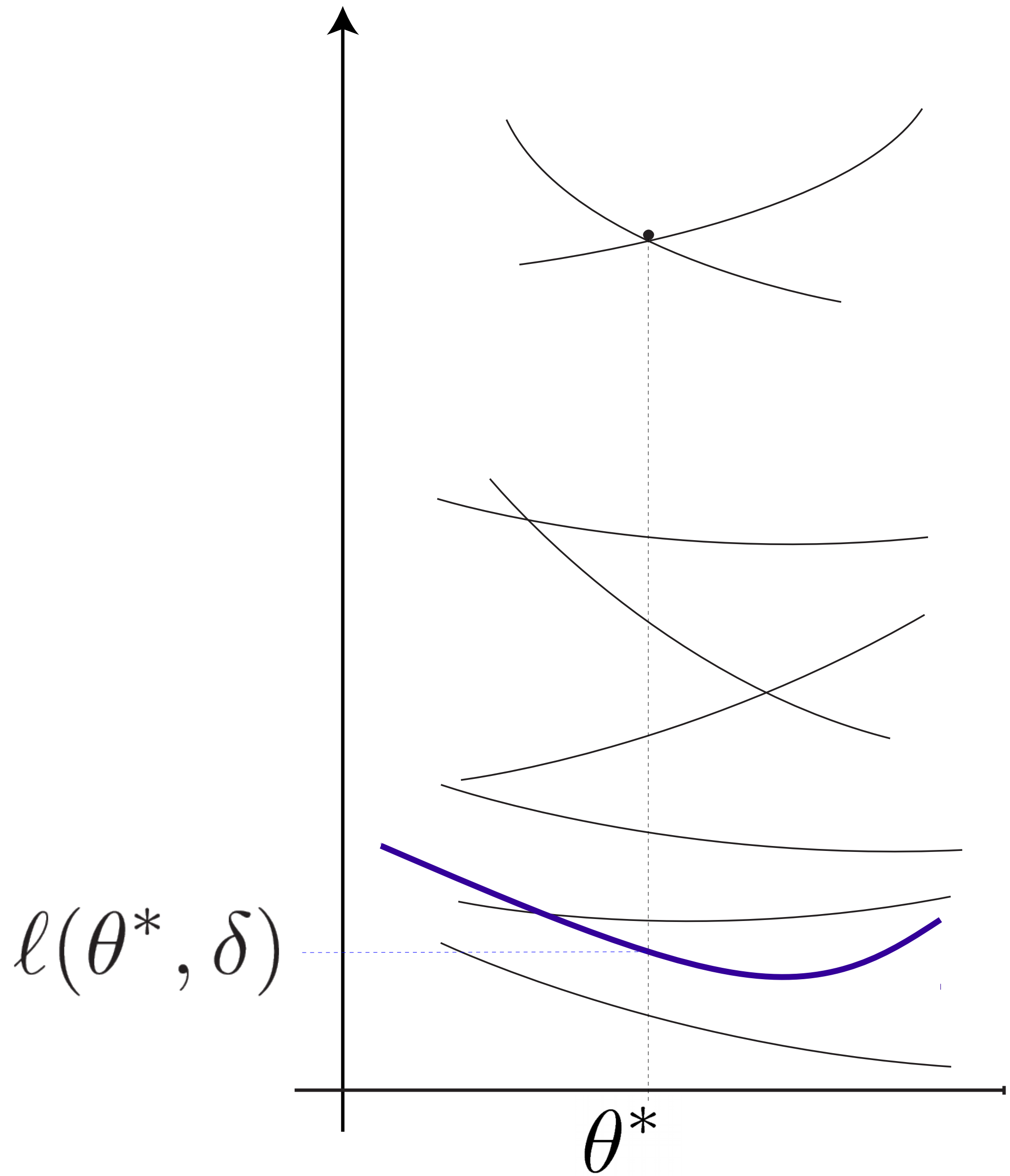
$$\vdots$$

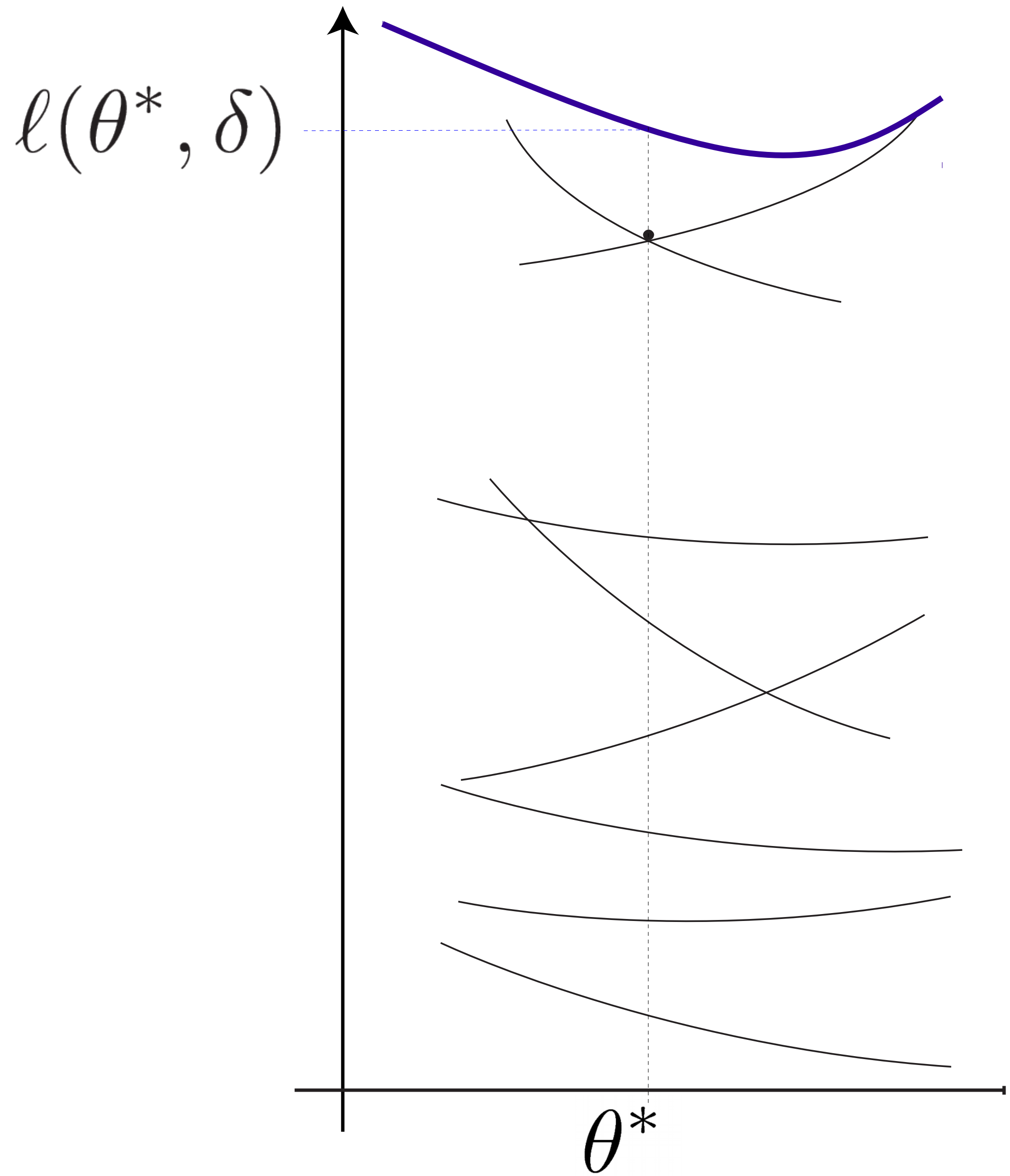
$$\ell(\theta^*, \delta^{(N)})$$



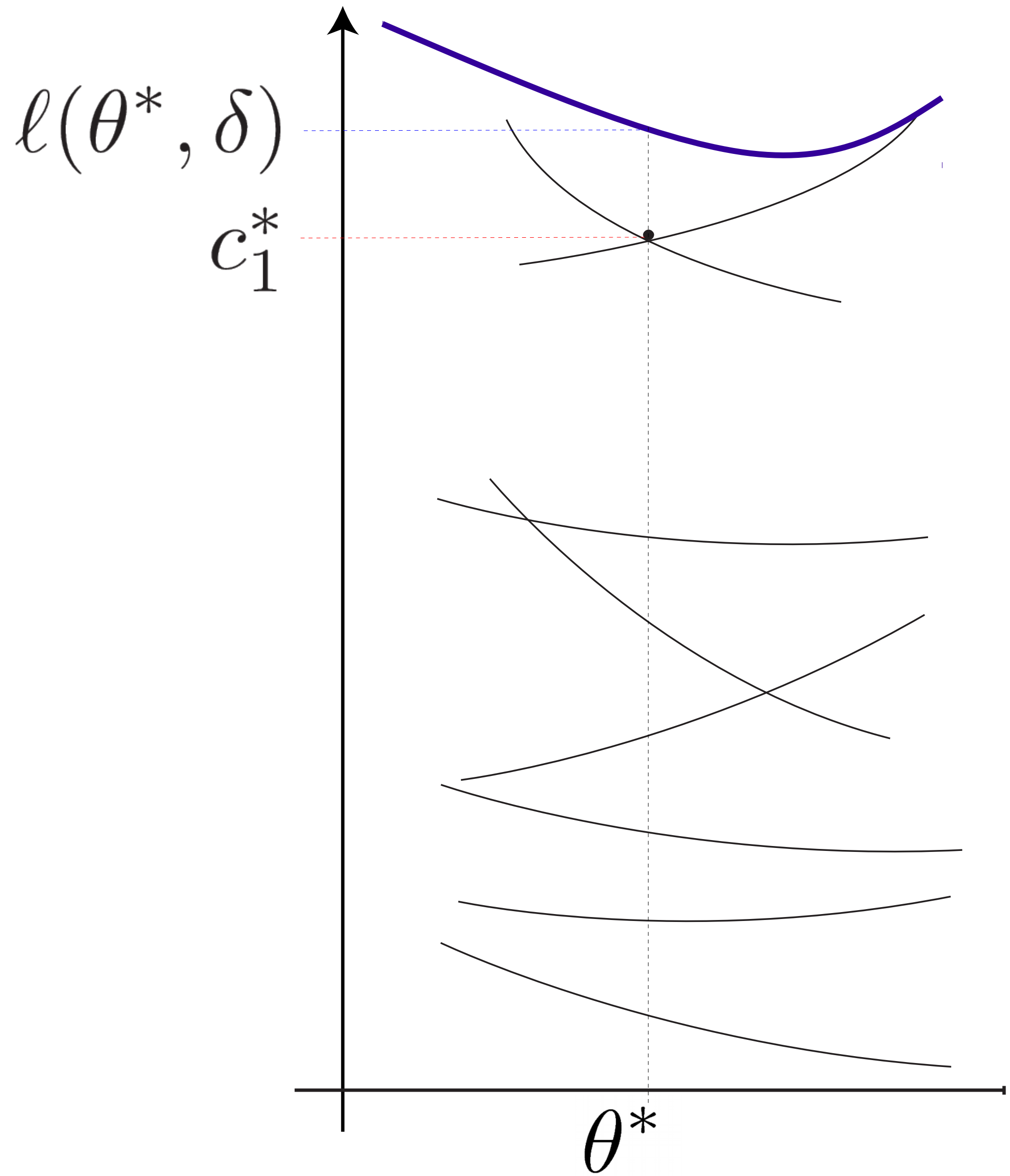


$\ell(\theta^*, \delta)$
θ^*

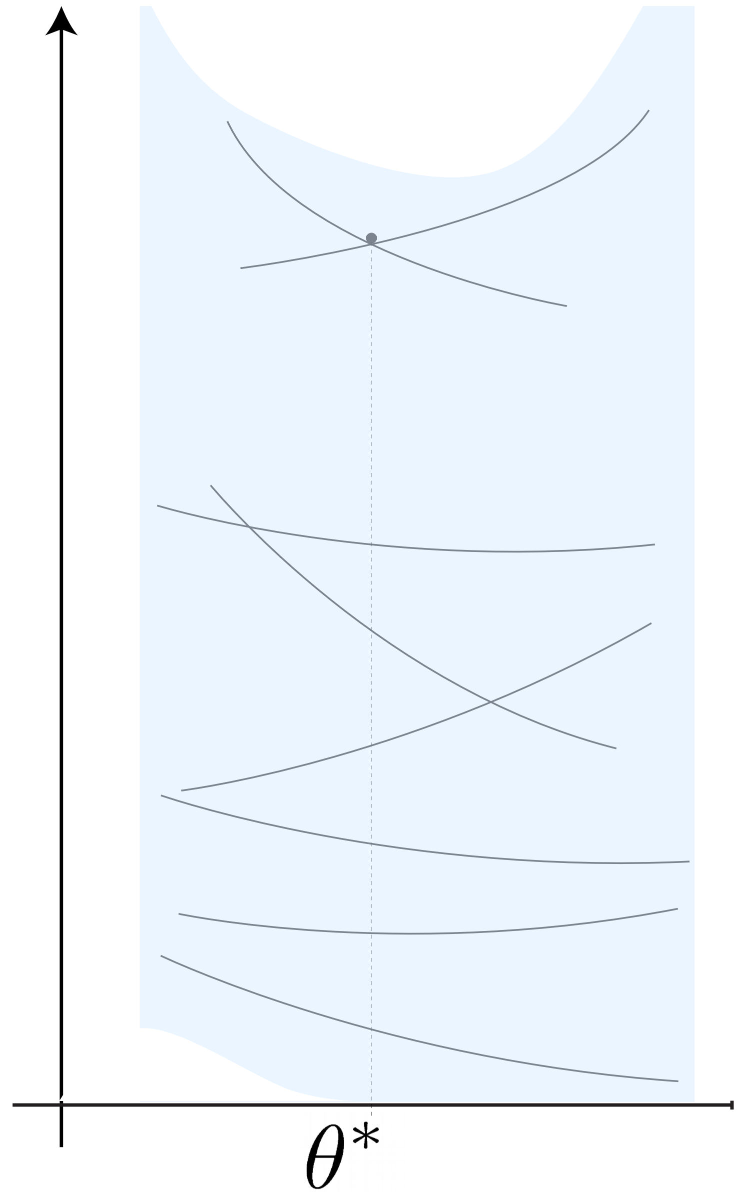

$\ell(\theta^*, \delta)$
θ^*



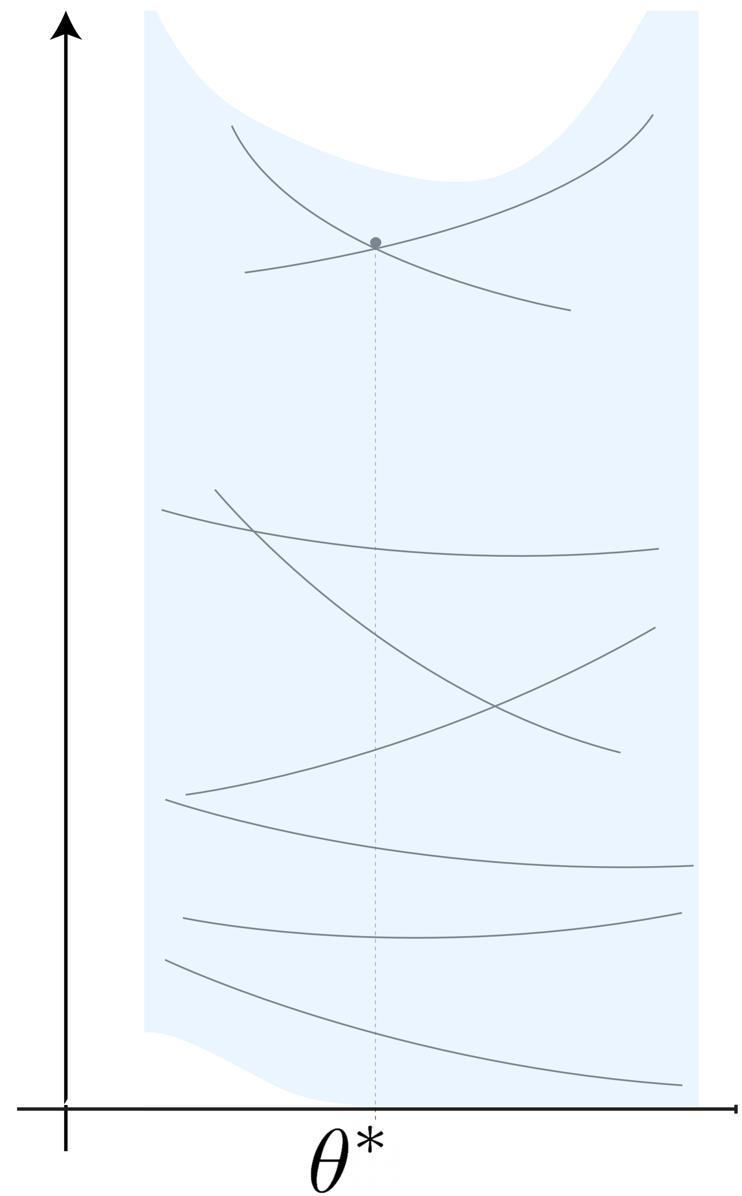
$\ell(\theta^*, \delta)$
θ^*



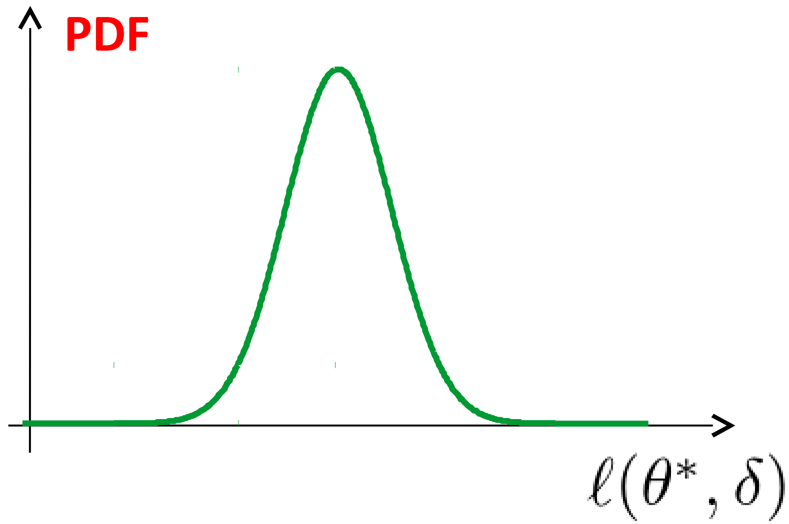
$\ell(\theta^*, \delta)$
c_1^*
θ^*

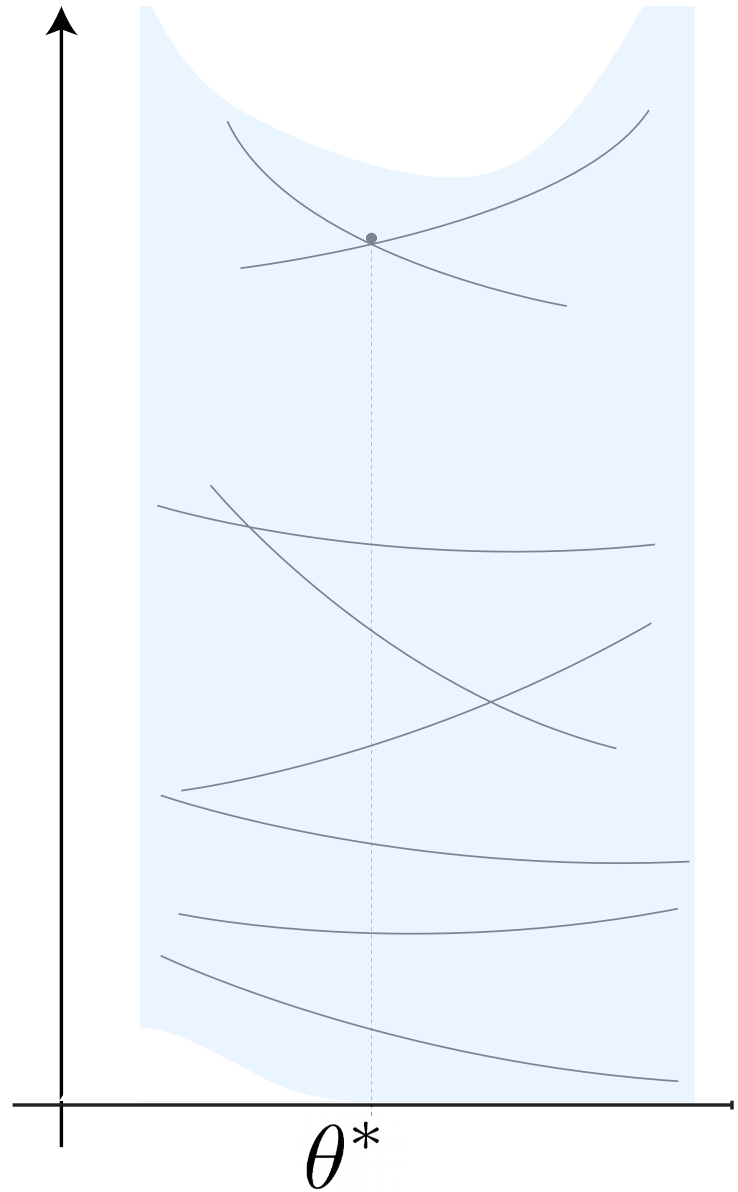

θ^*

Probability distribution of $\ell(\theta^*, \delta)$

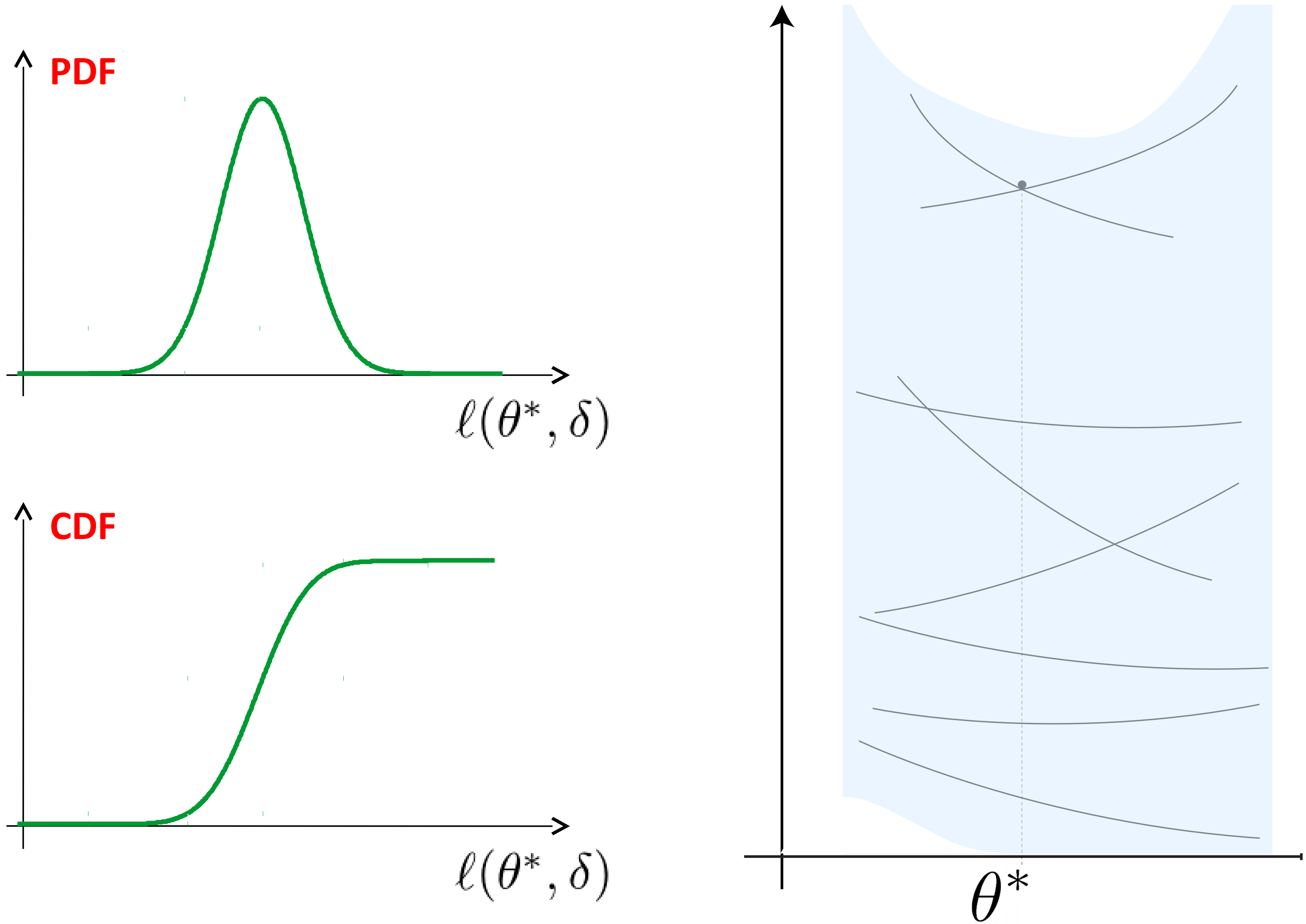


Probability distribution of $\ell(\theta^*, \delta)$

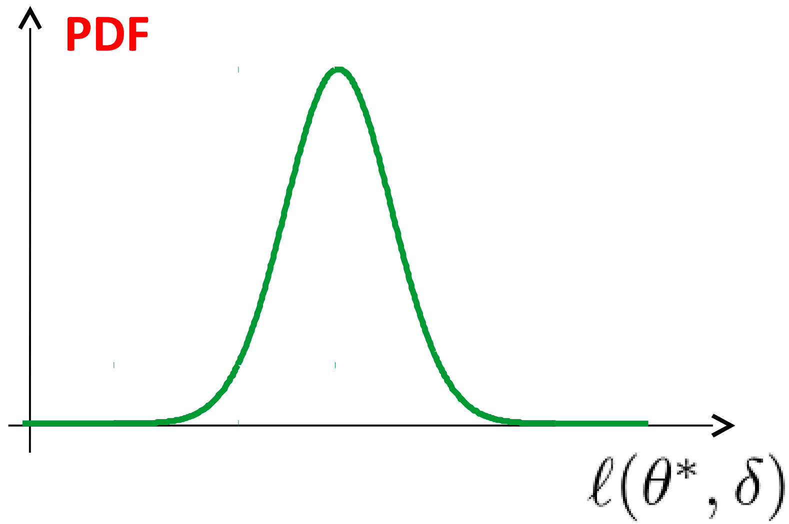


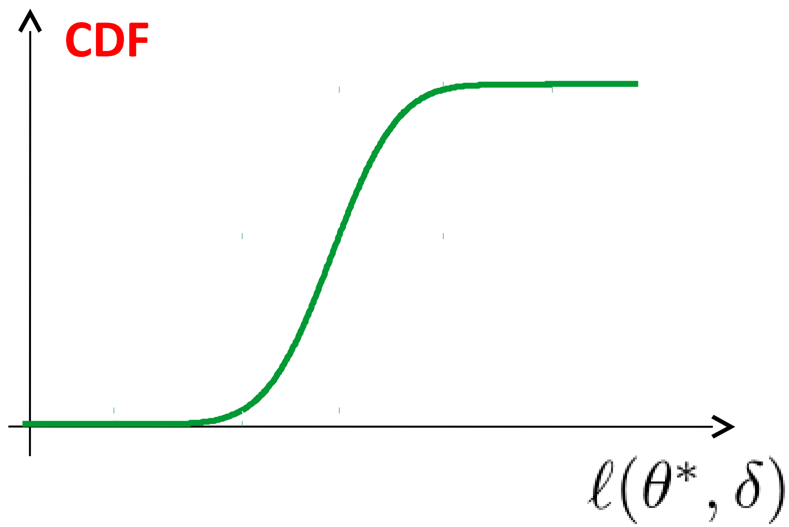


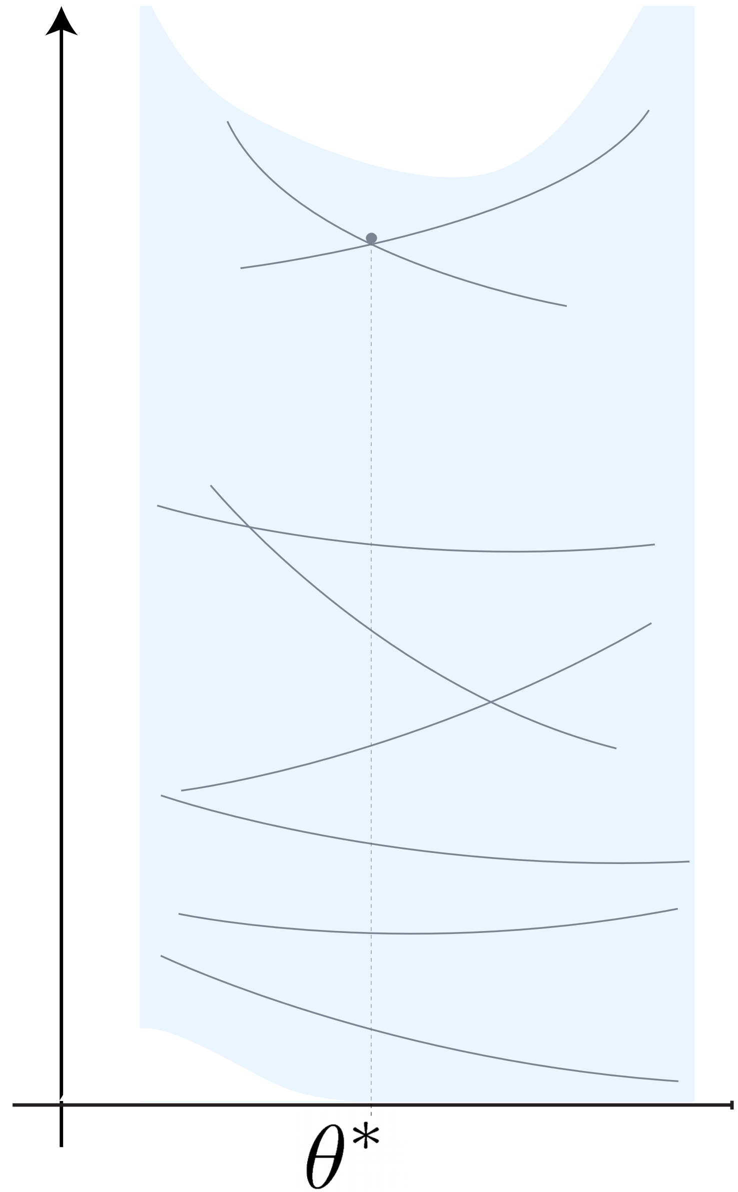
Probability distribution of $\ell(\theta^*, \delta)$



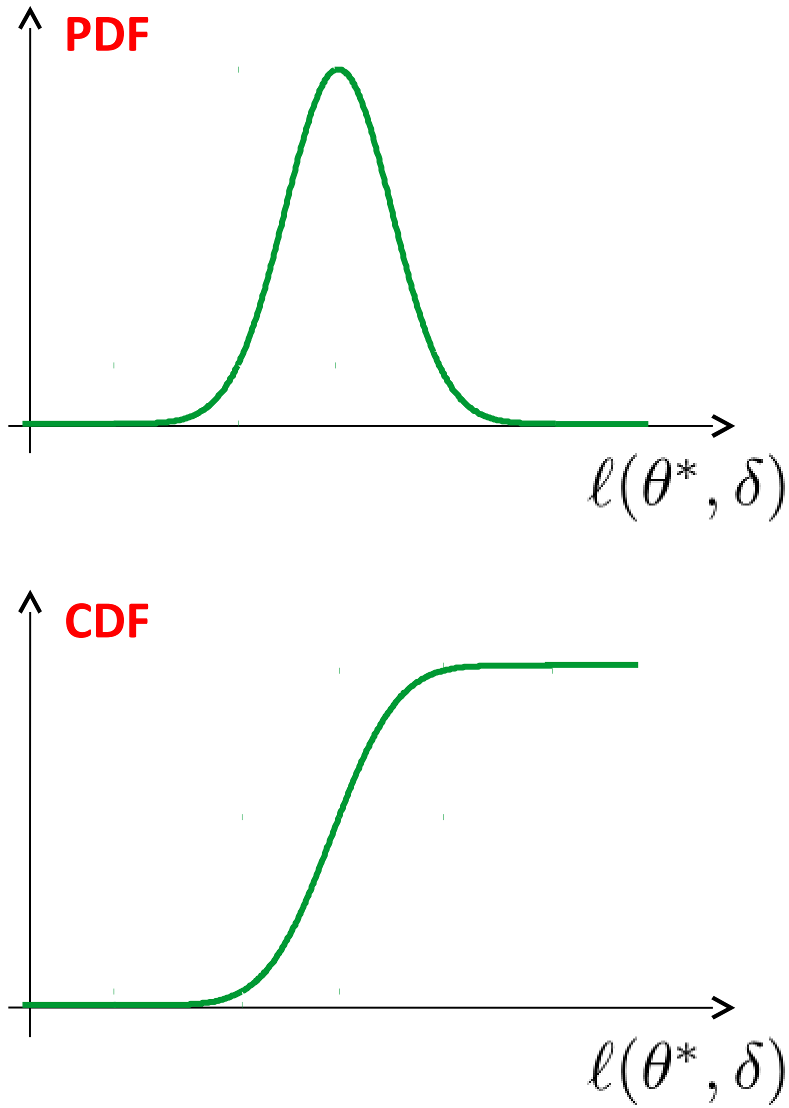
Probability distribution of $\ell(\theta^*, \delta)$

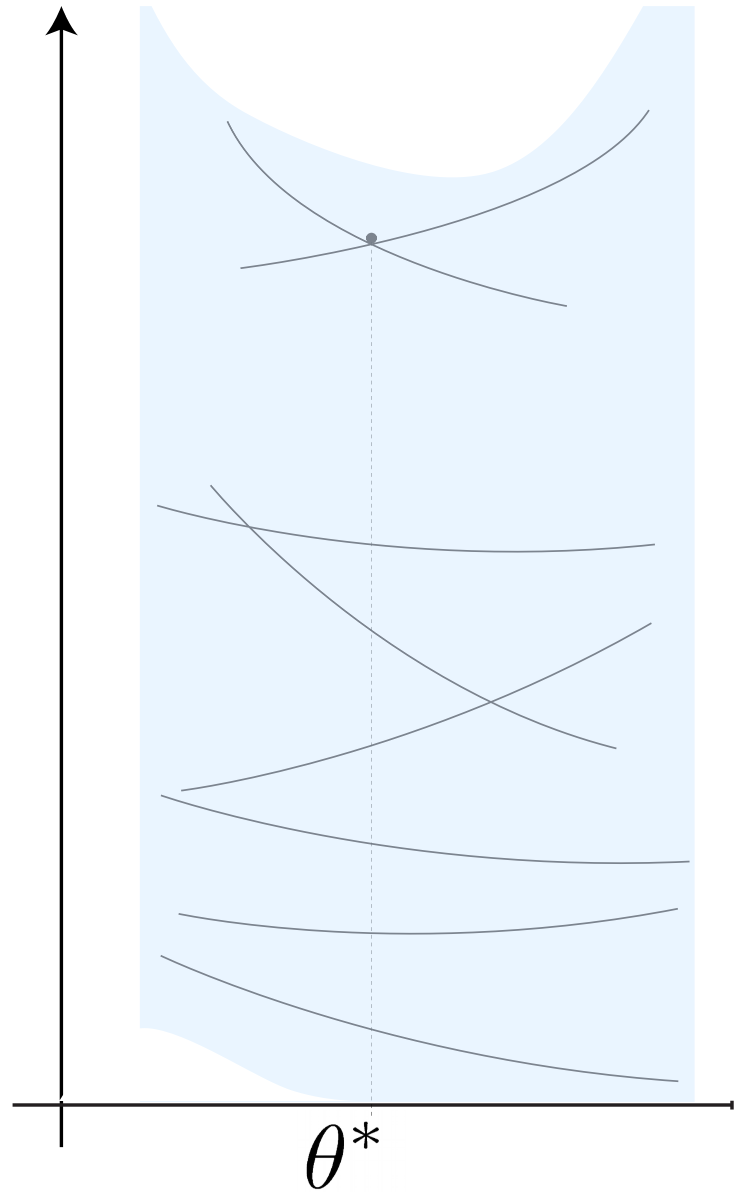




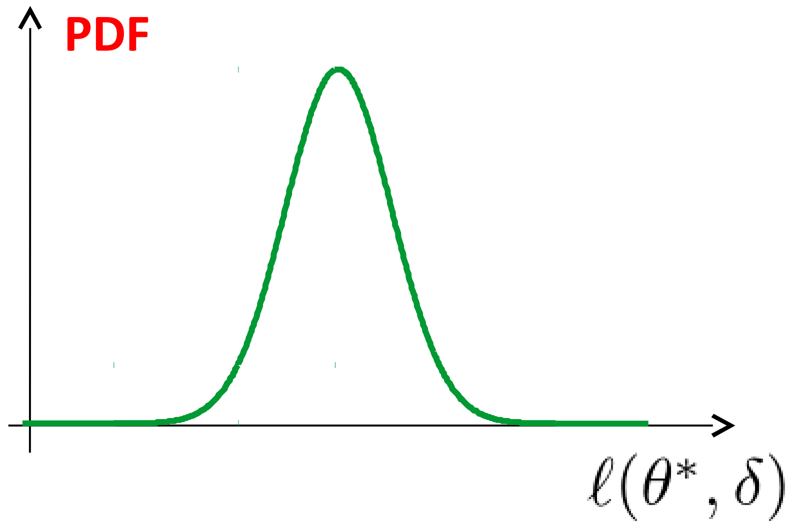


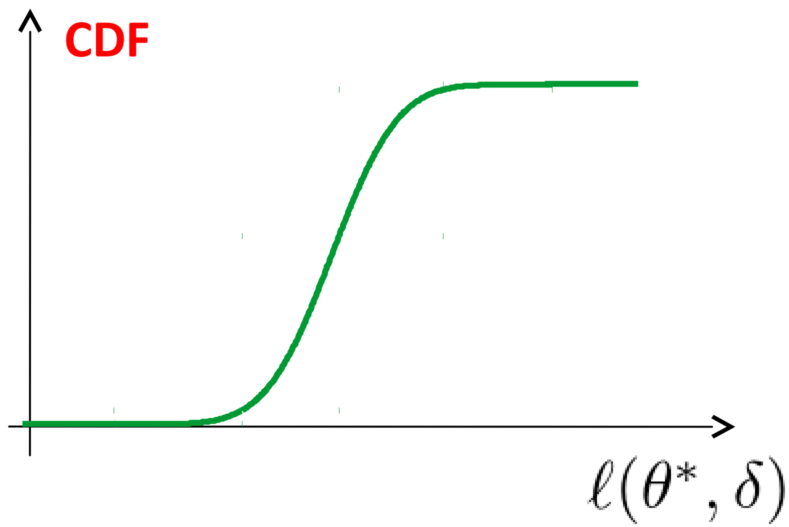
Probability distribution of $\ell(\theta^*, \delta)$ $\quad \delta \sim \mathbb{P}$

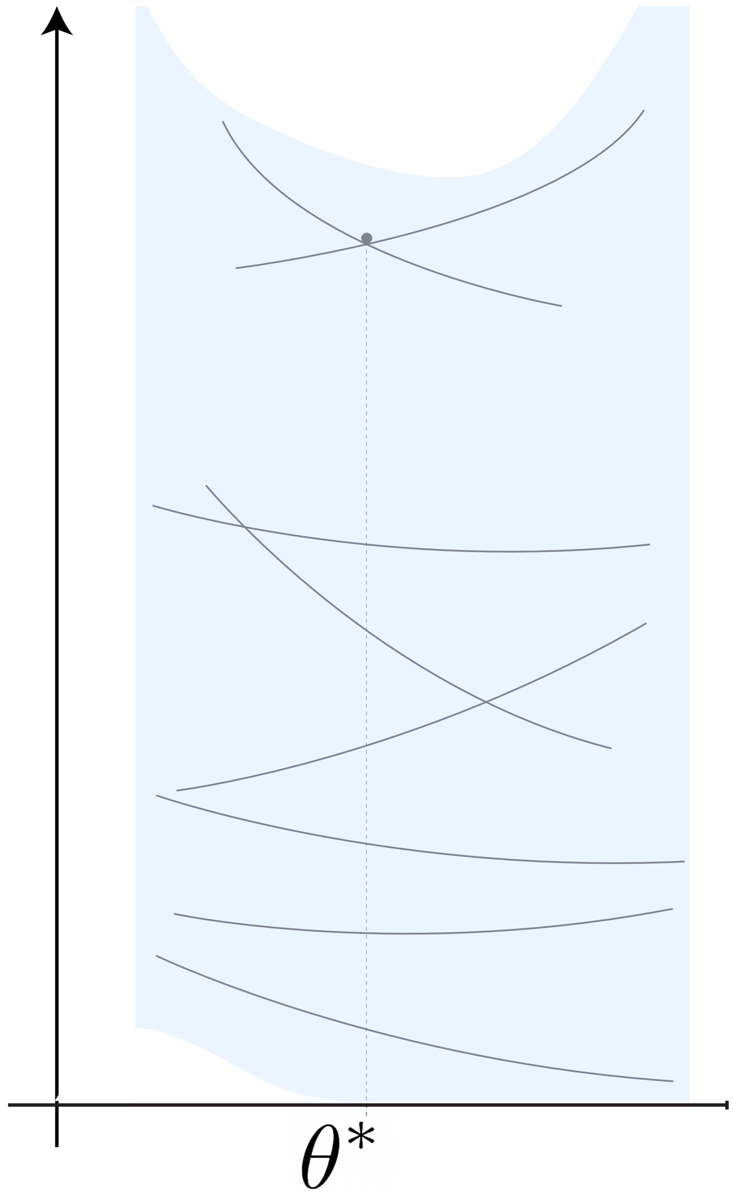




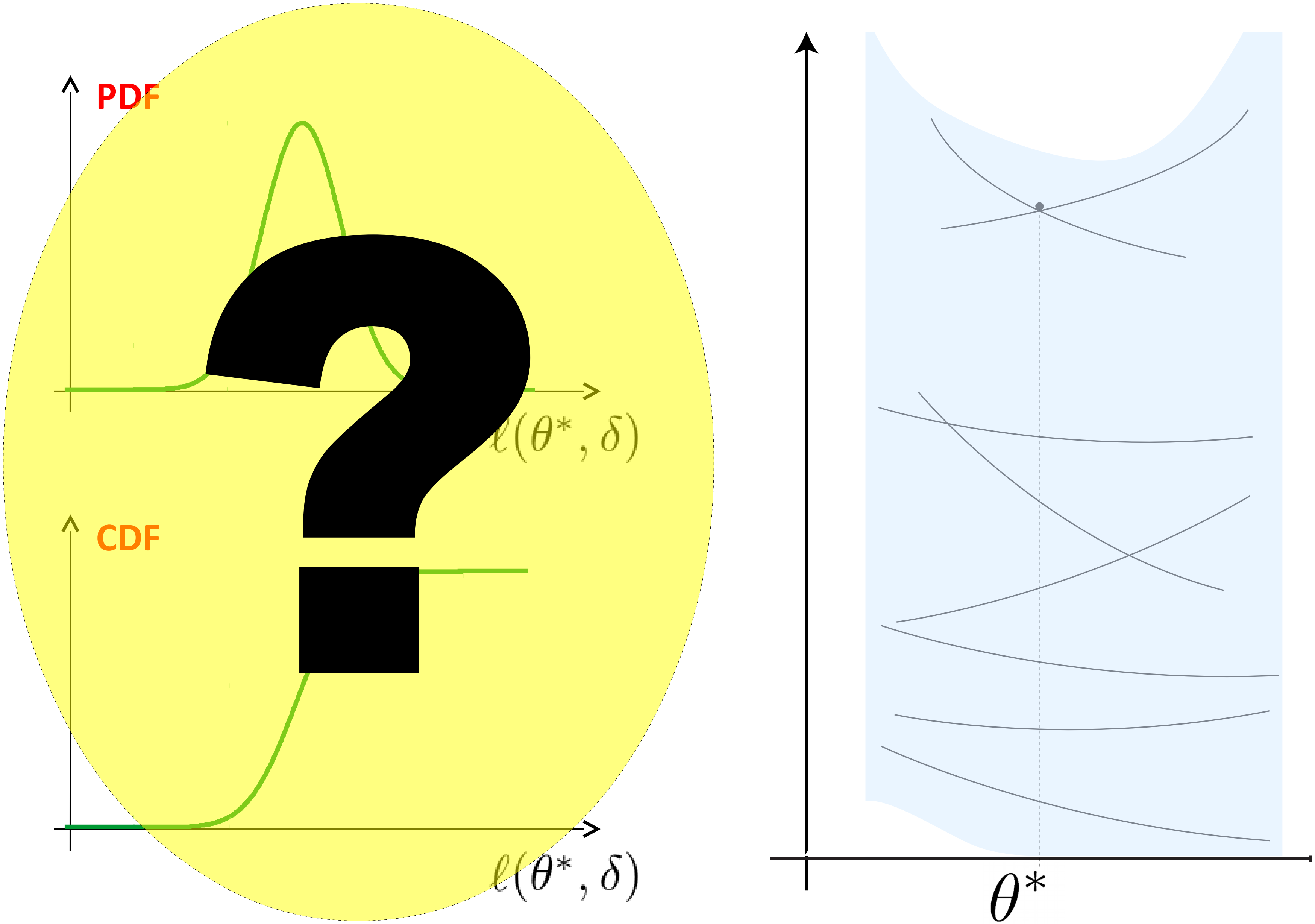
Probability distribution of $\ell(\theta^*, \delta)$ $\delta \sim \mathbb{P}$





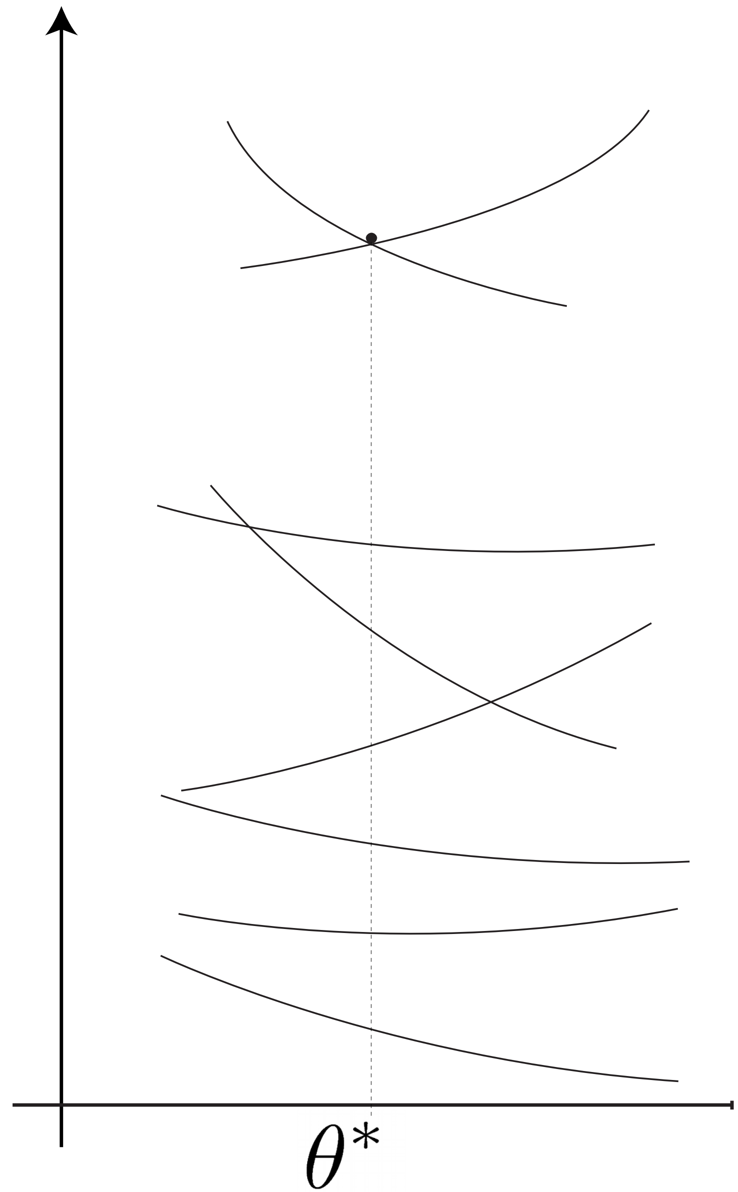


Probability distribution of $\ell(\theta^*, \delta)$ $\delta \sim \mathbb{P}$



Probability distribution of $\ell(\theta^*, \delta)$ $\quad \delta \sim \mathbb{P}$

?



TAKE-HOME MESSAGE:

It is possible to “reconstruct” the distribution of the cost

TAKE-HOME MESSAGE:

It is possible to “reconstruct”
the distribution of the cost
by using the sole N scenarios
that have been used to compute θ^*.

TAKE-HOME MESSAGE:

It is possible to "reconstruct"
the distribution of the cost
by using the sole N scenarios
that have been used to compute θ^*.

Without using any new observation
nor any specific knowledge of $\mathbb{P}$.

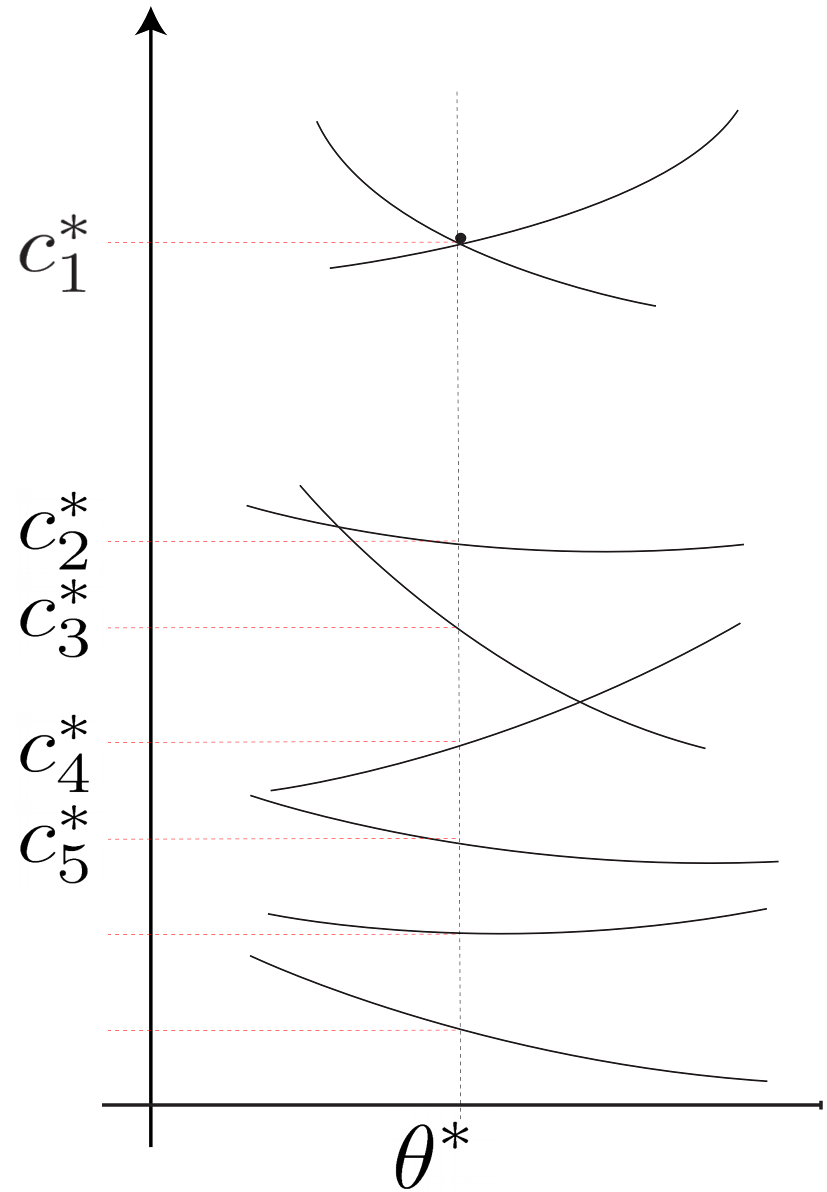
HOW?

HOW?

By combining a-posteriori knowledge

with distribution-free theorems

Empirical distribution of $\ell(\theta^*, \delta)$



HOW?

By combining a-posteriori knowledge
(the empirical distribution of the cost)

with distribution-free theorems

HOW?

By combining a-posteriori knowledge

(the empirical distribution of the cost)

with distribution-free theorems

(invariant properties of convex problems)

HOW?

By combining a-posteriori knowledge

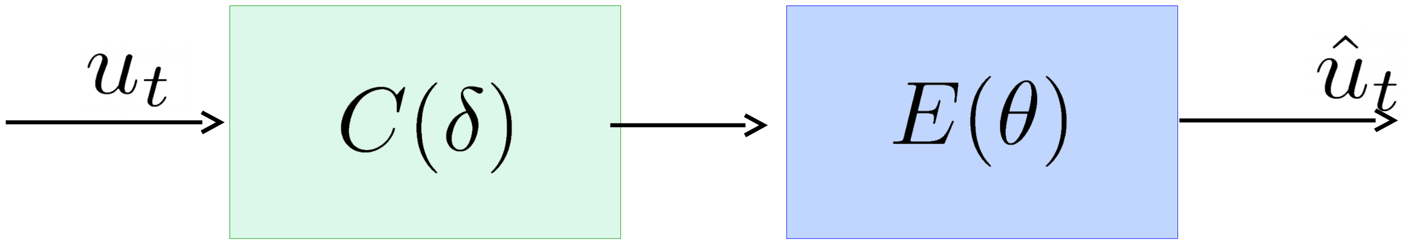
(the empirical distribution of the cost)

with distribution-free theorems

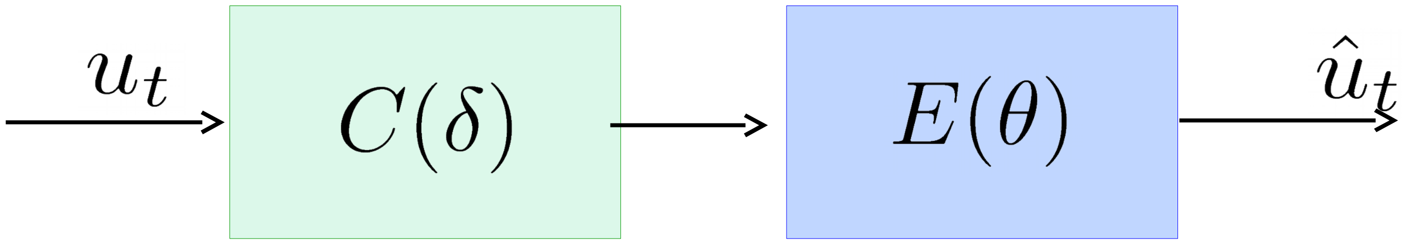
(invariant properties of convex problems)

example (and a few technical details) following...

Channel equalization example

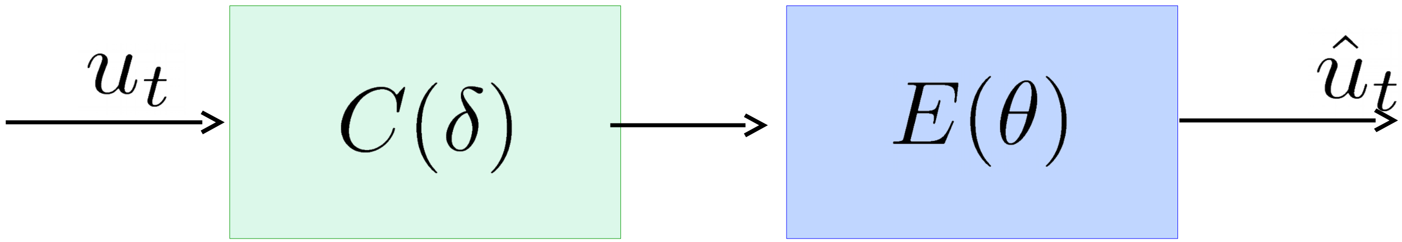


Channel equalization example



$$\ell(\theta, \delta) = \|C(\delta)E(\theta) - \text{IdealChannel}\|$$

Channel equalization example

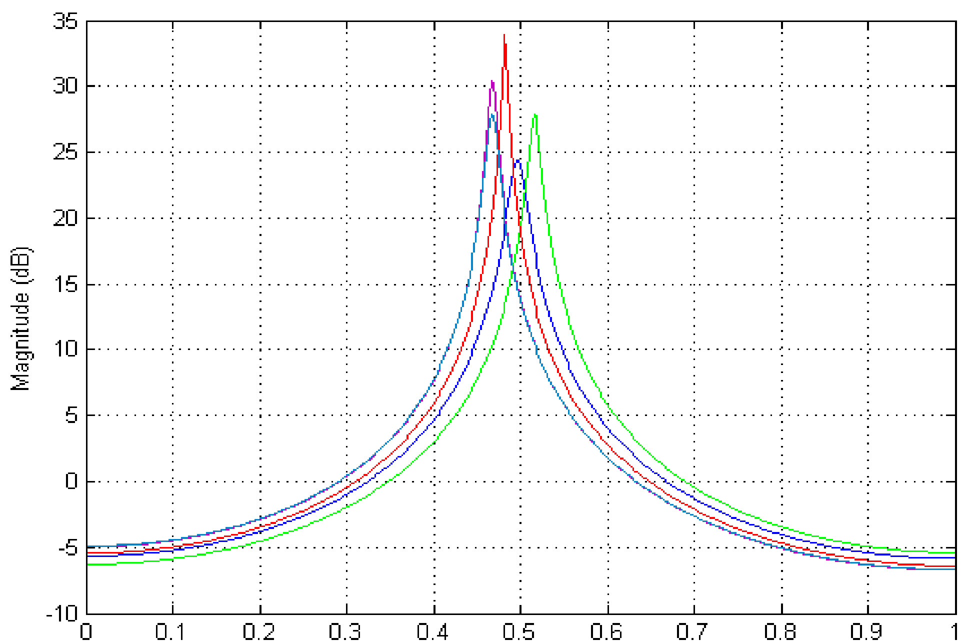


$$\ell(\theta, \delta) = \|C(\delta)E(\theta) - \mathrm{IdealChannel}\|$$

For details: see paper

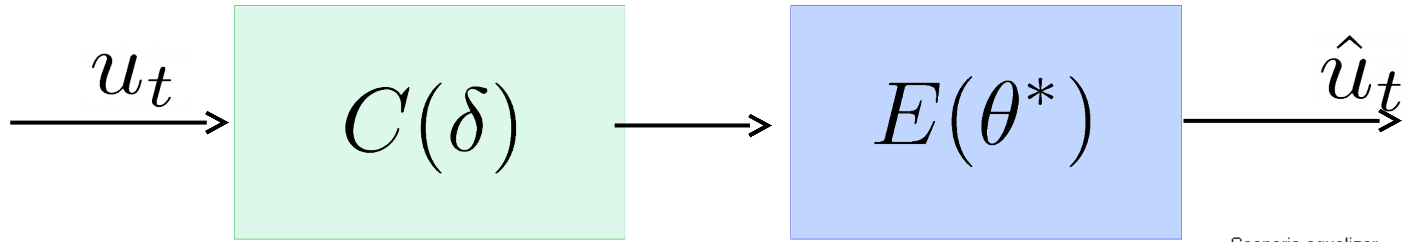
Channel equalization example

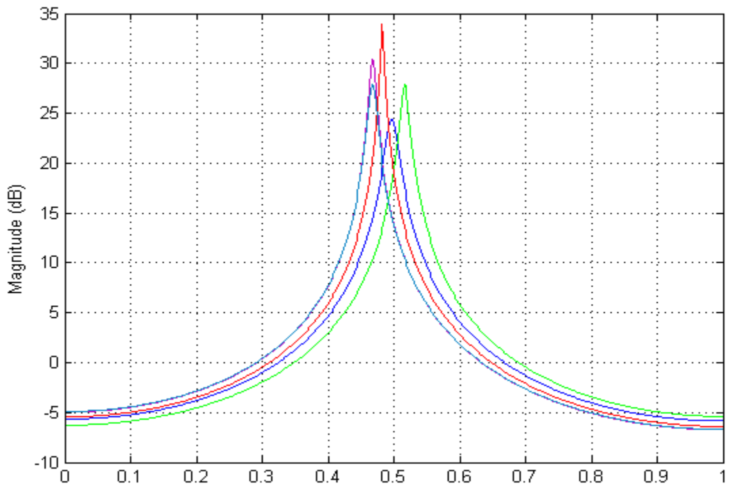
$$C(\delta)$$



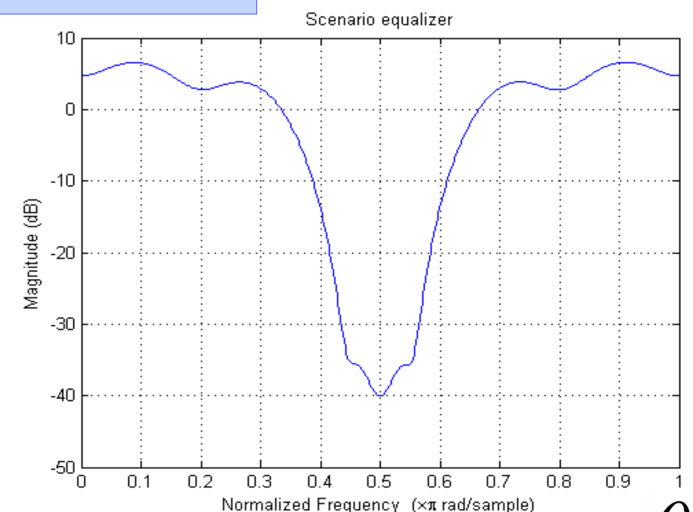
$$\delta^{(1)}, \delta^{(2)}, \ldots, \delta^{(N)}$$

Channel equalization example



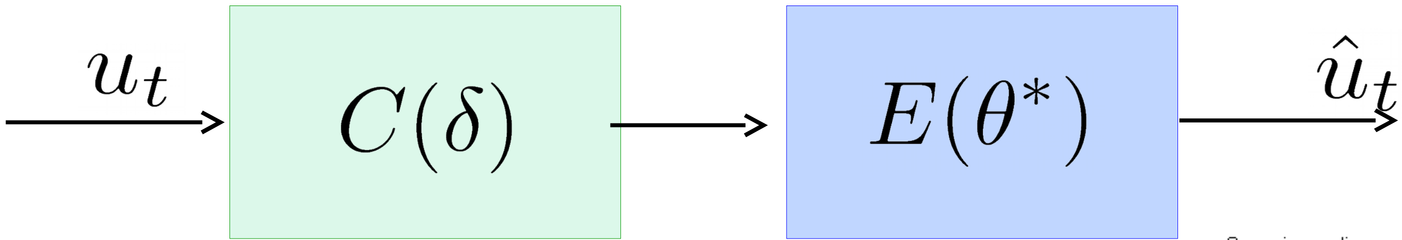


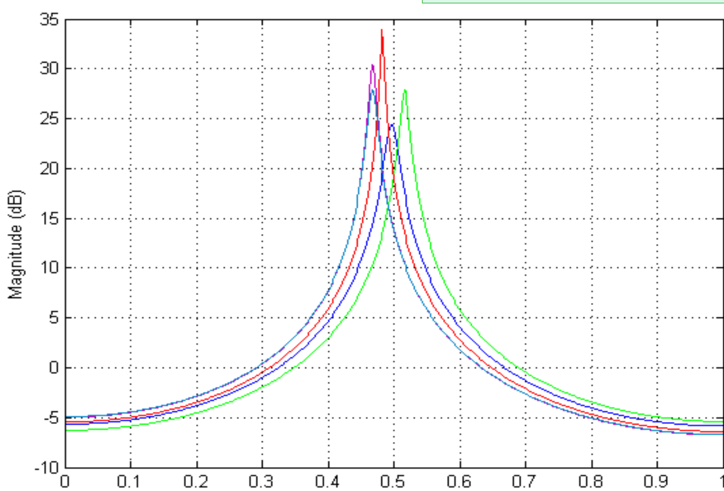
$\delta^{(1)}, \delta^{(2)}, \ldots, \delta^{(N)}$



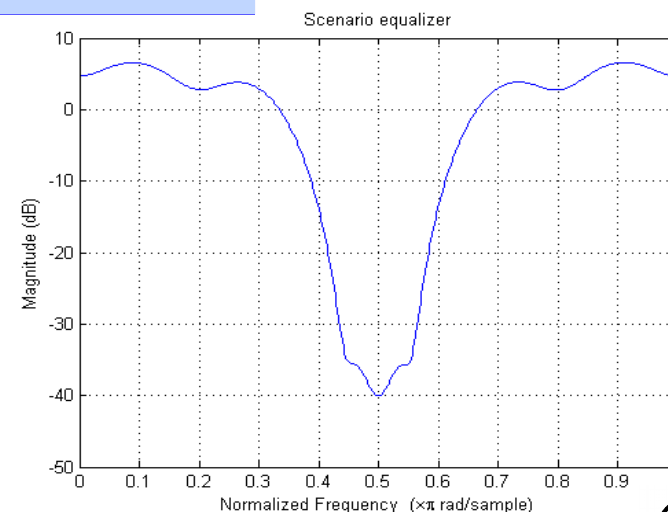
θ^*

Channel equalization example



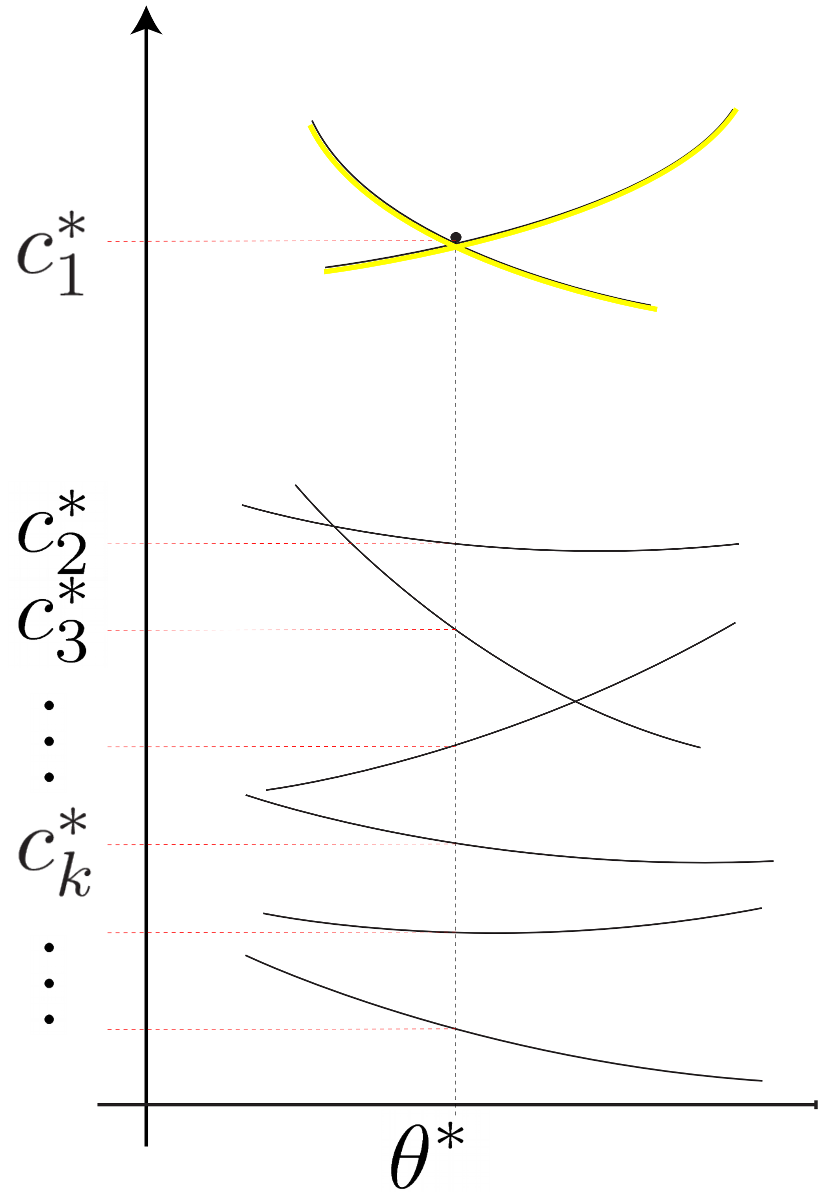


$\delta^{(1)}, \delta^{(2)}, \ldots, \delta^{(N)}$



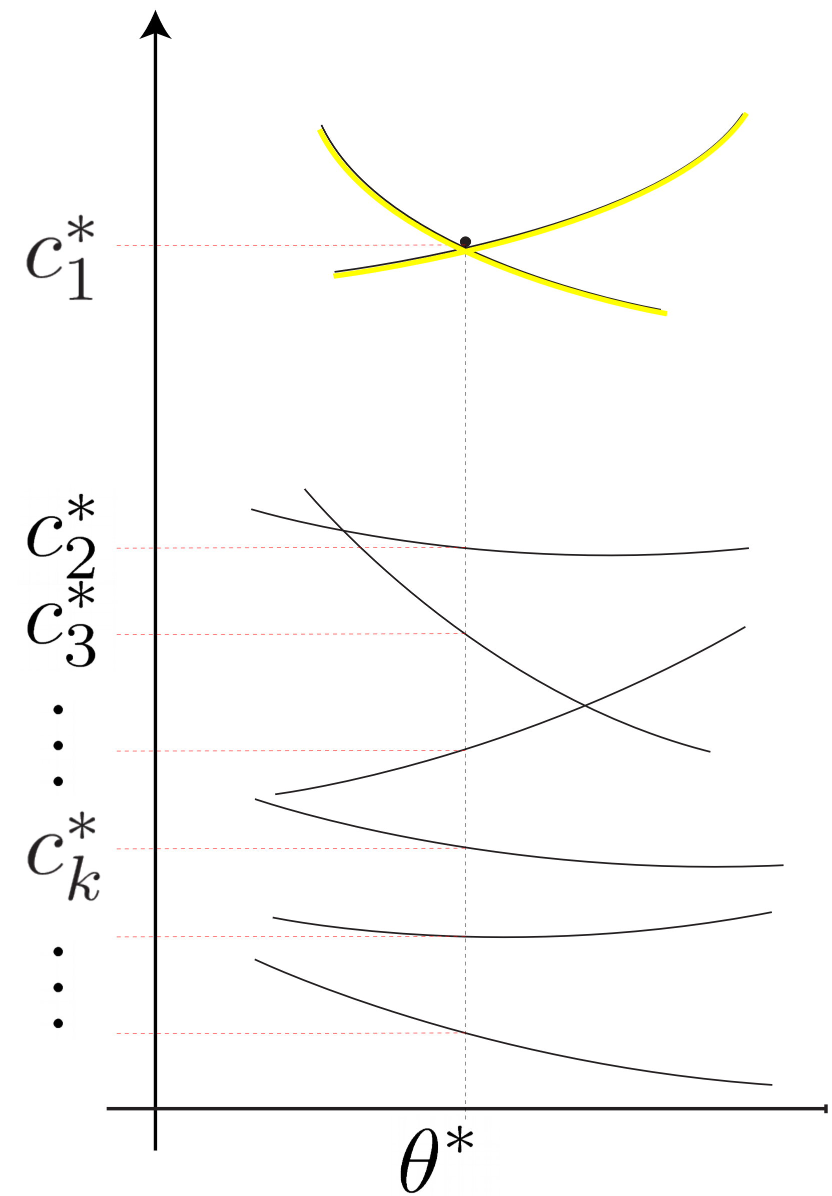
θ^*

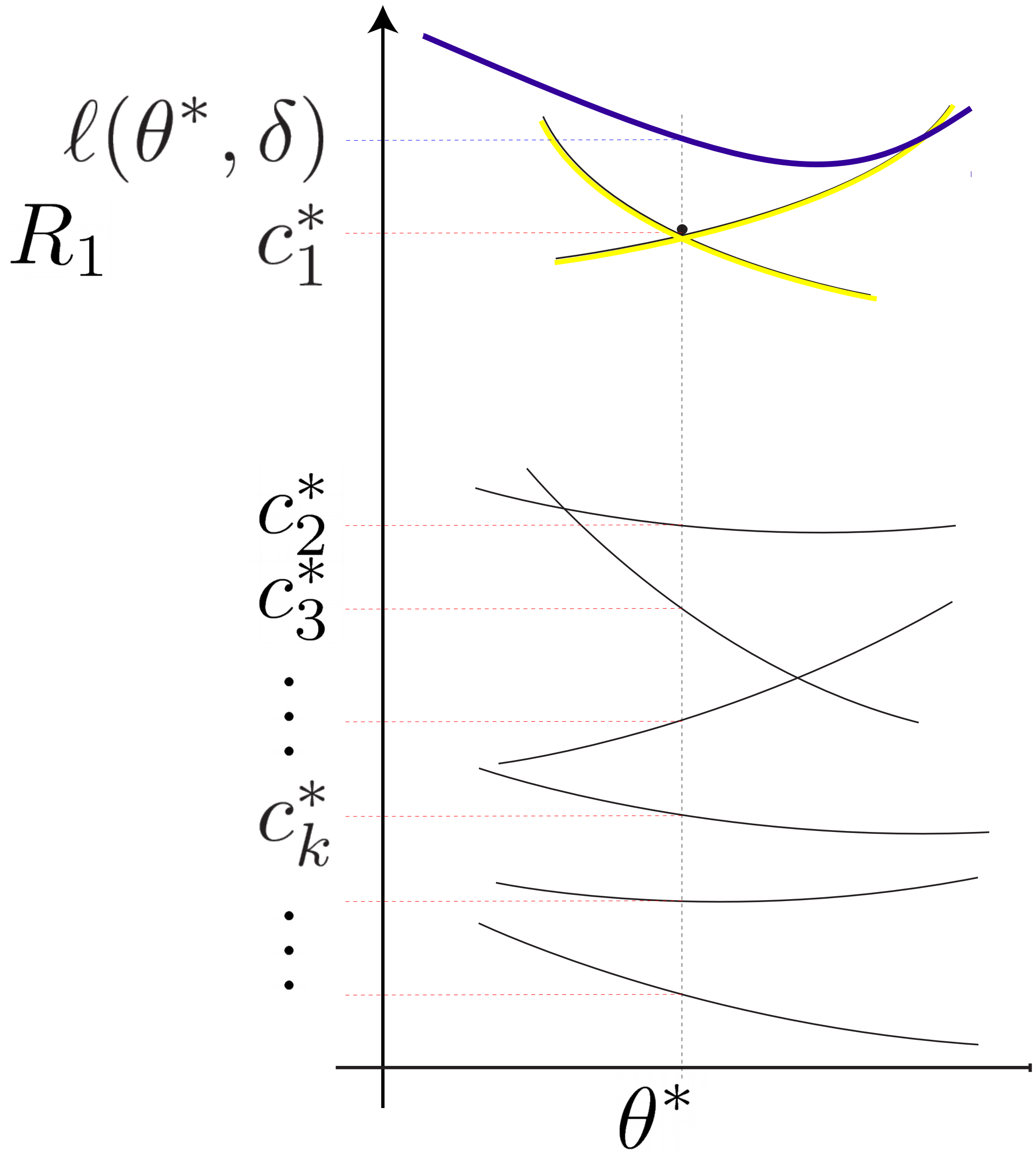
$$\ell(\theta^*, \delta^{(1)}), \ell(\theta^*, \delta^{(2)}), \ldots, \ell(\theta^*, \delta^{(N)})$$



c_1^*
c_2^*
c_3^*
c_k^*
θ^*

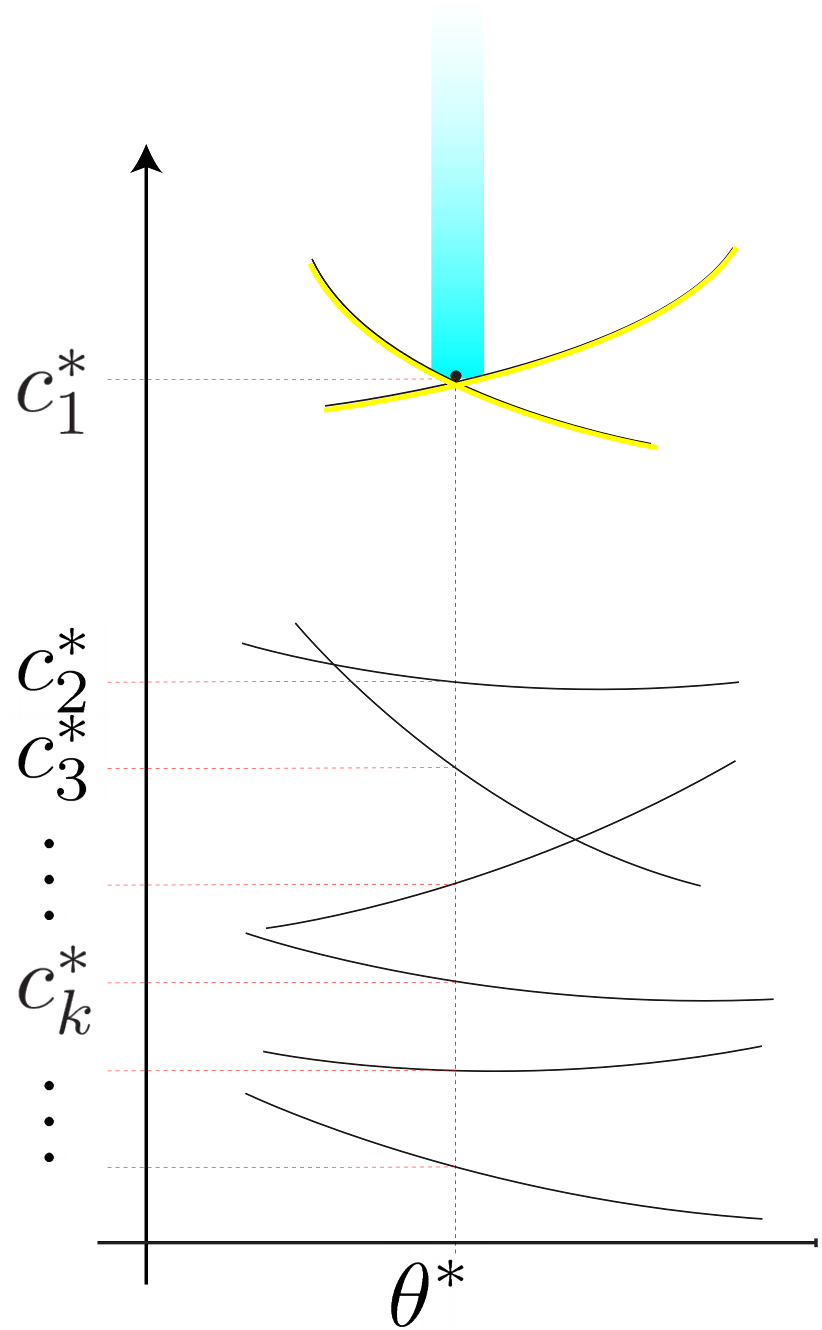
R_1



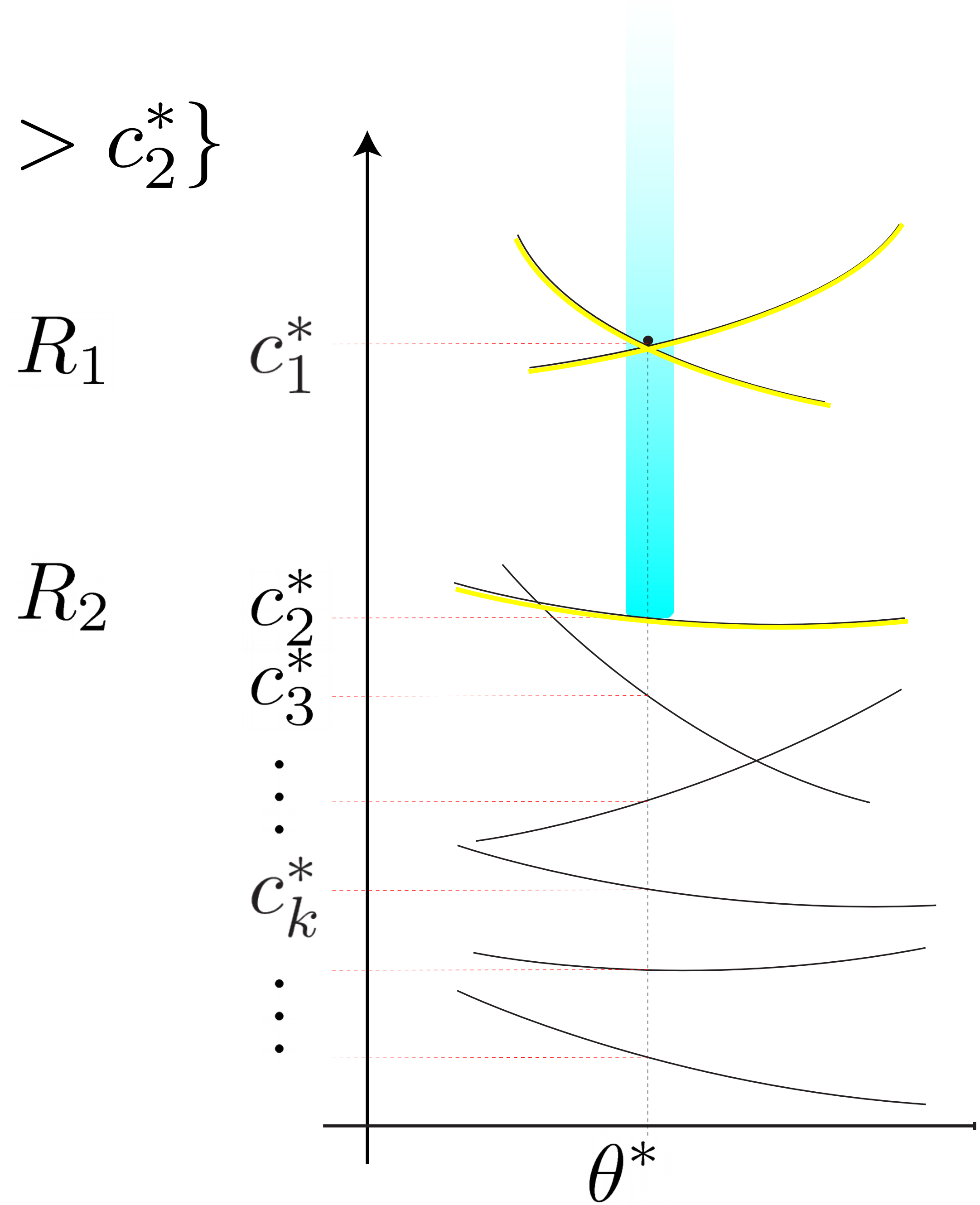

R_1
$\ell(\theta^*, \delta)$
c_1^*
c_2^*
c_3^*
c_k^*
θ^*

$$R_1 = \mathbb{P}\{\ell(\theta^*, \delta) > c_1^*\}$$

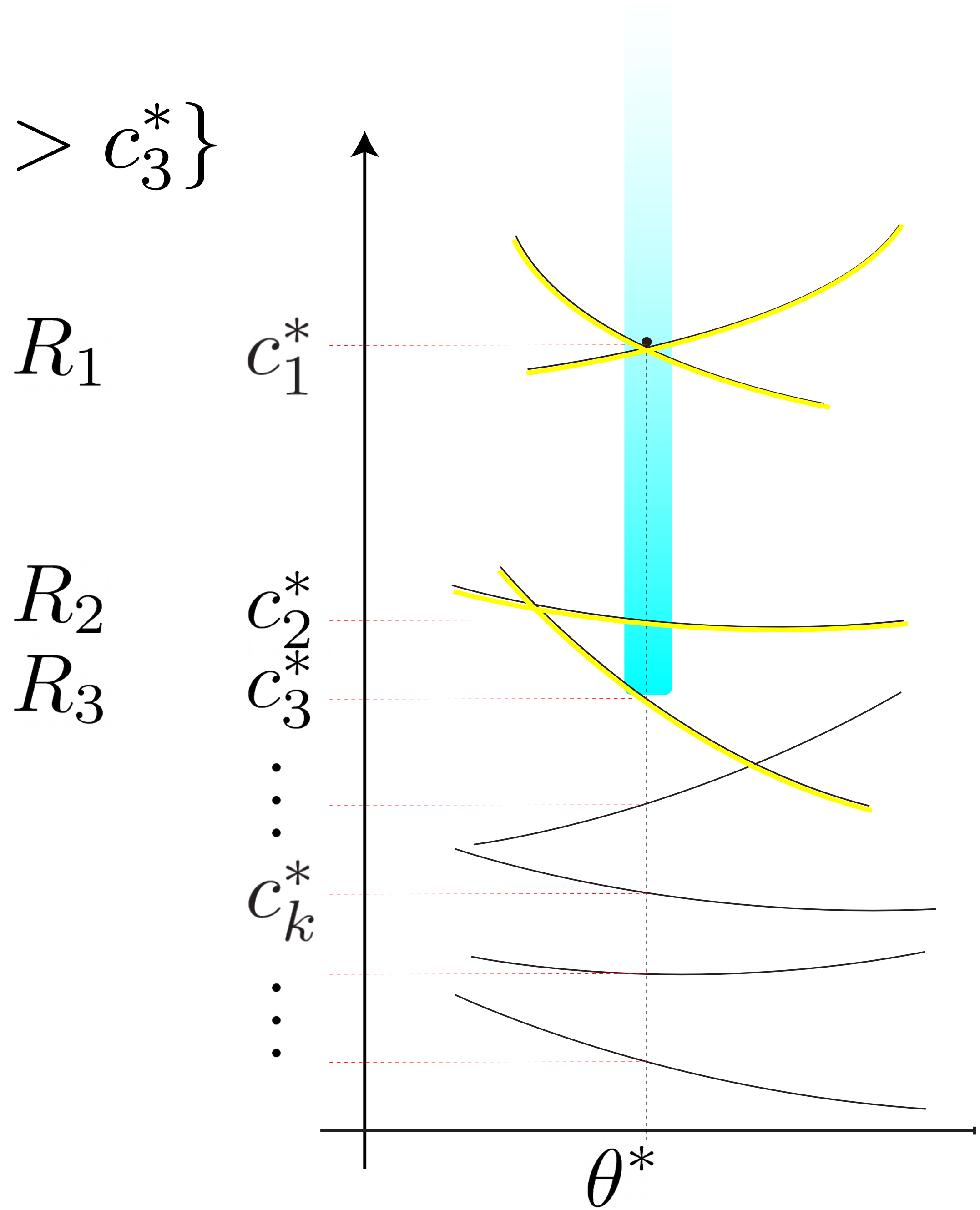
R_1



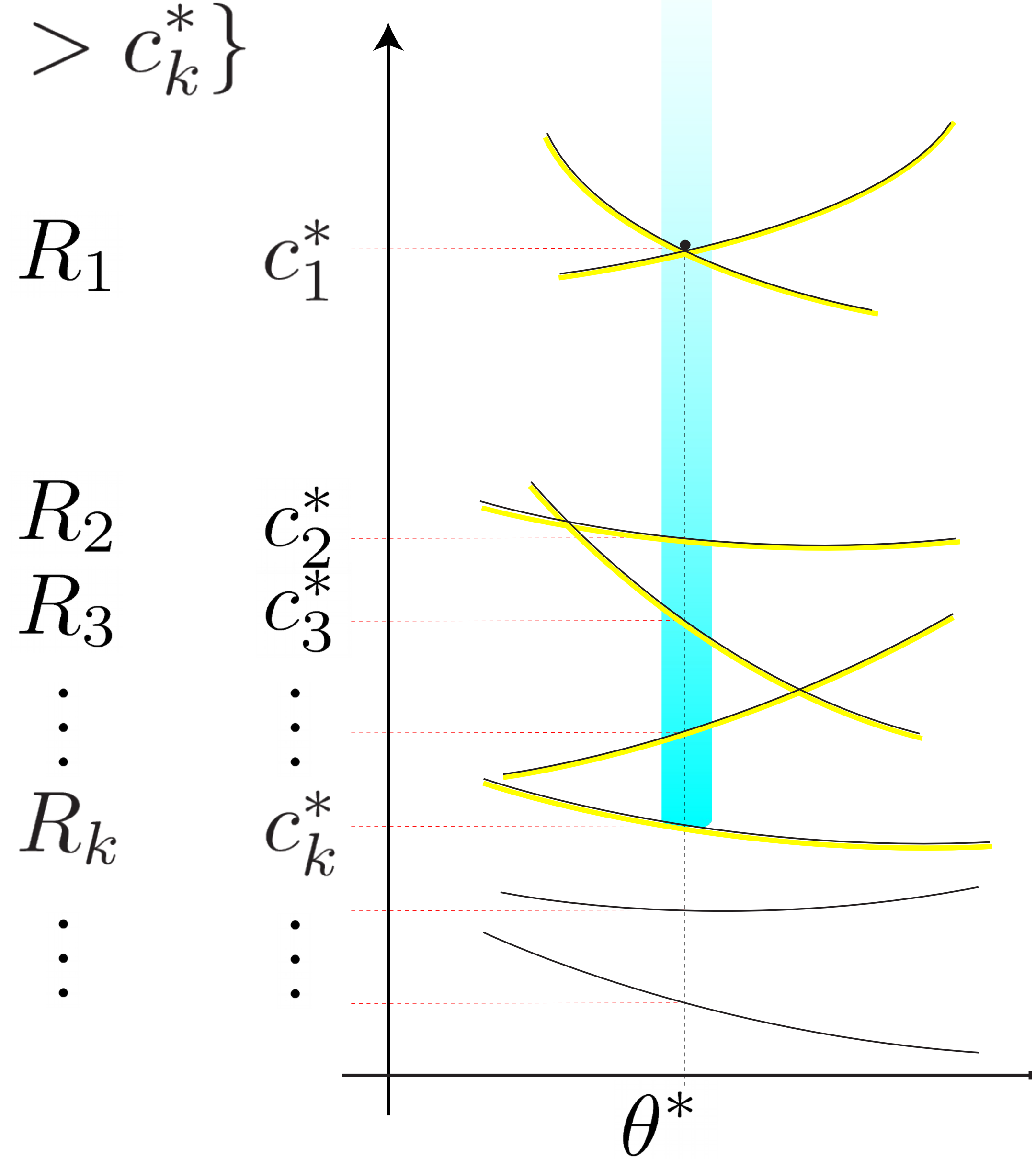
$$R_2 = \mathbb{P}\{\ell(\theta^*, \delta) > c_2^*\}$$



$$R_3 = \mathbb{P}\{\ell(\theta^*, \delta) > c_3^*\}$$



$$R_k = \mathbb{P}\{\ell(\theta^*, \delta) > c_k^*\}$$



$$R_k = \mathbb{P}\{\ell(\theta^*, \delta) > c_k^*\}$$

R_1

R_2

R_3

$\vdots$

R_k

$\vdots$

$$R_k = \mathbb{P}\{\ell(\theta^*, \delta) > c_k^*\}$$

$$R_1$$
$$R_2$$
$$R_3$$
$$\vdots$$
$$R_k$$
$$\vdots$$

$$R_k = \mathbb{P}\{\ell(\theta^*, \delta) > c_k^*\}$$

$$R_1$$
$$R_2$$
$$R_3$$
$$\vdots$$
$$R_k$$
$$\vdots$$

N and **d** are all that matters

$$R_k = \mathbb{P}\{\ell(\theta^*, \delta) > c_k^*\}$$

$$R_1$$

$$R_2$$

$$R_3$$

$$\vdots$$

$$R_k \in [\underline{\epsilon}_k, \bar{\epsilon}_k] \qquad R_k$$

$$\vdots$$

N and **d** are all that matters

$$R_k = \mathbb{P}\{\ell(\theta^*, \delta) > c_k^*\}$$

$$R_1$$

$$R_2$$

$$R_3$$

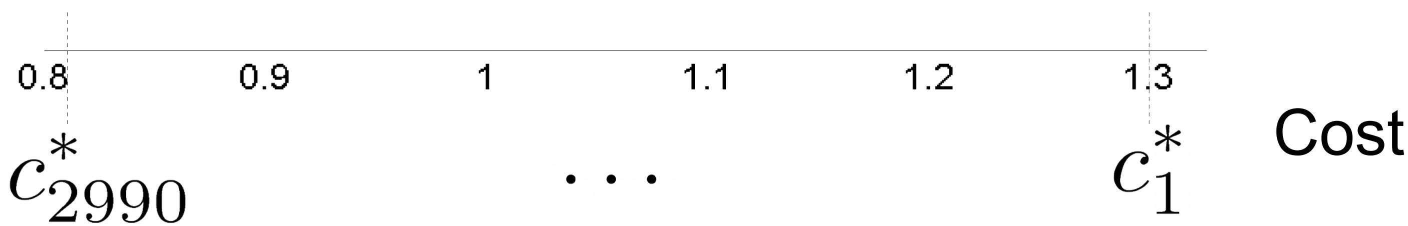
$$\vdots$$

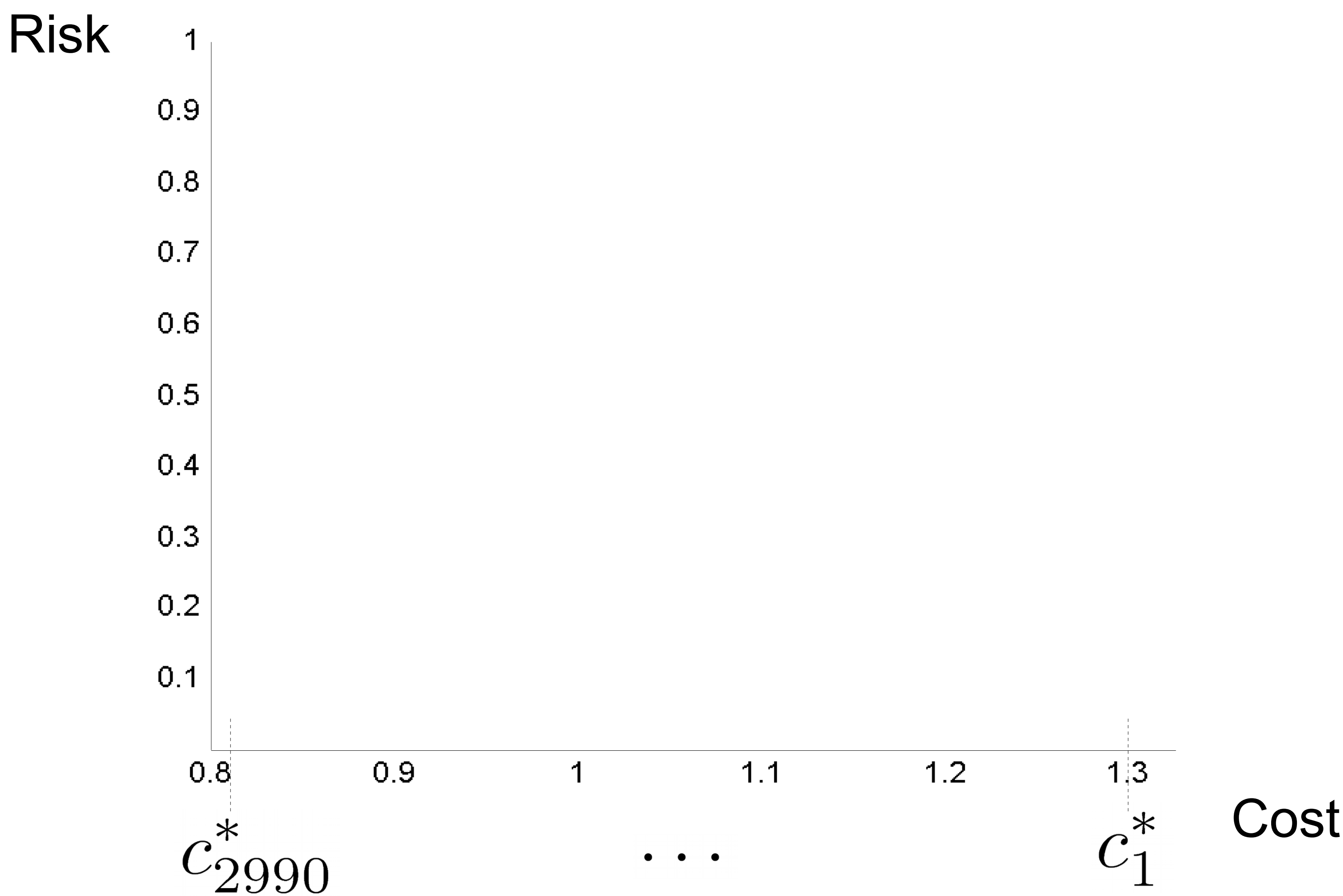
$$R_k \in [\underline{\epsilon}_k, \bar{\epsilon}_k] \qquad R_k$$

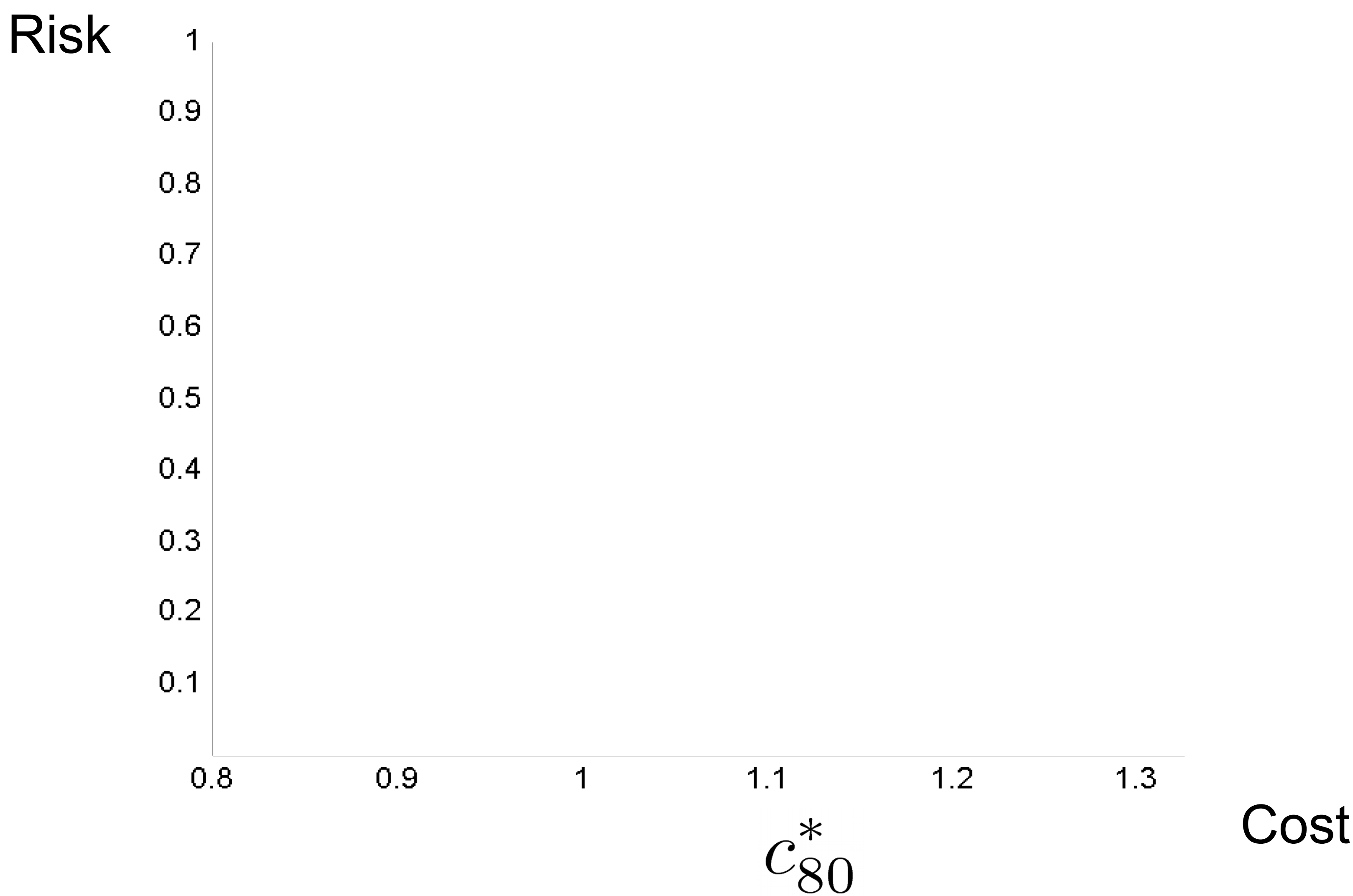
with confidence 1-10^{-6}

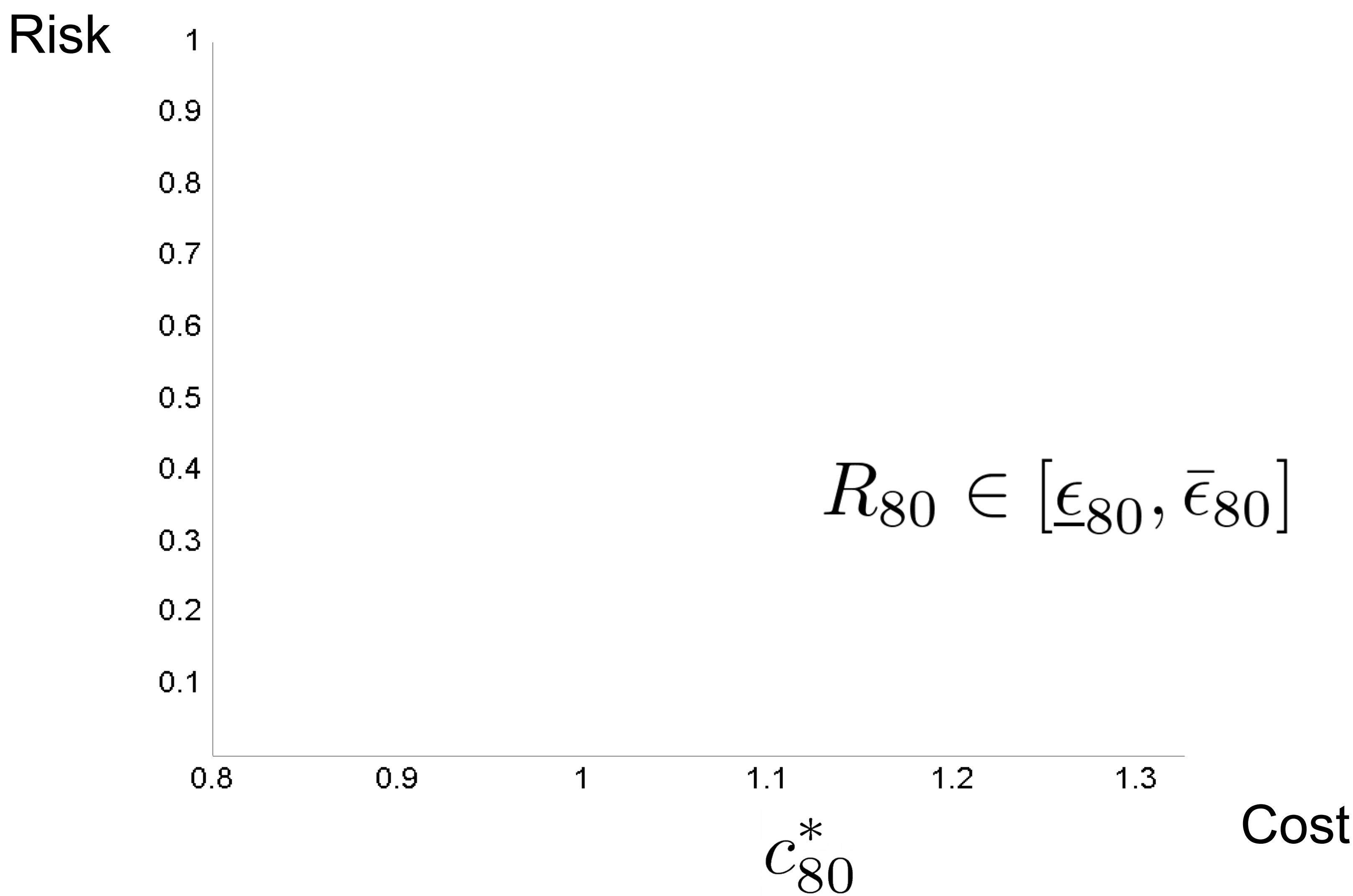
$$\vdots$$

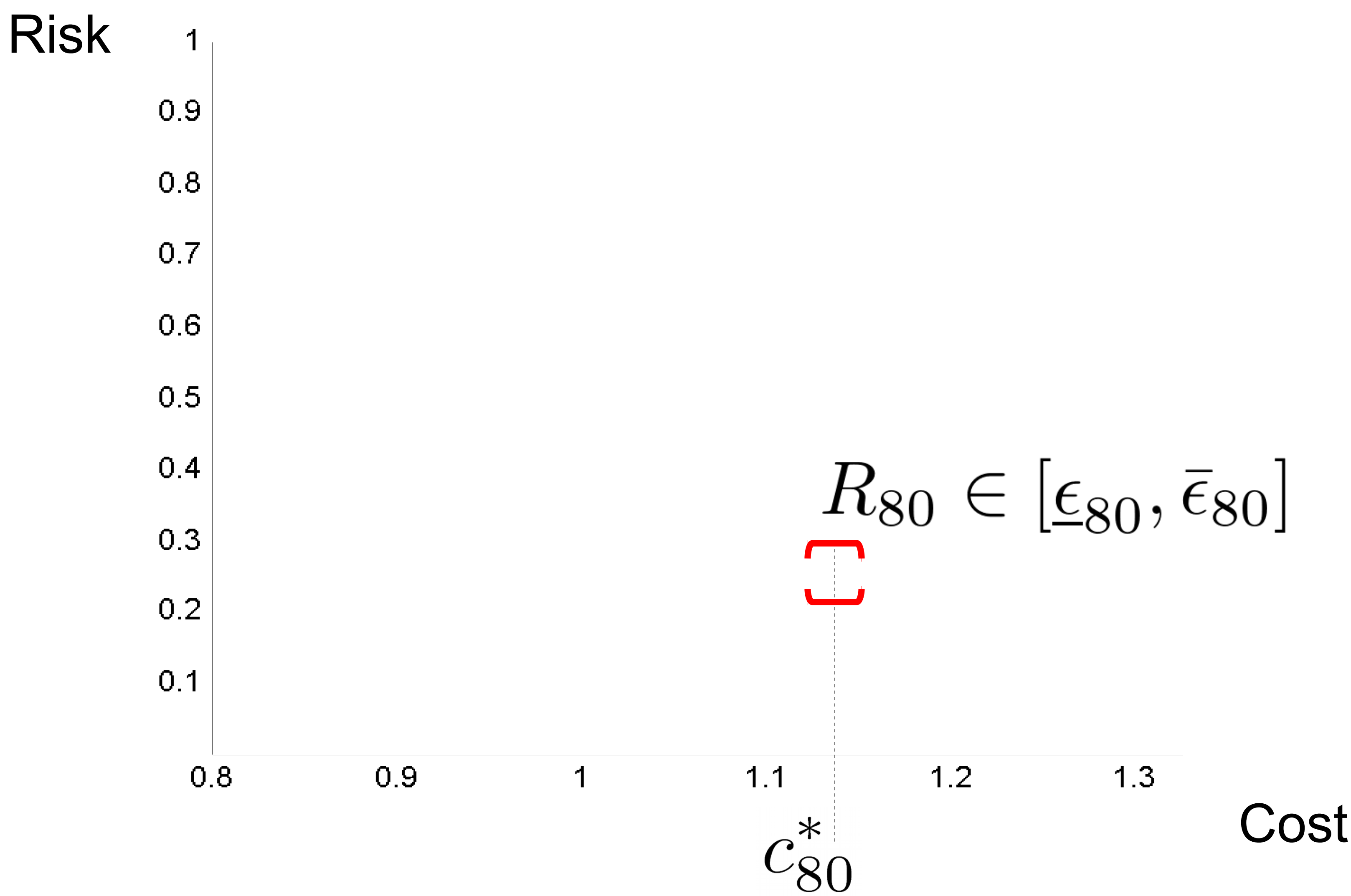
N and **d** are all that matters


0.8
0.9
1
1.1
1.2
1.3
Cost
c^*_{2990}
. . .
c^*_1

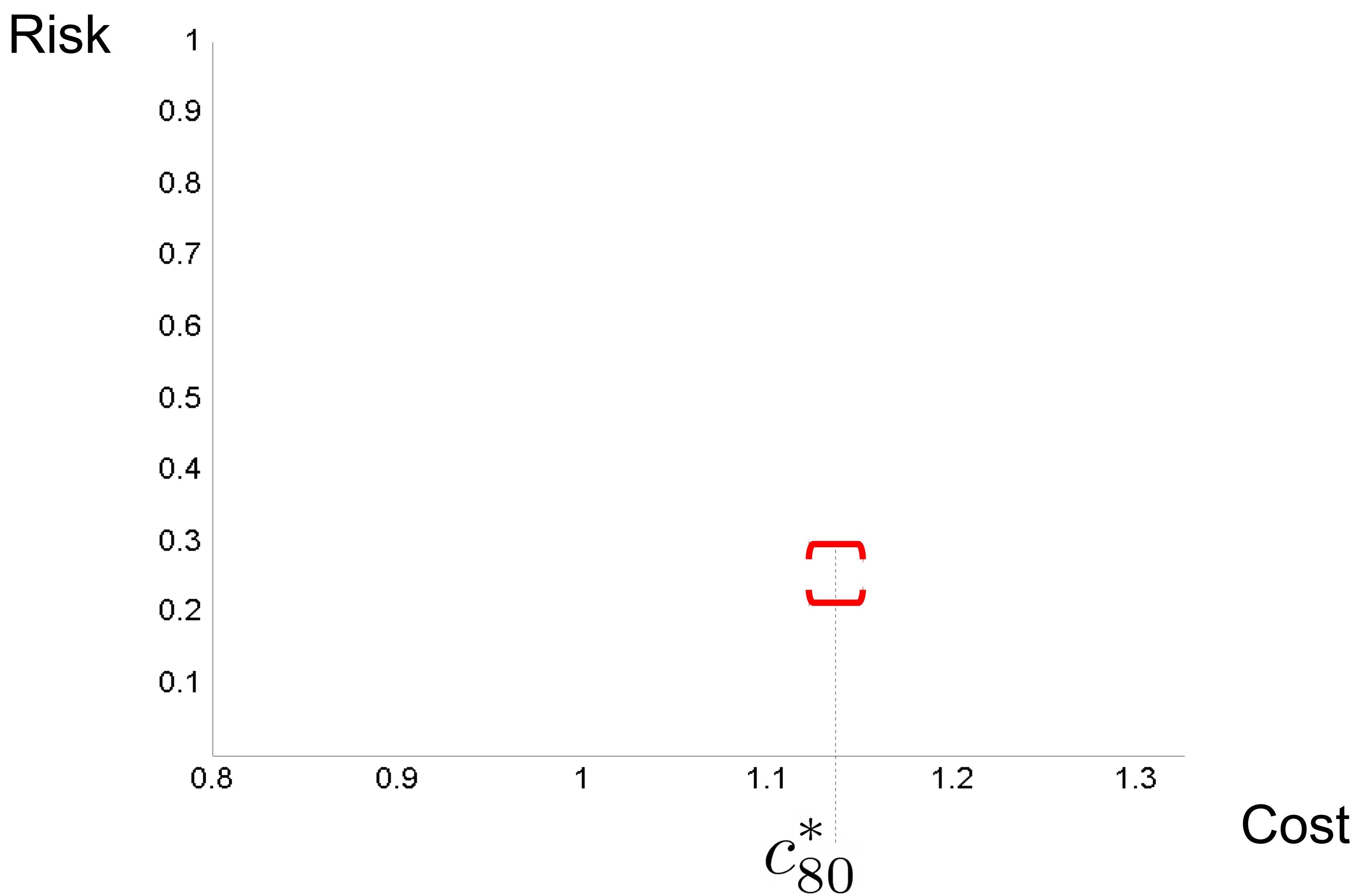

Risk
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.8
0.9
1
1.1
1.2
1.3
Cost
c^*_{2990}
$\cdots$
c^*_1

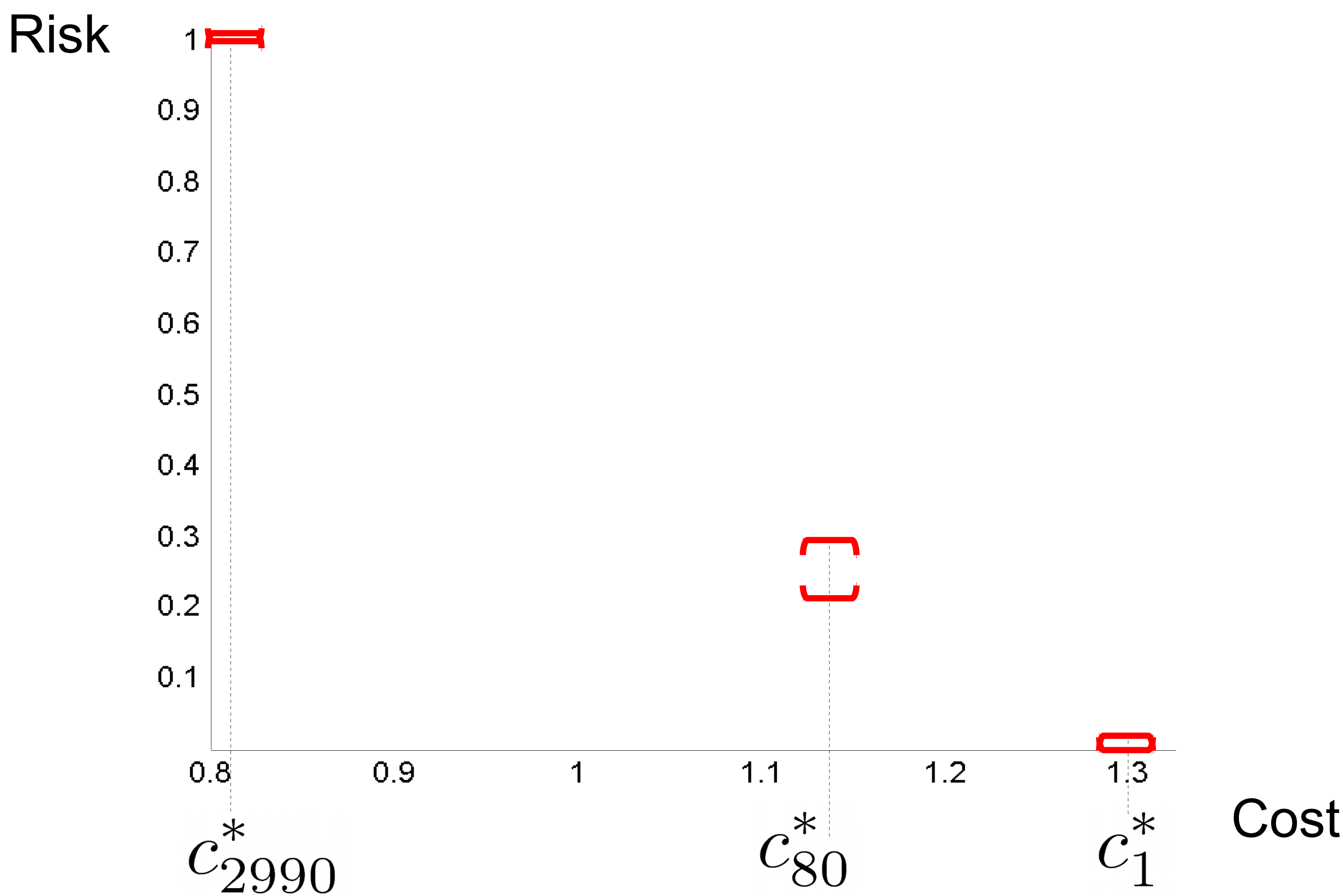

Risk
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.8
0.9
1
1.1
1.2
1.3
Cost
c_{80}^*


Risk
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.8
0.9
1
1.1
1.2
1.3
Cost
$R_{80} \in [\underline{\epsilon}_{80}, \bar{\epsilon}_{80}]$
c^*_{80}

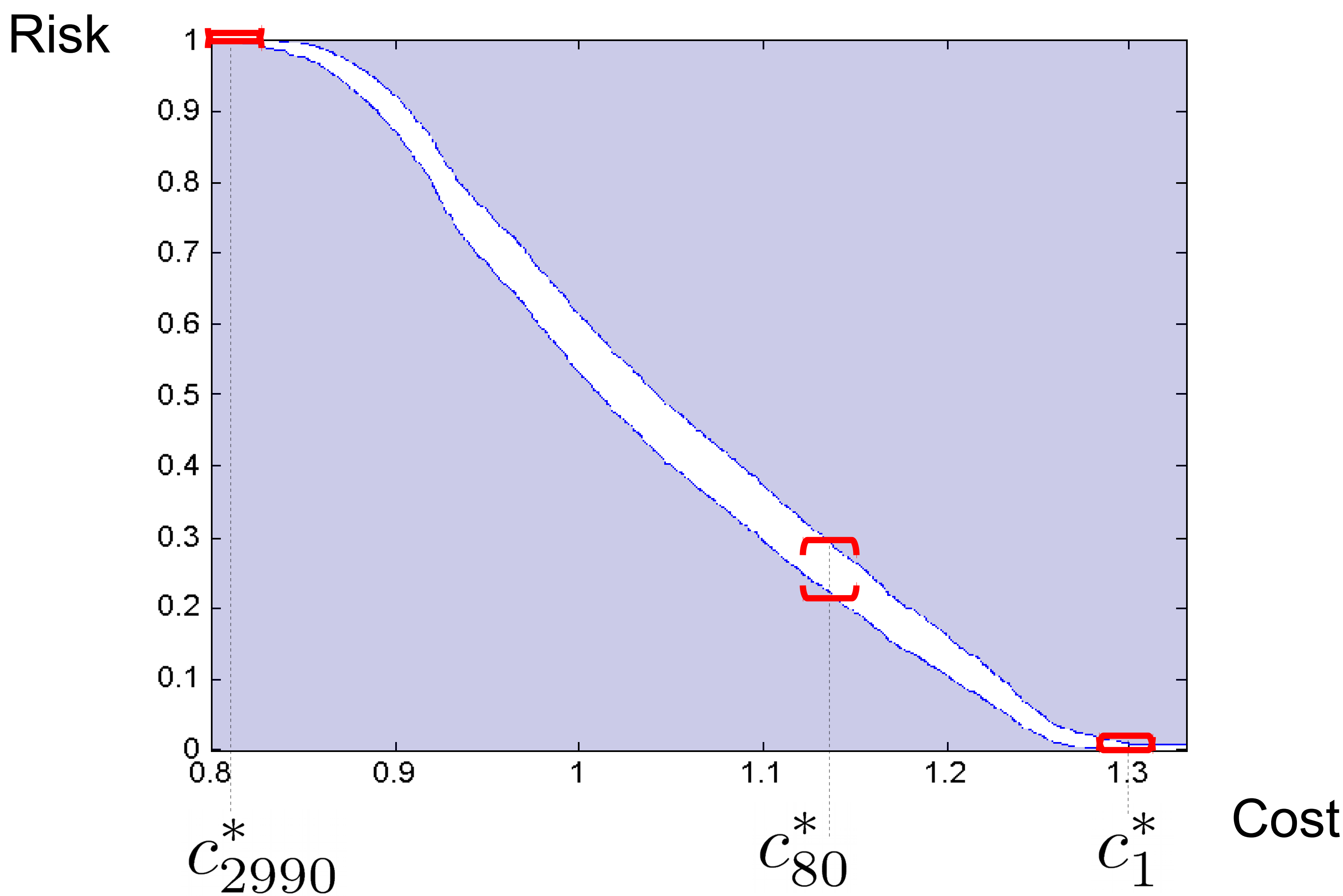


Risk
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
$R_{80} \in [\underline{\epsilon}_{80}, \bar{\epsilon}_{80}]$
0.8
0.9
1
1.1
1.2
1.3
Cost
c^*_{80}

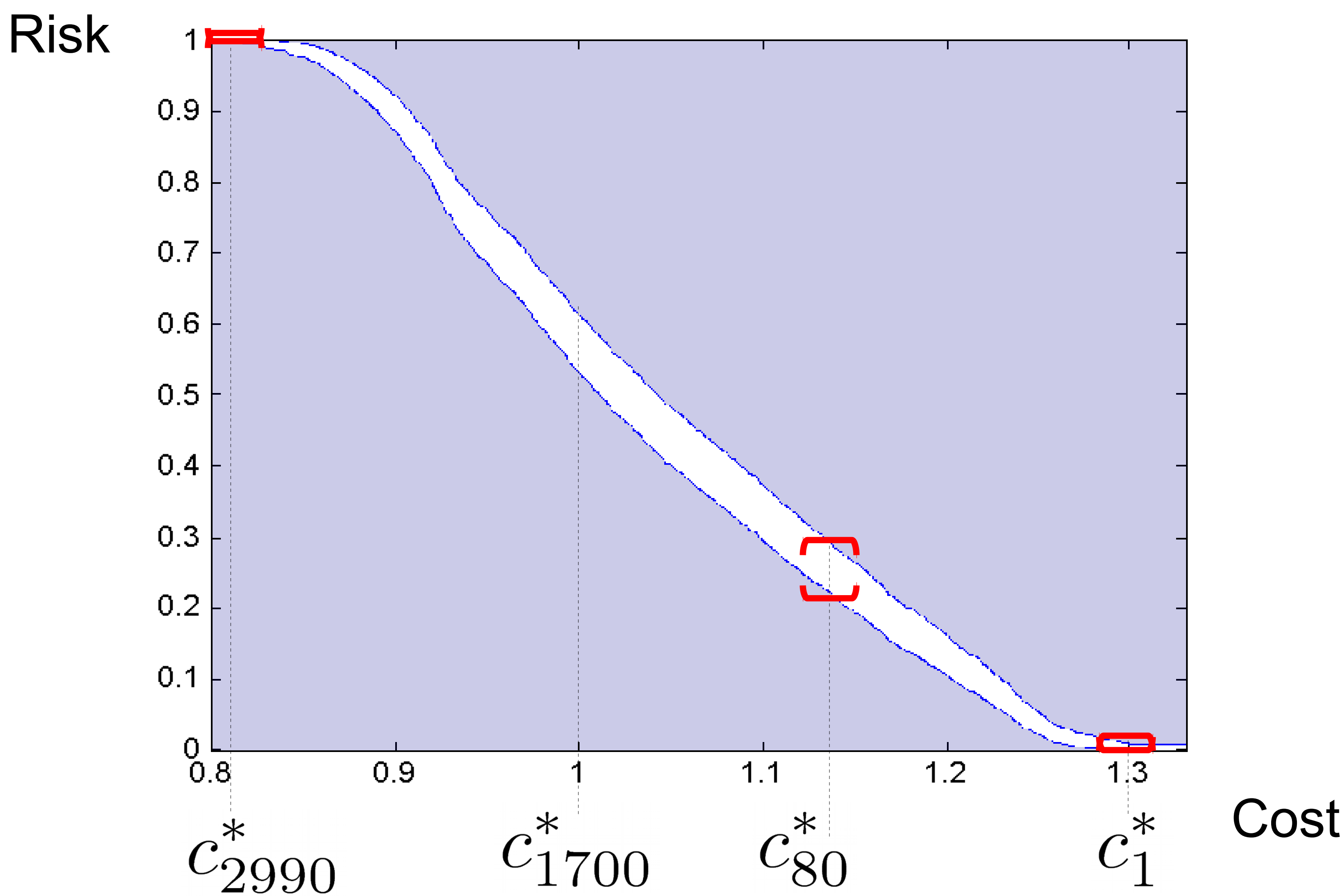

Risk
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.8
0.9
1
1.1
1.2
1.3
Cost
c^*_{80}



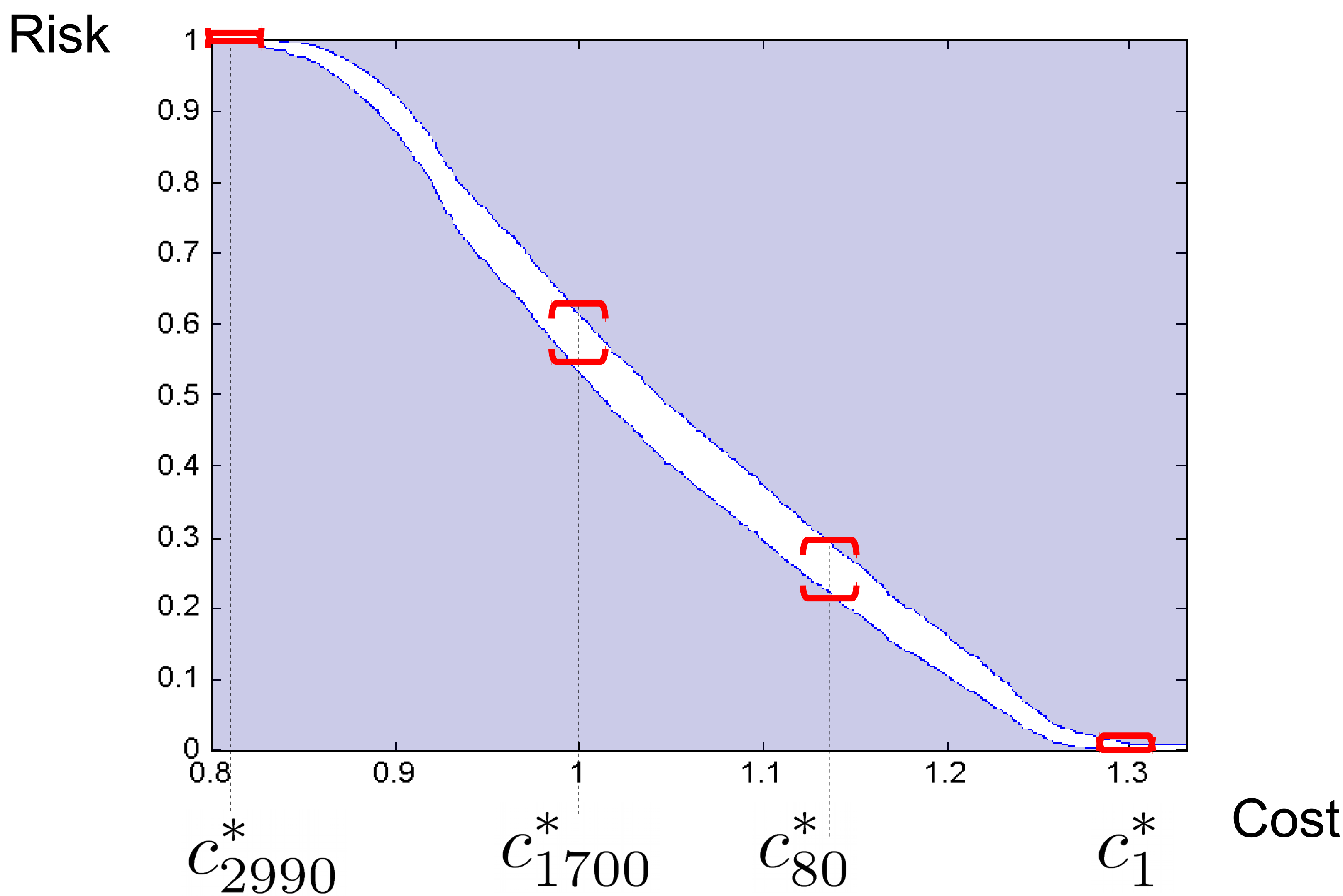
Risk
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.8
0.9
1
1.1
1.2
1.3
Cost
c^*_{2990}
c^*_{80}
c^*_1



Risk
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.8
0.9
1
1.1
1.2
1.3
Cost
c^*_{2990}
c^*_{80}
c^*_1

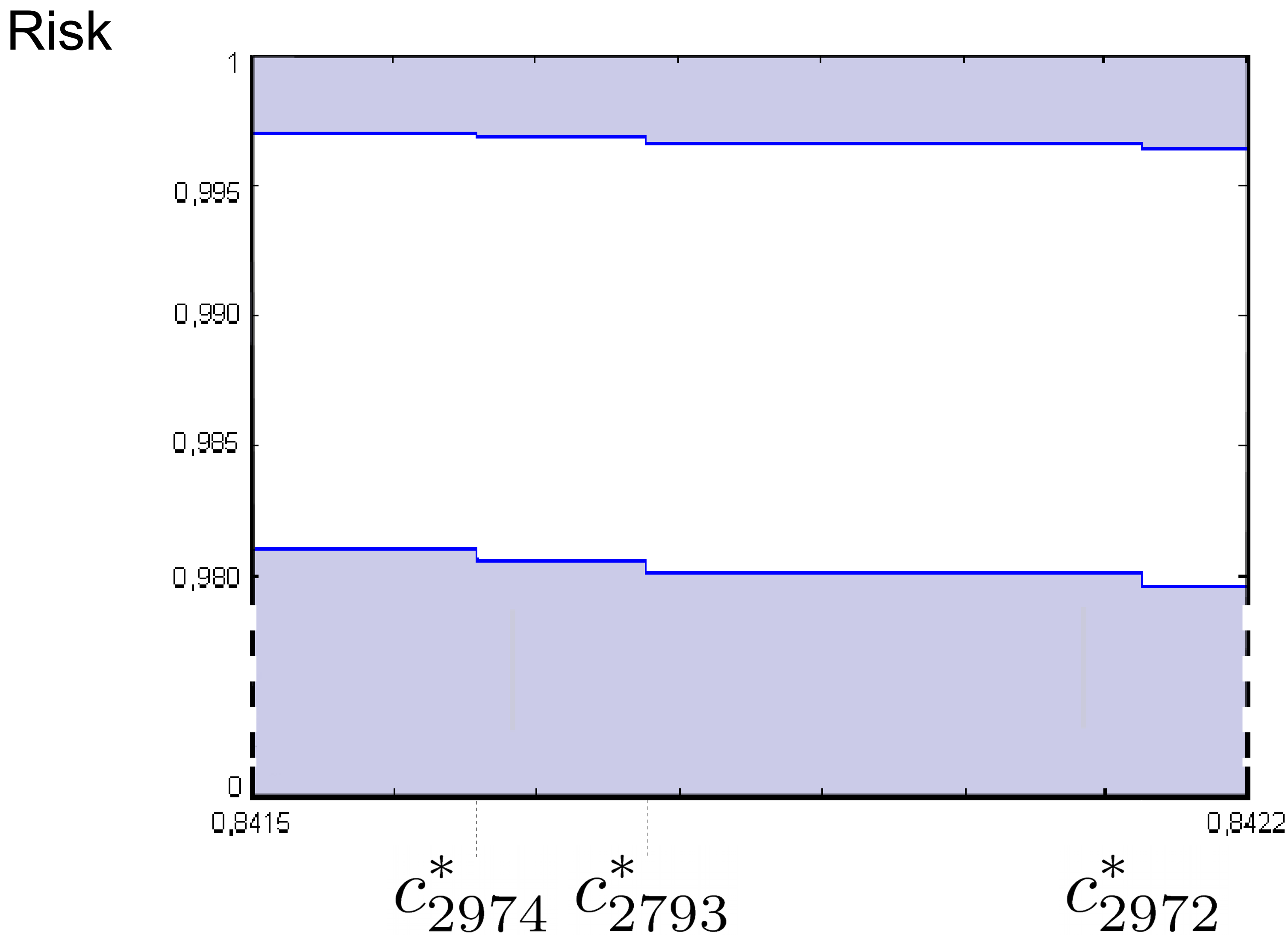


Risk
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.8
0.9
1
1.1
1.2
1.3
Cost
c^*_{2990}
c^*_{1700}
c^*_{80}
c^*_{1}

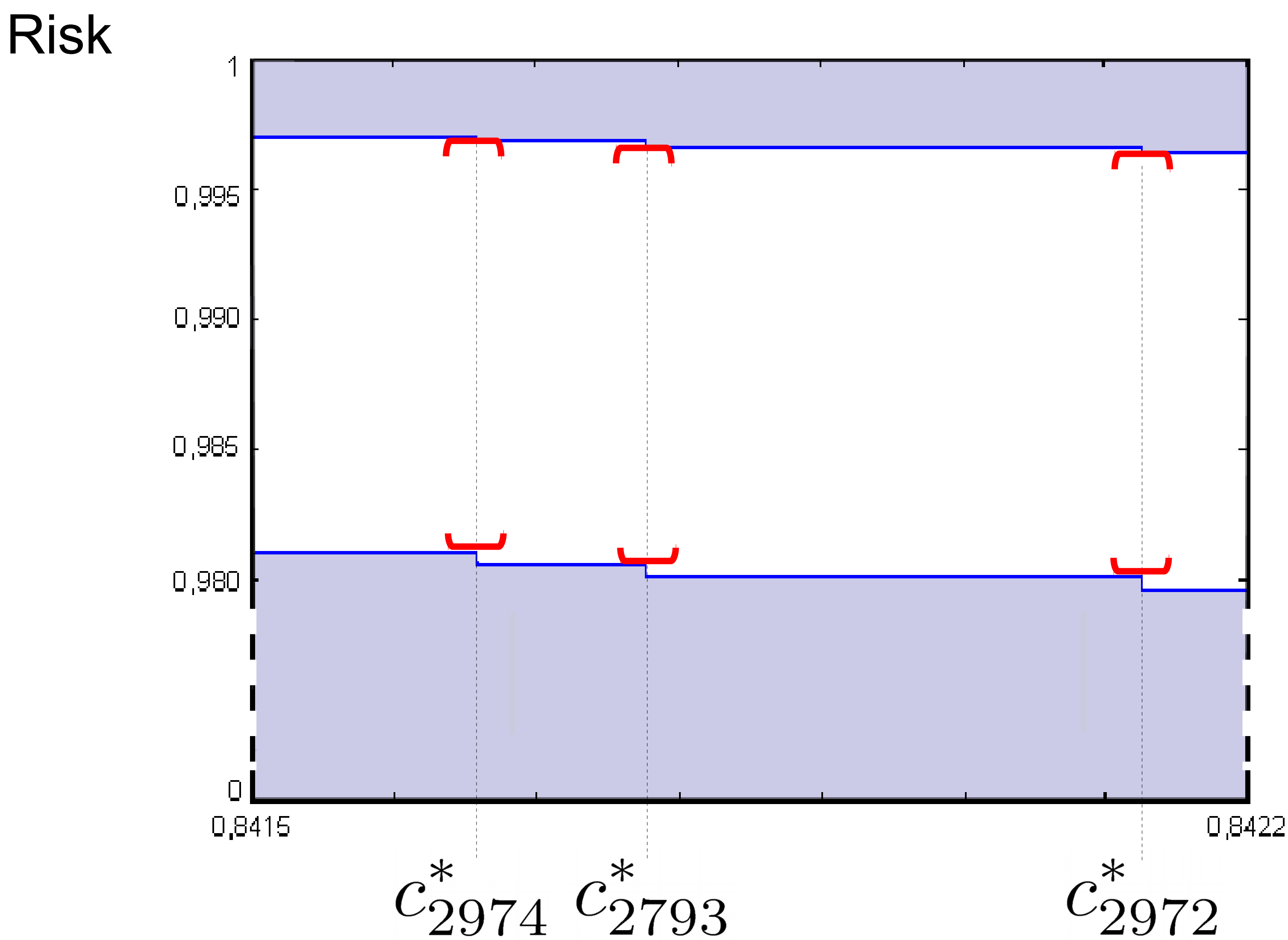


Risk
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.8
0.9
1
1.1
1.2
1.3
Cost
c^*_{2990}
c^*_{1700}
c^*_{80}
c^*_{1}

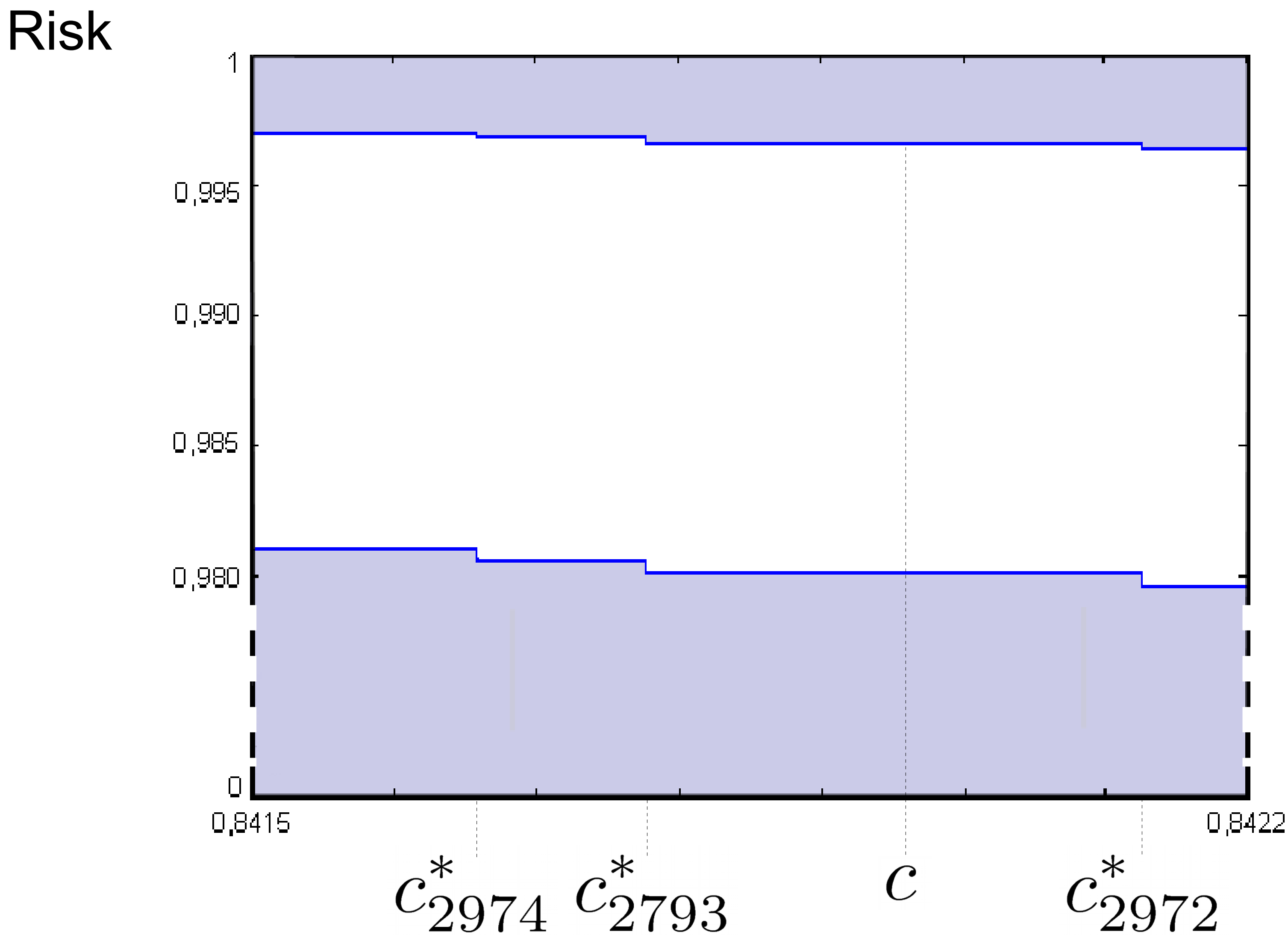
Zoomed detail



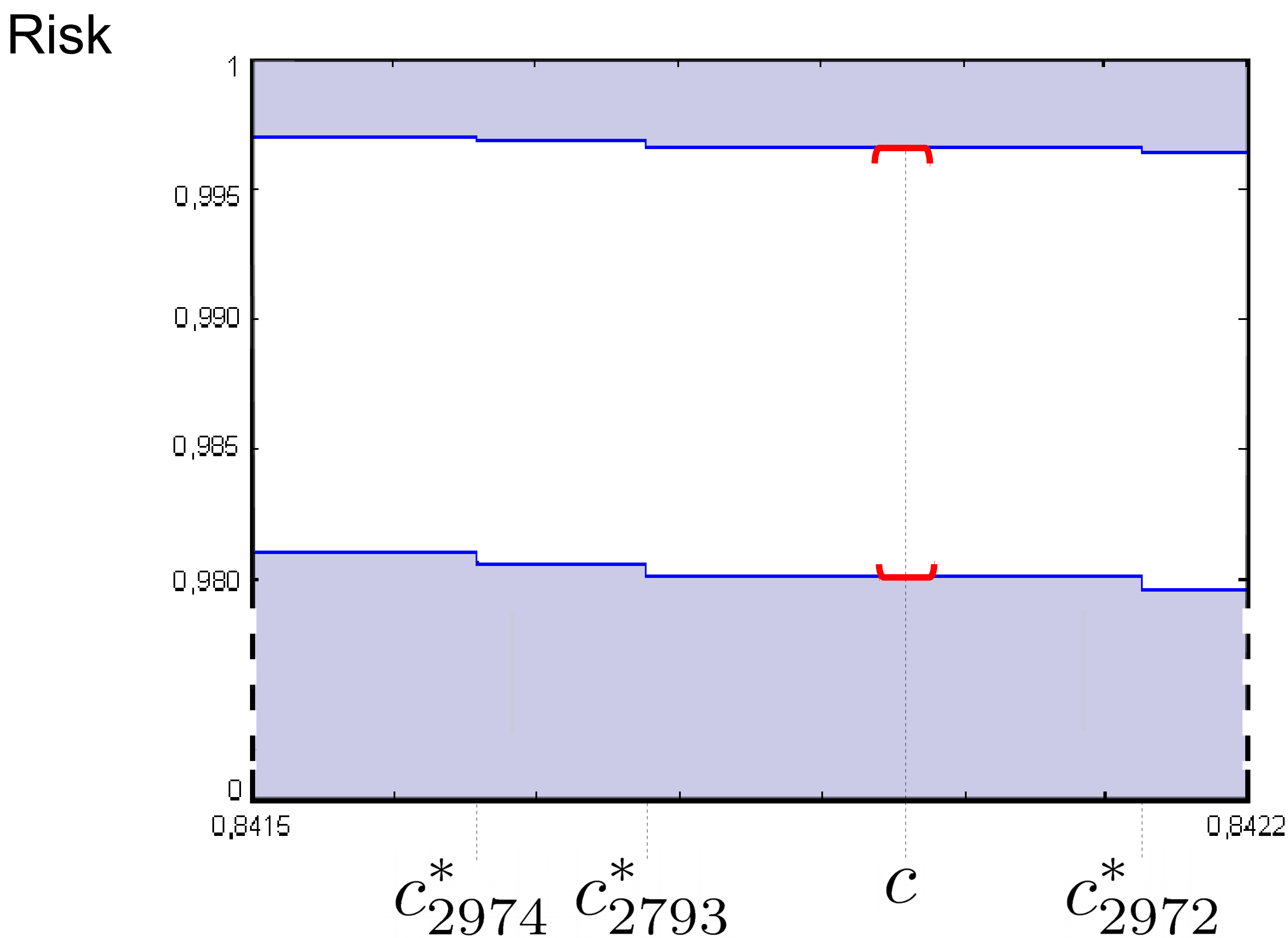
Zoomed detail

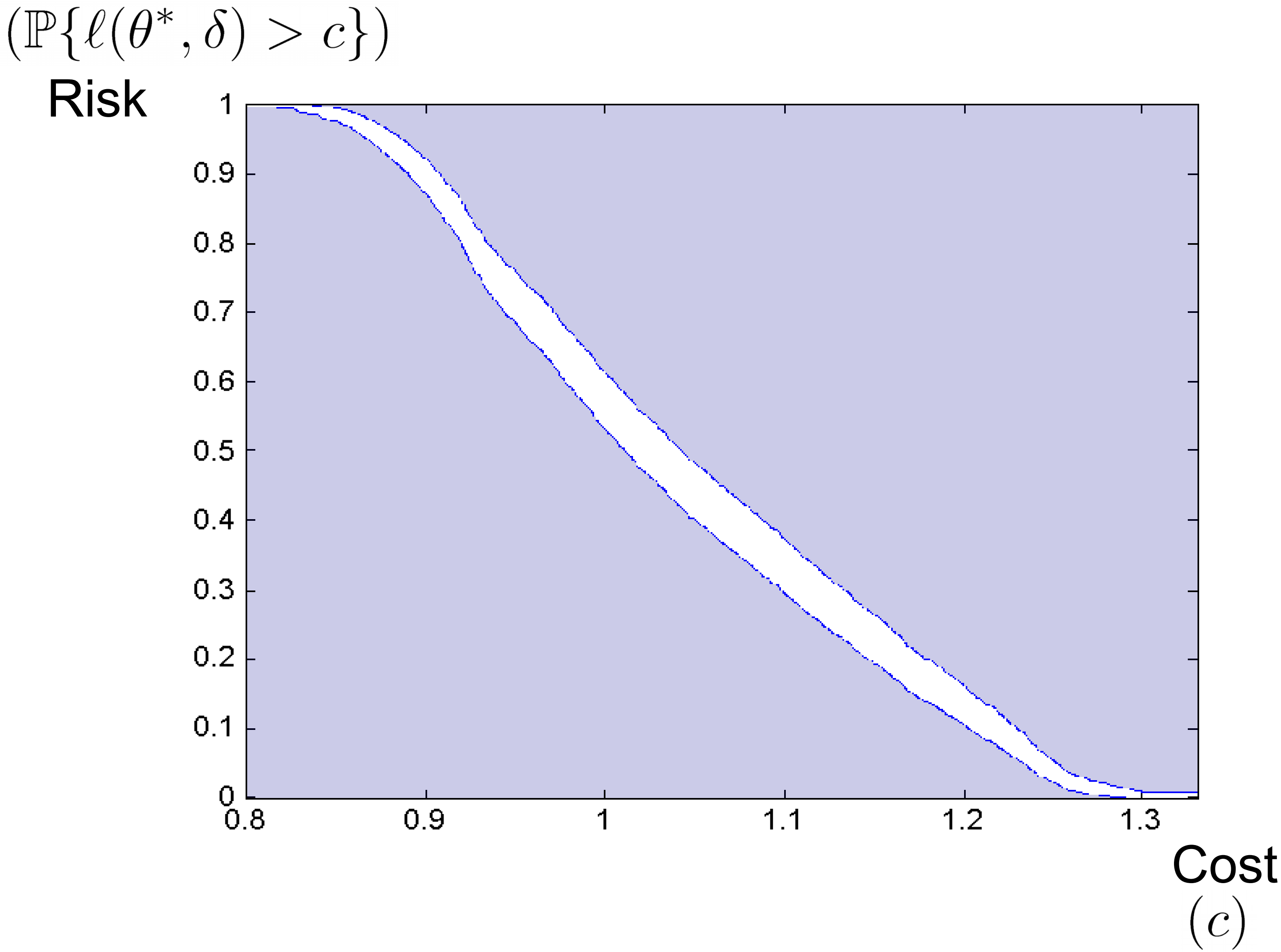


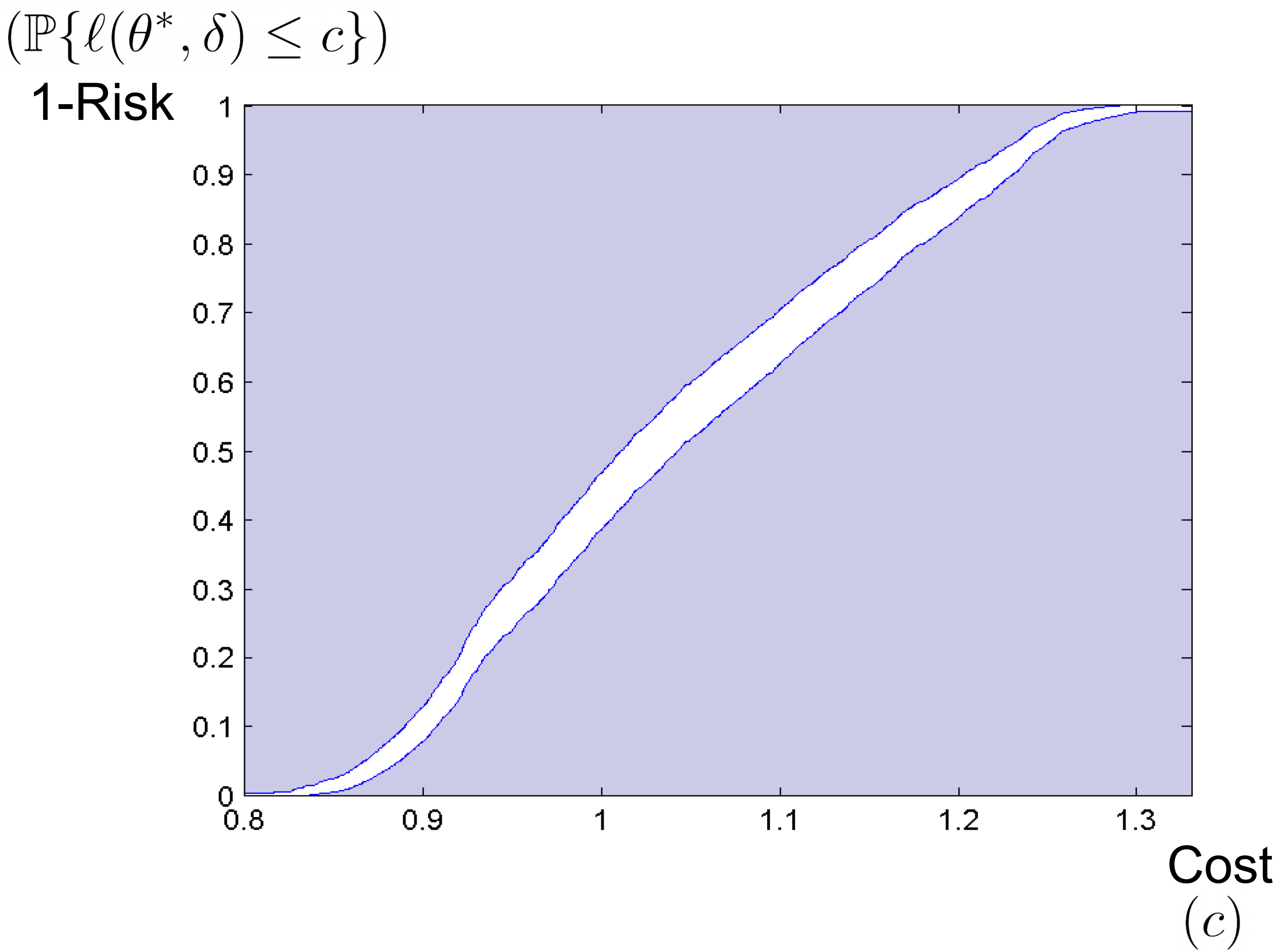
Zoomed detail



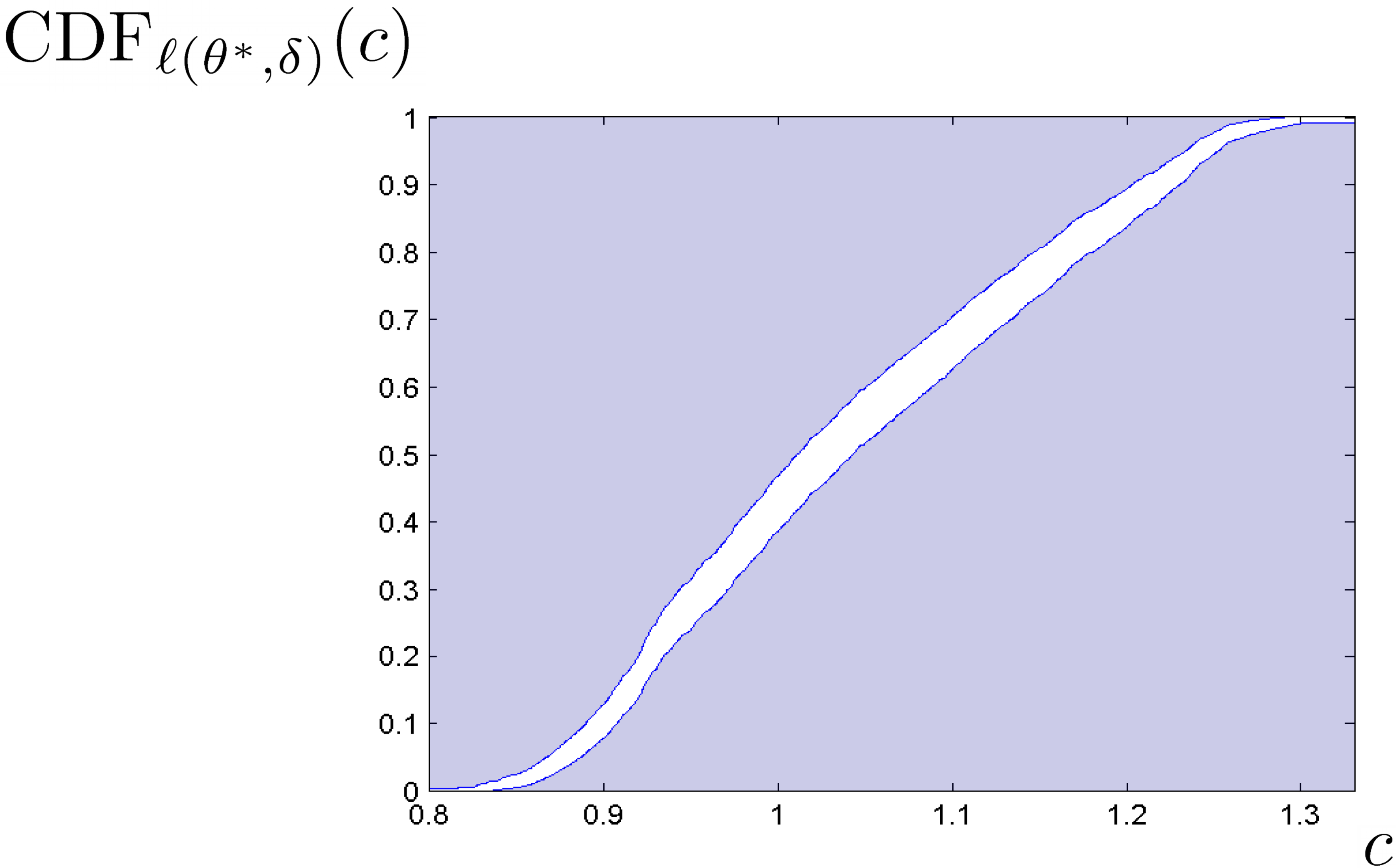
Zoomed detail



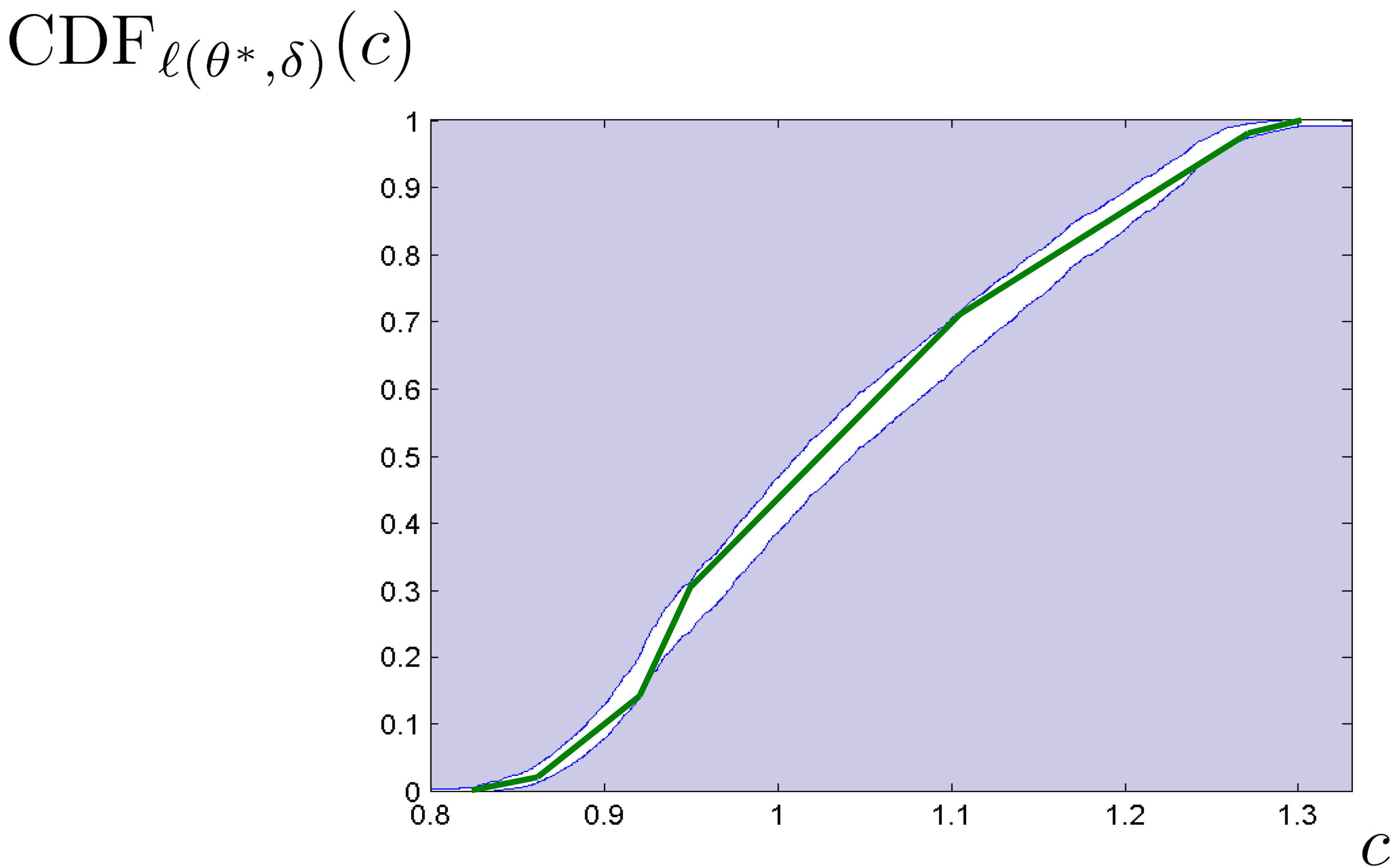

$(\mathbb{P}\{\ell(\theta^*, \delta) > c\})$
Risk
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.8
0.9
1
1.1
1.2
1.3
Cost
(c)



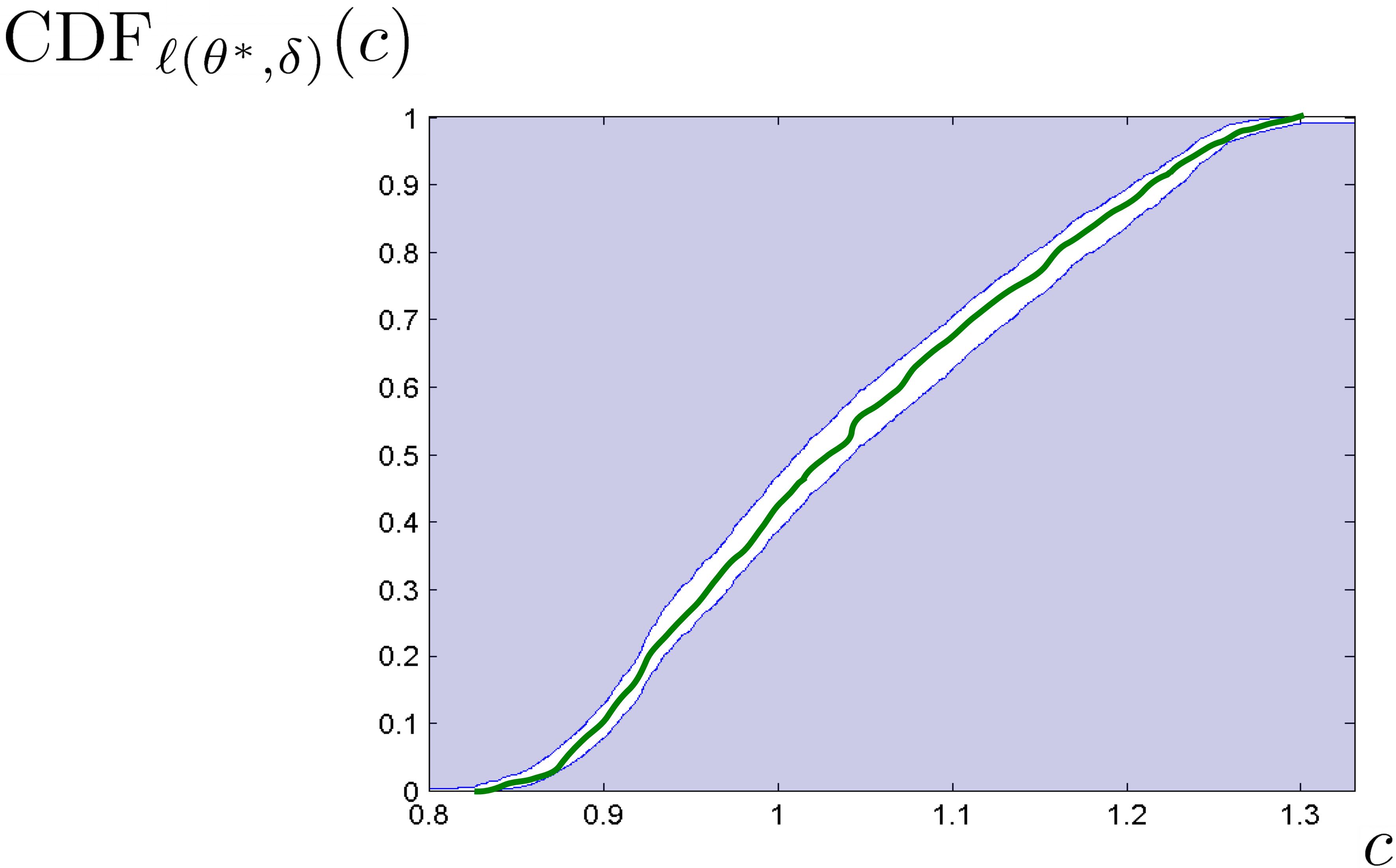
$(\mathbb{P}\{\ell(\theta^*, \delta) \leq c\})$
1-Risk
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.8
0.9
1
1.1
1.2
1.3
Cost
(c)

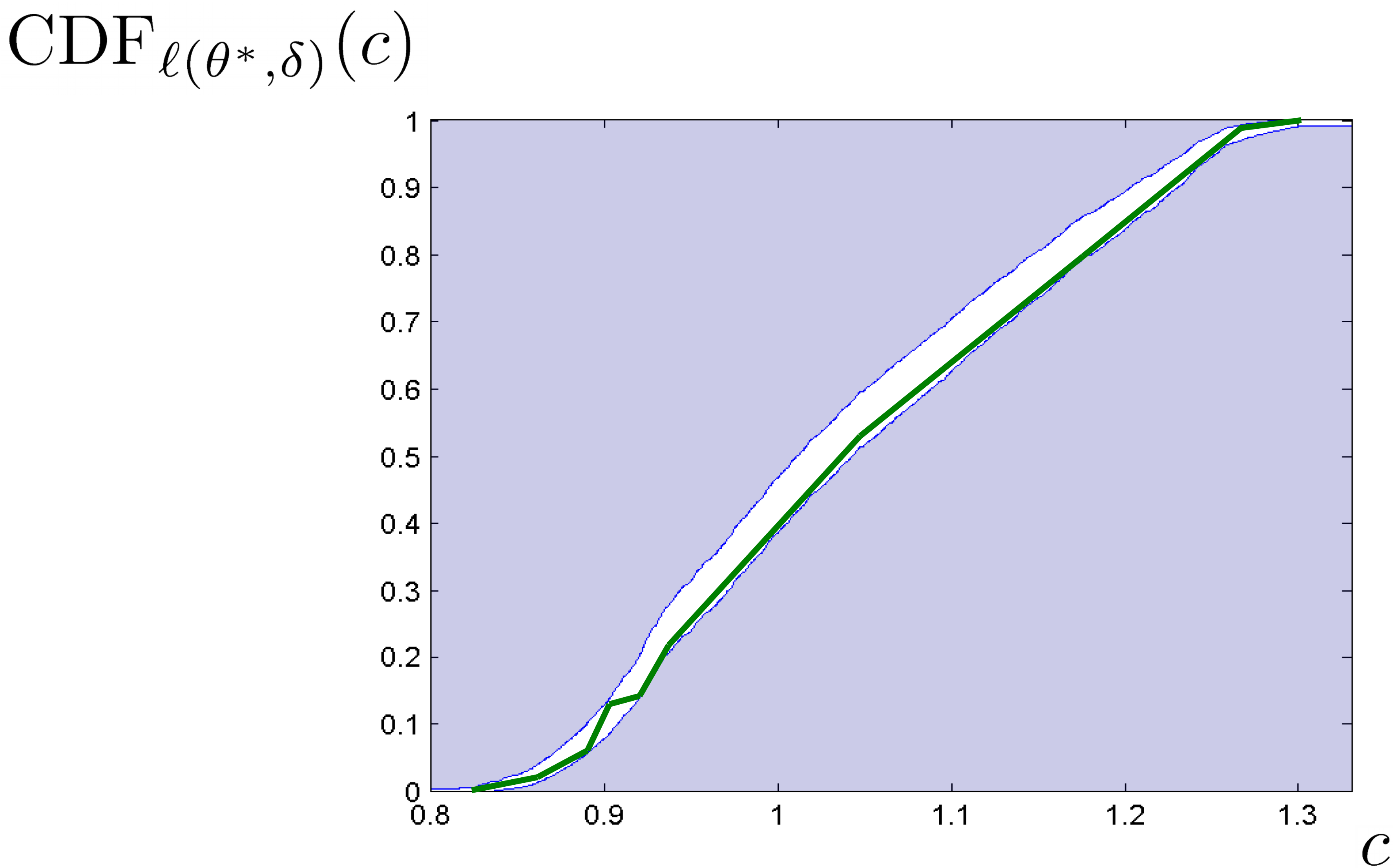


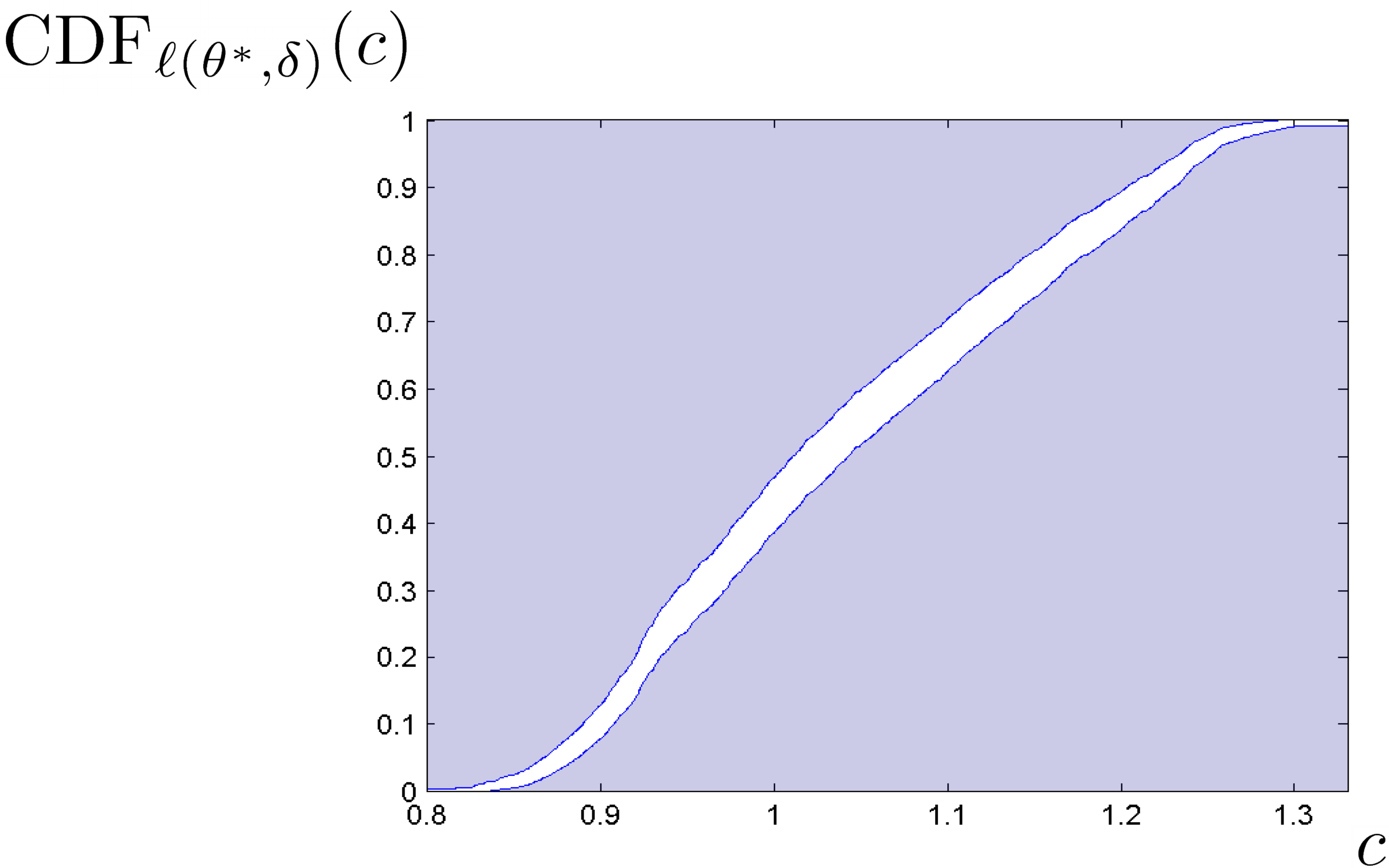
$\mathrm{CDF}_{\ell(\theta^*,\delta)}(c)$
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.8
0.9
1
1.1
1.2
1.3
c



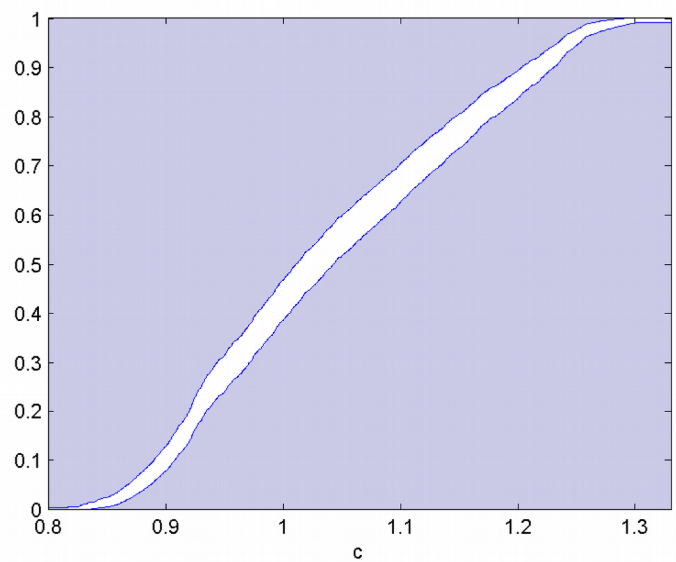
$\mathrm{CDF}_{\ell(\theta^*,\delta)}(c)$
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.8
0.9
1
1.1
1.2
1.3
c

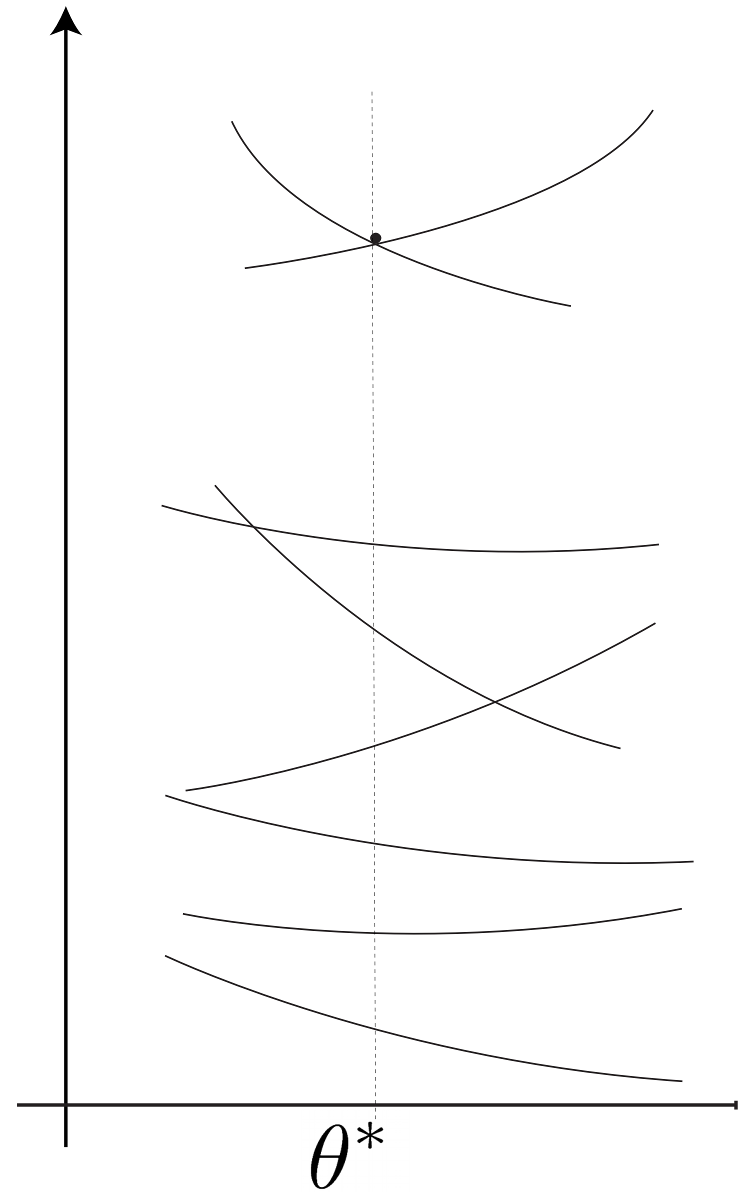

$\mathrm{CDF}_{\ell(\theta^*,\delta)}(c)$
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.8
0.9
1
1.1
1.2
1.3
c


$\mathrm{CDF}_{\ell(\theta^*,\delta)}(c)$
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.8
0.9
1
1.1
1.2
1.3
c


$\mathrm{CDF}_{\ell(\theta^*,\delta)}(c)$
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.8
0.9
1
1.1
1.2
1.3
c

Probability distribution of $\ell(\theta^*, \delta)$





We have reconstructed (wrapped) the real distribution of the cost from data (N scenarios).

No specific knowledge of $\ell(\theta, \delta), \Delta$ and $\mathbb{P}$ has been used.

We have reconstructed (wrapped) the real distribution of the cost from data (N scenarios).

No specific knowledge of $\ell(\theta, \delta), \Delta$ and $\mathbb{P}$ has been used.

Thank you!

REFERENCE

A. Carè, S. Garatti, and M.C. Campi,
Scenario Min-Max Optimization and the Risk of Empirical Costs.
SIAM Journal on Optimization, 25(4):2061-2080, 2015.

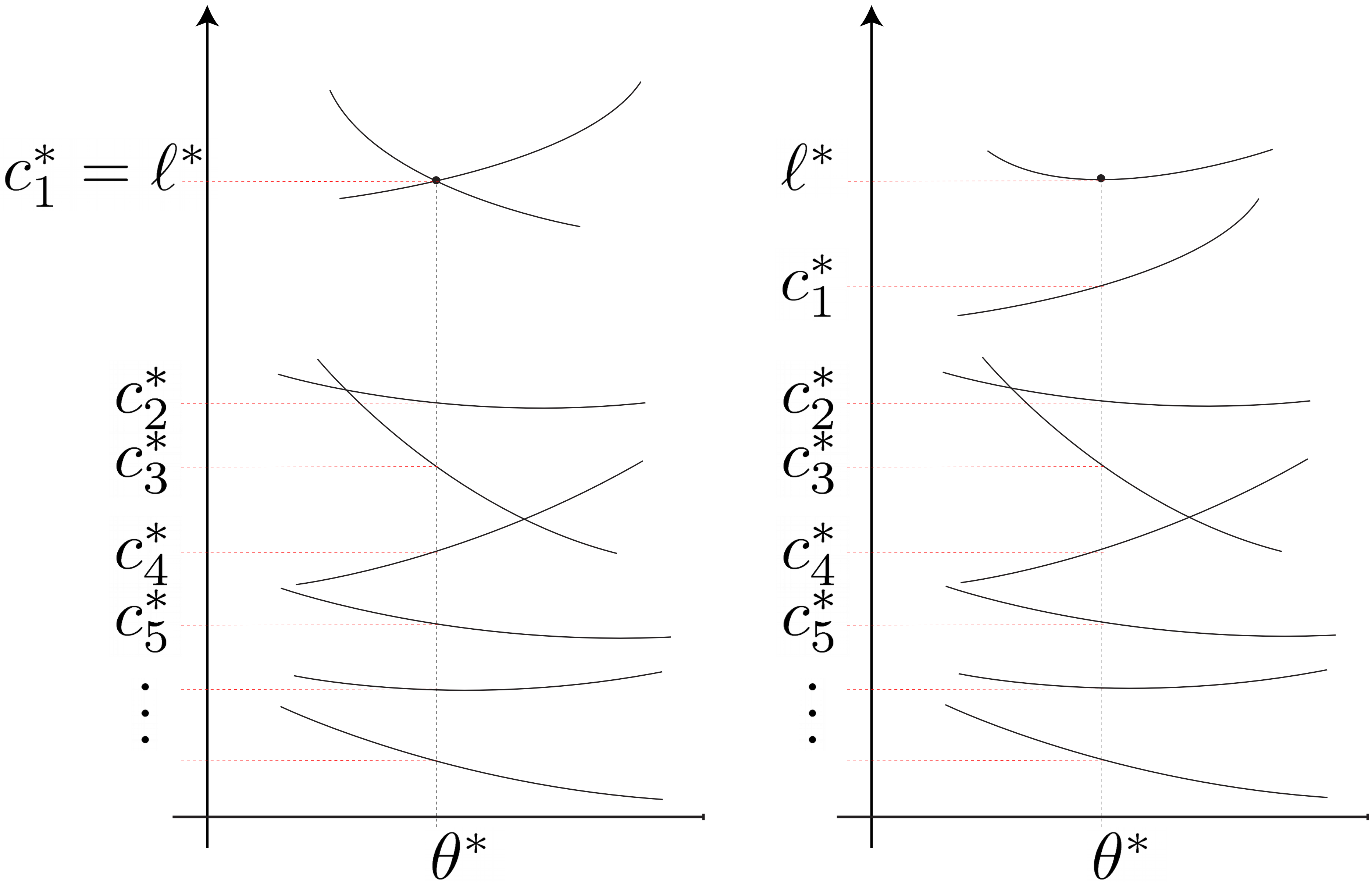
algo.care@gmail.com

Ordered Dirichlet C.D.F.

$$\mathbb{P}^N\{R_1 \leq \epsilon_1, R_2 \leq \epsilon_2, \ldots, R_{N-d} \leq \epsilon_{N-d}\}$$

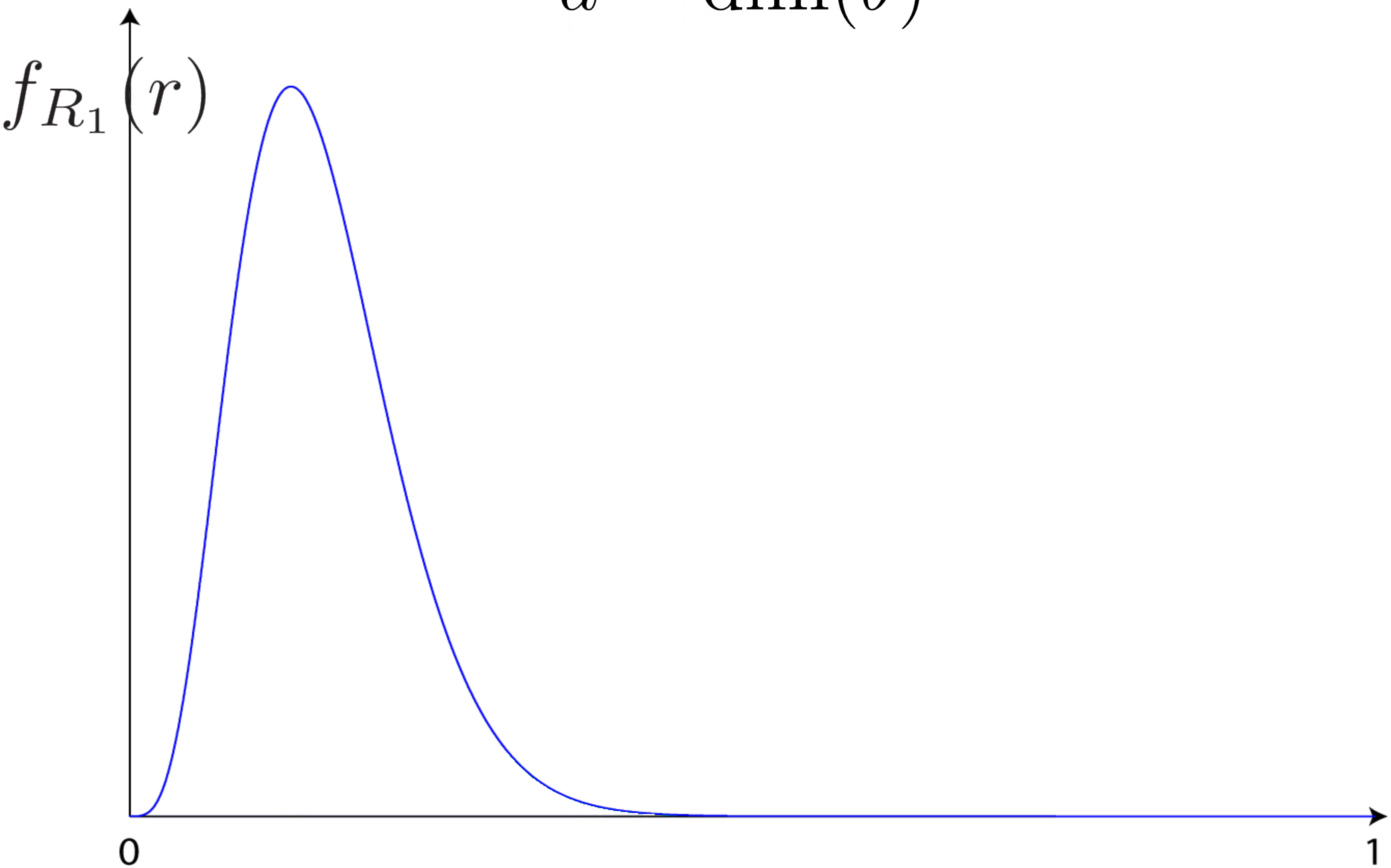
$$= \frac{N!}{d!}\int_0^{\epsilon_1}\int_{r_1}^{\epsilon_2}\cdots\int_{r_{N-d-1}}^{\epsilon_{N-d}} r_1^d \mathrm{d}r_{N-d}\cdots\mathrm{d}r_2\mathrm{d}r_1$$

“Fully” vs “non-fully” supported



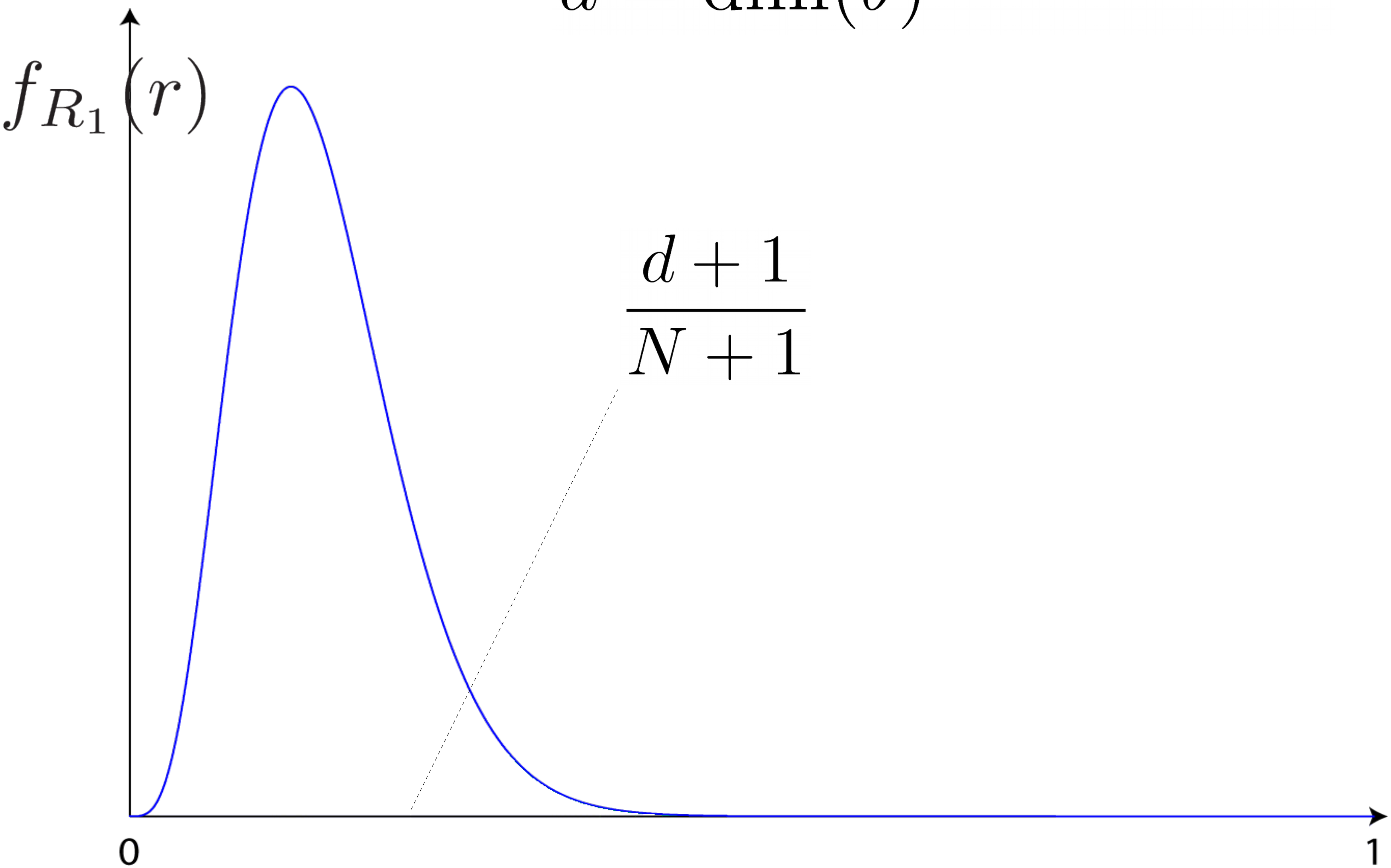
N number of scenarios

$d = \dim(\theta)$



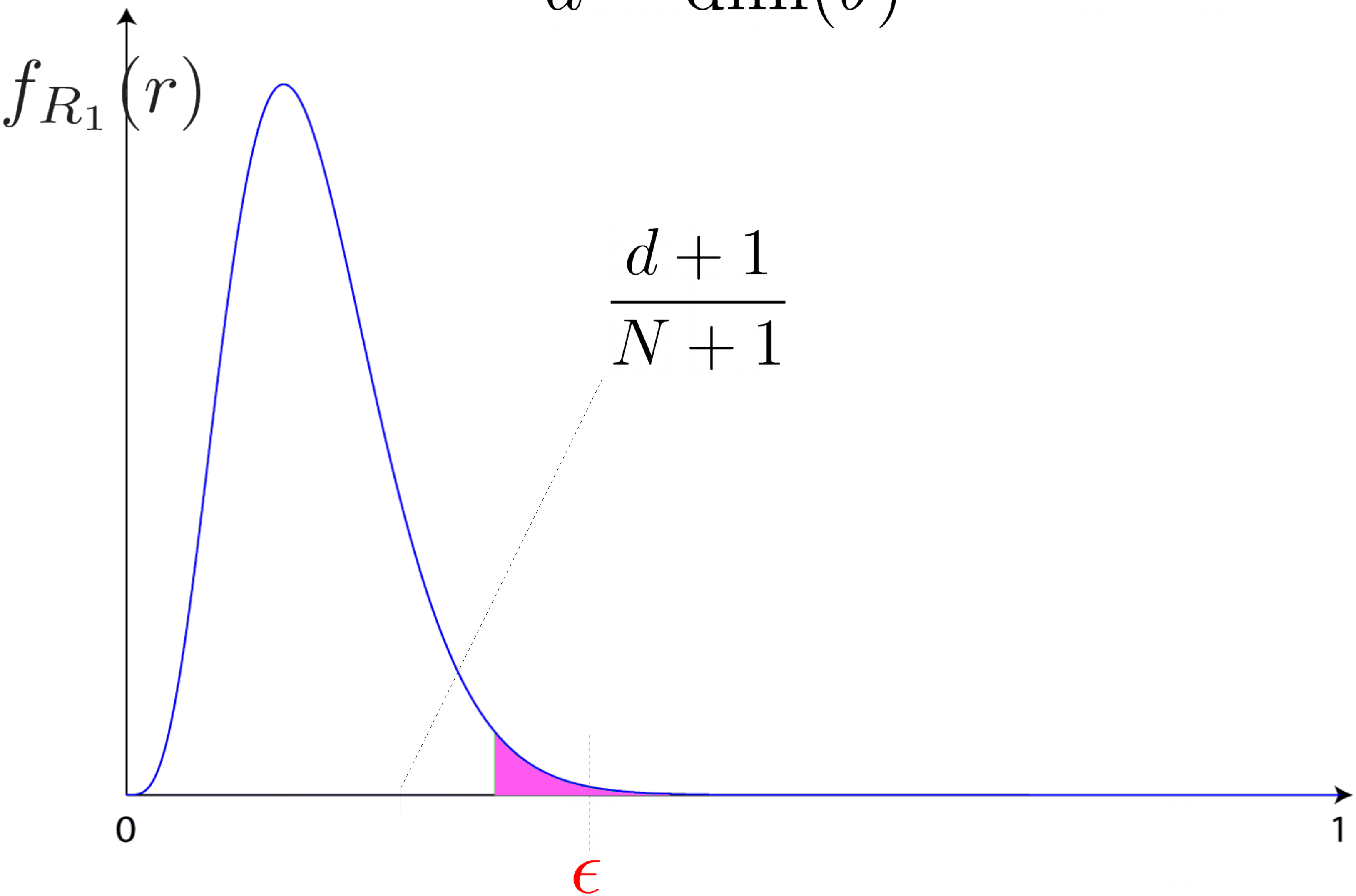
N number of scenarios

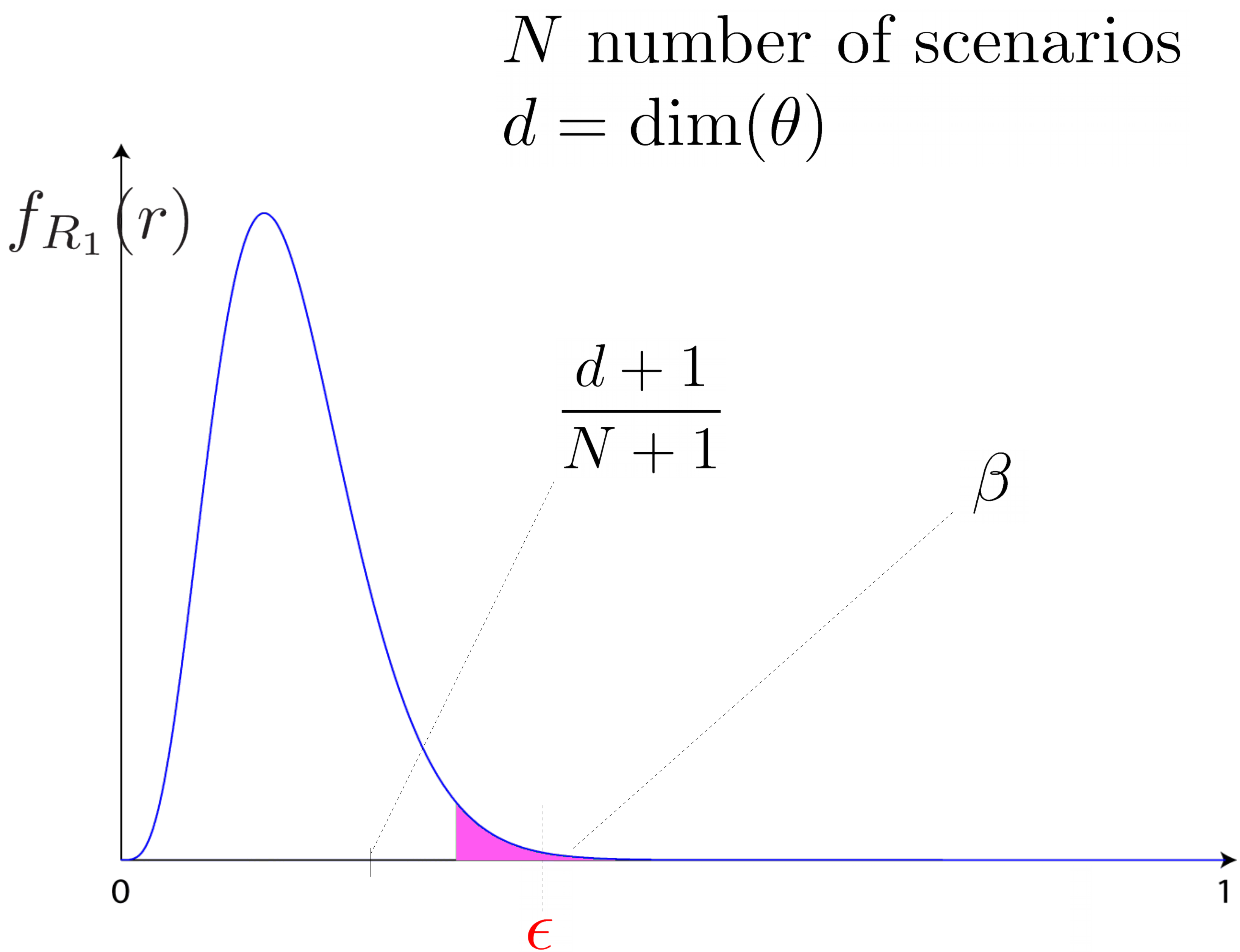
$d = \dim(\theta)$



N number of scenarios

$d = \dim(\theta)$

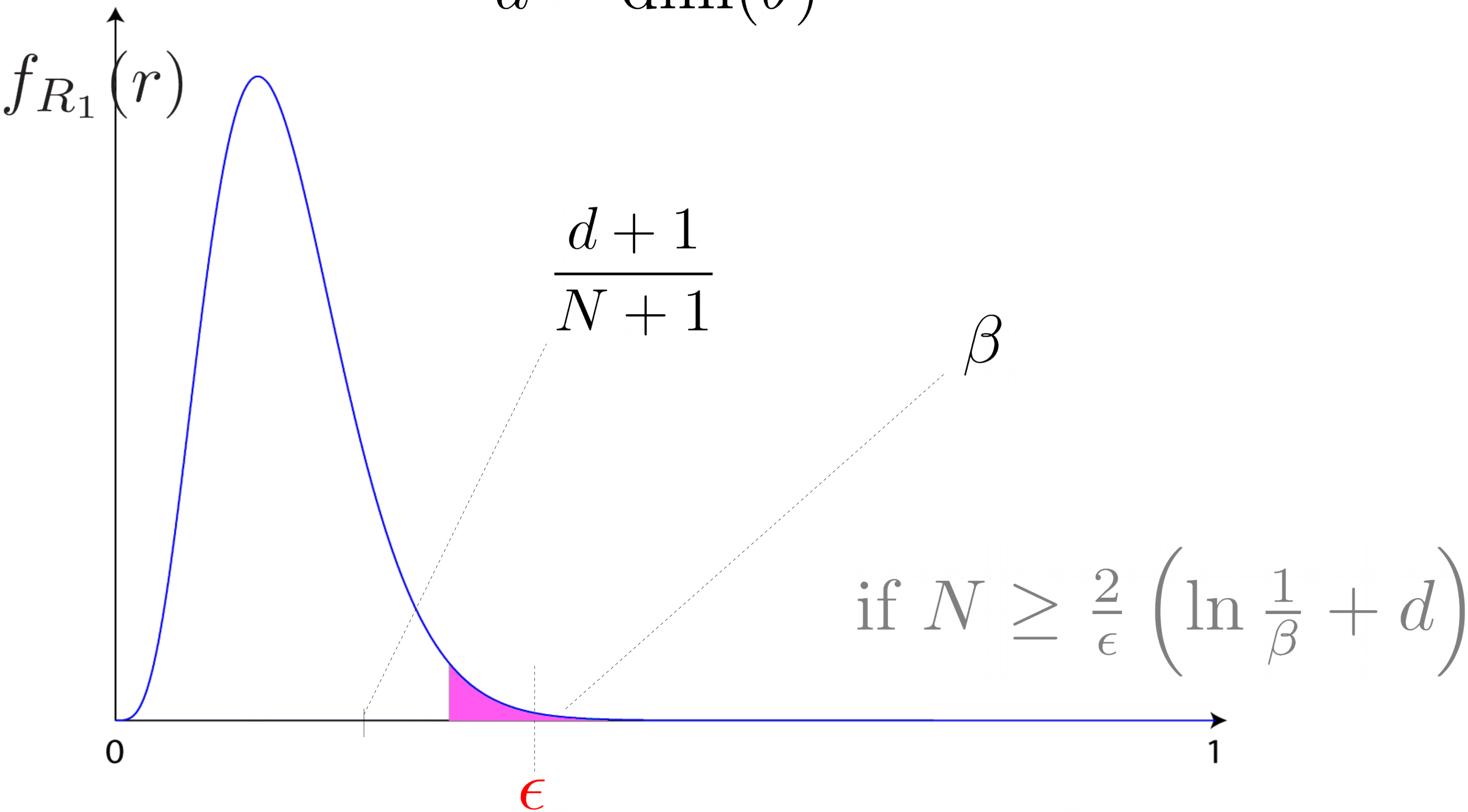


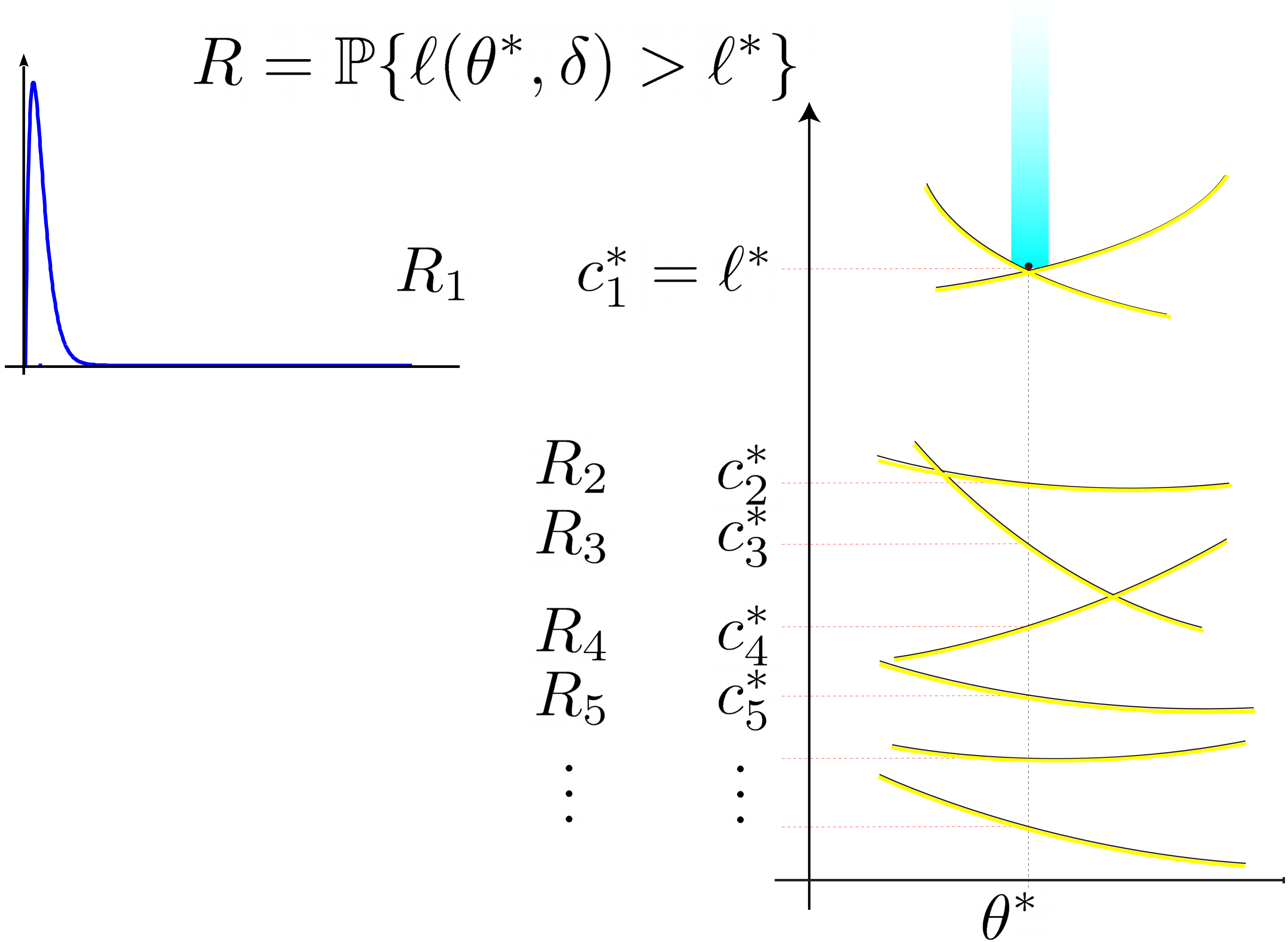


N number of scenarios
$d = \dim(\theta)$
$f_{R_1}(r)$
$\frac{d+1}{N+1}$
β
0
1
ϵ

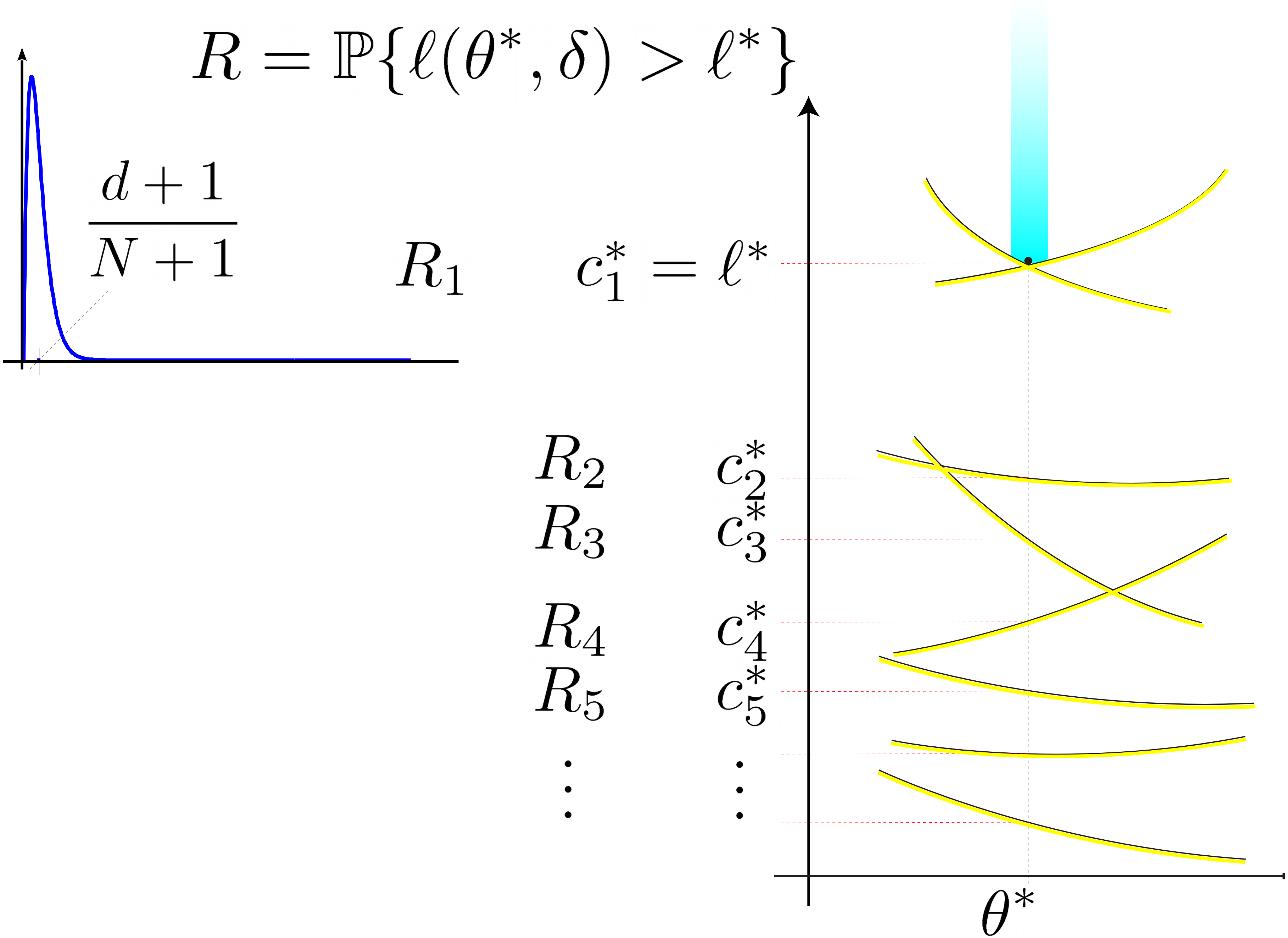
N number of scenarios

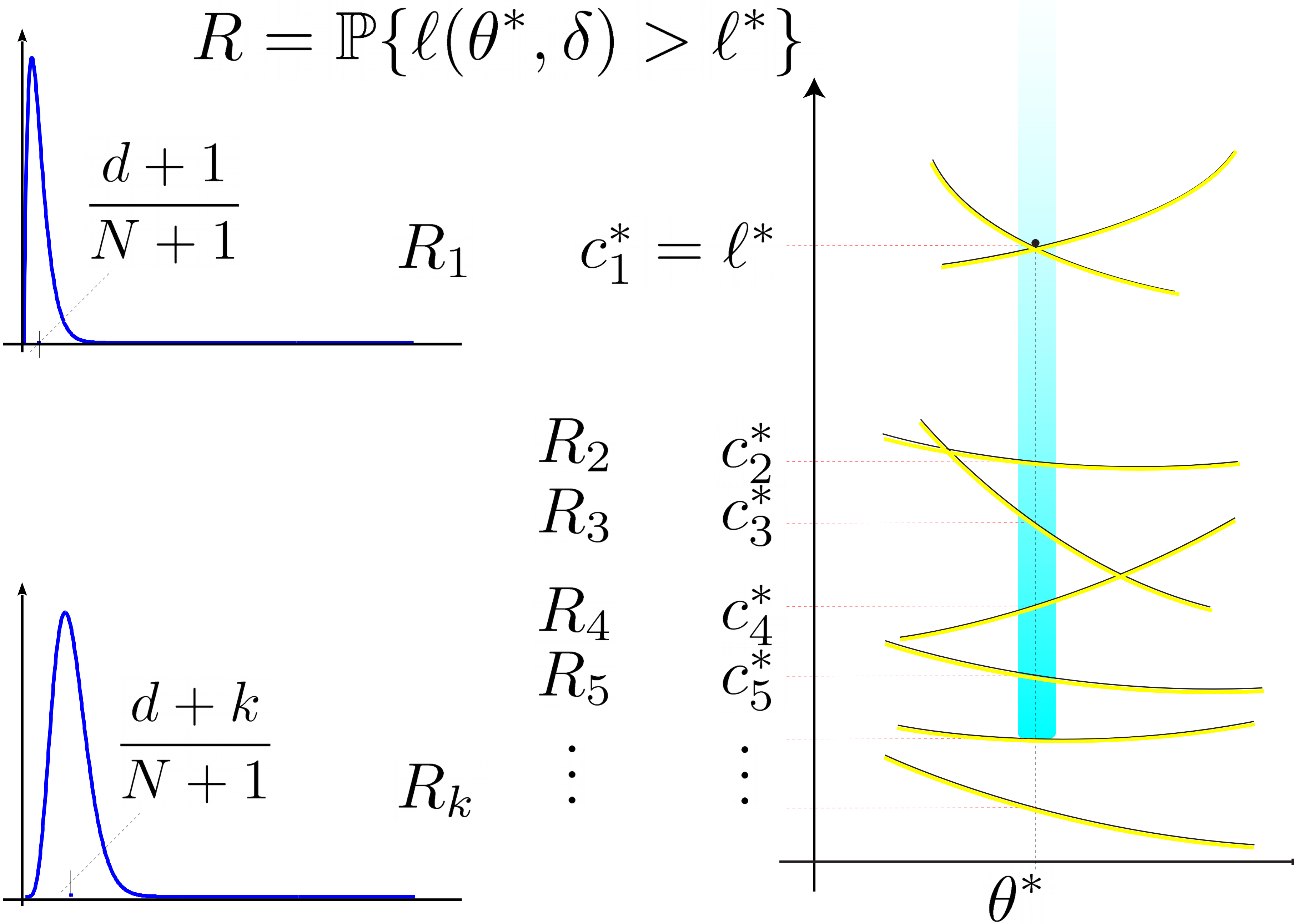
$d = \dim(\theta)$





$R = \mathbb{P}\{\ell(\theta^*, \delta) > \ell^*\}$
R_1
$c_1^* = \ell^*$
R_2
c_2^*
R_3
c_3^*
R_4
c_4^*
R_5
c_5^*
θ^*


$R = \mathbb{P}\{\ell(\theta^*, \delta) > \ell^*\}$
$\frac{d+1}{N+1}$
R_1
$c_1^* = \ell^*$
R_2
c_2^*
R_3
c_3^*
R_4
c_4^*
R_5
c_5^*
$\vdots$
$\vdots$
θ^*



$R = \mathbb{P}\{\ell(\theta^*, \delta) > \ell^*\}$
$\frac{d+1}{N+1}$
R_1
$c_1^* = \ell^*$
R_2
R_3
R_4
R_5
c_2^*
c_3^*
c_4^*
c_5^*
$\frac{d+k}{N+1}$
R_k
θ^*