Chapter 3:

Data-driven Uncertainty Set Design

Chance constraints

- Charnes and Cooper introduced in 1959, a concept now referred as chance constraint. Namely, given a distribution F for a random vector Z and a tolerance $\epsilon > 0$.
  One can impose that

$$\mathbb{P}(a(Z)^T x \leq b(Z)) \geq 1 - \epsilon$$

- This also gives rise to the notion of Value at Risk for a return

$$\mathrm{VaR}_{1-\epsilon}(c(Z)^T x + d(Z)) := -\sup\{y \in \mathbb{R} \,|\, \mathbb{P}(c(Z)^T x + d(Z) \geq y) \geq 1 - \epsilon\}$$

- Both involve verifying whether a constraint is satisfied with high probability

- E.g. minimizing the value at risk of a portfolio of stocks

$$\min_{x : x \geq 0, \sum_i x_i = 1} \quad \mathrm{VaR}_{1-\epsilon}(r^T x)$$

SOCP reformulation for normal distribution

- The three following constraints are equivalent when the random return vector « r » is normally distributed $N(\mu, \Sigma)$

(1) $\mathbb{P}(r^T x \geq y) \geq 1 - \epsilon$

(2) $\mu^T x - \Phi^{-1}(1-\epsilon)\sqrt{x^T \Sigma x} \geq y$

(3) $r^T x \geq y \,,\, \forall\, r : \|\Sigma^{-1/2}(r-\mu)\| \leq \Phi^{-1}(1-\epsilon)$

- For general distribution, verifying whether the chance constraint is satisfied for a fixed « x » is NP-hard.

Robust optimization as an approximation to chance constraints

Theorem 3.2. : *Given some* $\epsilon > 0$ *and some random vector* Z *distributed according to* F*, let* $\mathcal{Z}$ *be a set such that*

$$\mathbb{P}(Z \in \mathcal{Z}) \geq 1 - \epsilon \,,$$

then one has the guarantee that any x *satisfying the robust constraint*

$$a(z)^T x \leq b(z) \,, \forall\, z \in \mathcal{Z} \,,$$

will also satisfy the following chance constraint

$$\mathbb{P}(a(Z)^T x \leq b(Z)) \geq 1 - \epsilon \,.$$

Note that the converse is not true so that the two constraints are generally not equivalent.

Robust optimization as an approximation to chance constraints

Theorem 3.2. : *Given some $\epsilon > 0$ and some random vector Z distributed according to F, let $\mathcal{Z}$ be a set such that*

$$\mathbb{P}(Z \in \mathcal{Z}) \geq 1 - \epsilon \,,$$

then one has the guarantee that any x satisfying the robust constraint

In particular, in the example:

$$\mathbb{P}(r^T x \geq y) \geq 1 - \epsilon$$

if r is normally distributed and an ellipsoidal set is used one would get the following uncertainty set

$$r^T x \geq y \,,\, \forall\, r : \|\Sigma^{-1/2}(r - \mu)\| \leq \sqrt{F^{-1}_{\chi^2_m}(1 - \epsilon)} \neq \Phi^{-1}(1 - \epsilon)$$

How RO approximates chance constraints

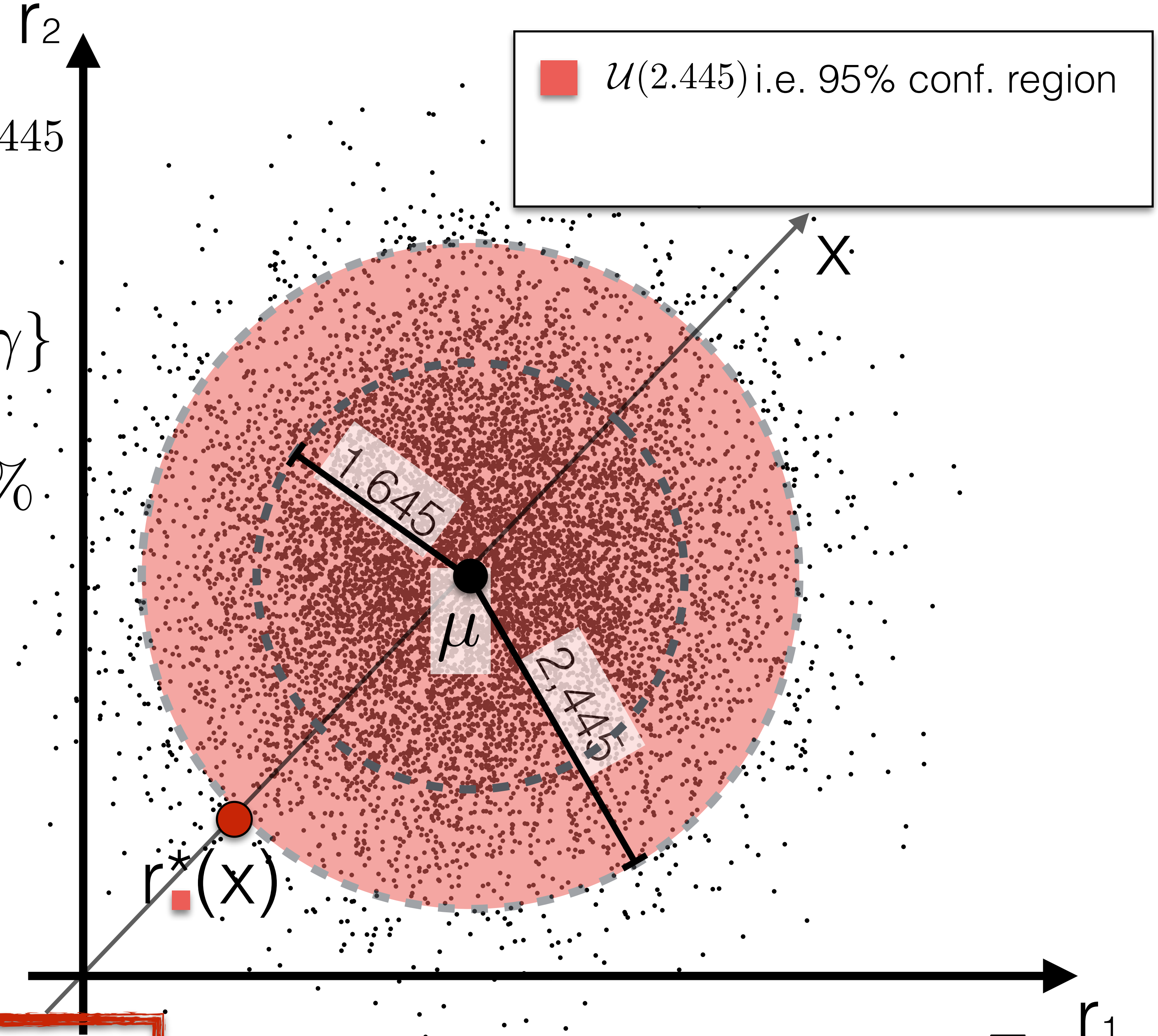
Let $m = 2$, $\epsilon = 5\%$, and $\Sigma = I$,

$\Phi^{-1}(1-\epsilon) = 1.645, \sqrt{F^{-1}_{\chi^2_2}(1-\epsilon)} = 2.445$

Let
$\mathcal{U}(\gamma) := \{r \mid \|\Sigma^{-1/2}(r-\mu)\|_2 \leq \gamma\}$

Since $P(r \in \mathcal{U}(2.445)) = 95\%$

$$r^T x \geq y \, , \forall \, r \in \mathcal{U}(2.445)$$
$$\Downarrow$$
$$P(r^T x \geq y) \geq 95\%$$



Actually, $P(r^T x \geq y) \geq 99\%$

$$r^*(x) = \arg \min_{r \in \mathcal{U}(2.445)} r^T x$$

How RO approximates chance constraints

Let $m = 2$, $\epsilon = 5\%$, and $\Sigma = I$,

$\Phi^{-1}(1-\epsilon) = 1.645, \sqrt{F^{-1}_{\chi^2_2}(1-\epsilon)} = 2.445$

Let

$\mathcal{U}(\gamma) := \{r \,|\, \|\Sigma^{-1/2}(r-\mu)\|_2 \leq \gamma\}$

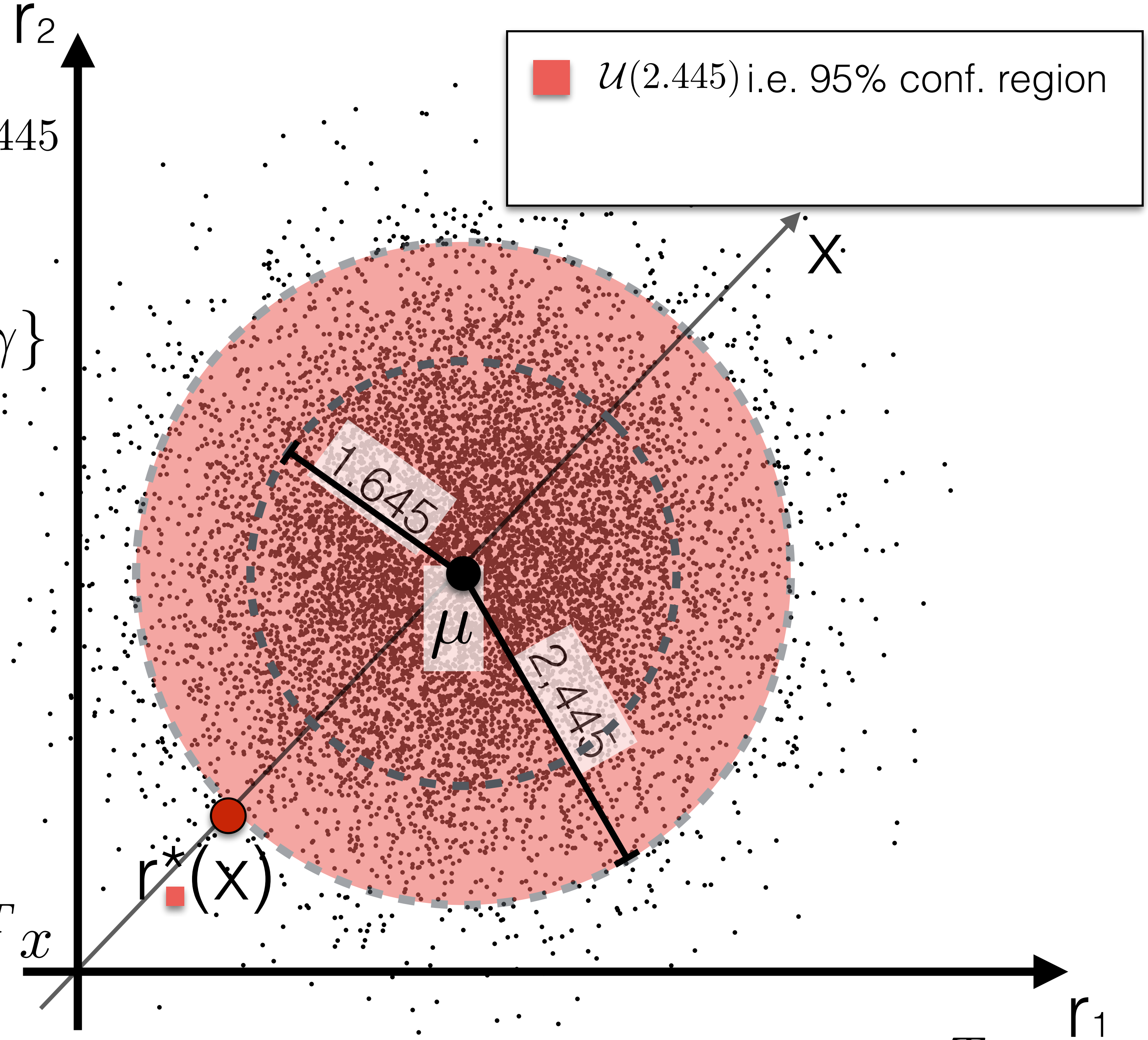
$$r^T x \geq y \,, \forall\, r \in \mathcal{U}(2.445)$$

$$\Downarrow$$

$$\min_{r \in \mathcal{U}(2.445)} r^T x \geq y$$

$$\Downarrow$$

$$r^T x \geq y \,, \forall\, r : r^T x \geq \min_{\bar{r} \in \mathcal{U}(2.445)} \bar{r}^T x$$



$r^*(x) = \arg \min_{z \in \mathcal{U}(2.445)} r^T x$

How RO approximates chance constraints

Let $m = 2$, $\epsilon = 5\%$, and $\Sigma = I$,

$\Phi^{-1}(1-\epsilon) = 1.645, \sqrt{F^{-1}_{\chi^2_2}(1-\epsilon)} = 2.445$

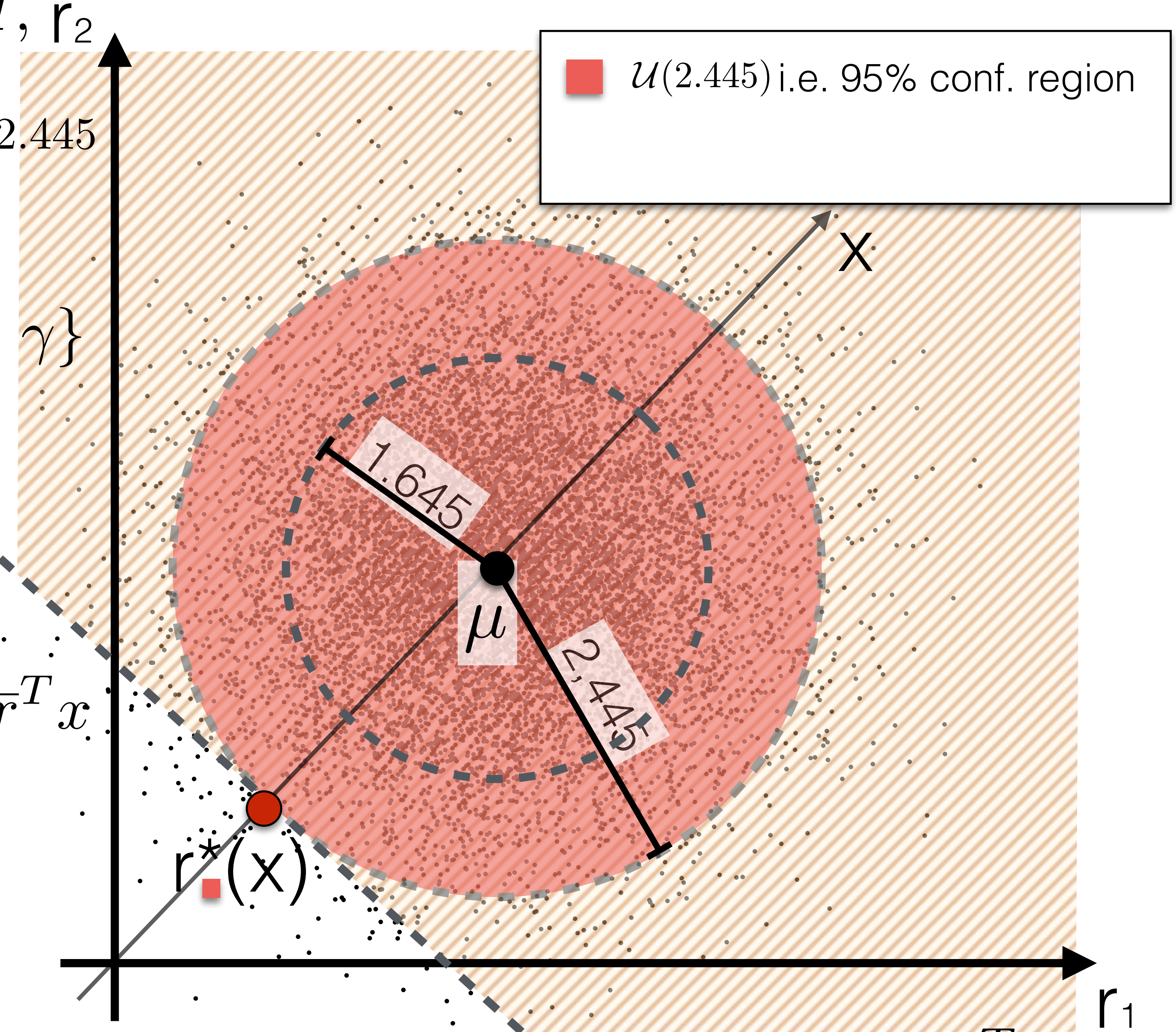
Let

$\mathcal{U}(\gamma) := \{r \mid \|\Sigma^{-1/2}(r-\mu)\|_2 \leq \gamma\}$

$$r^T x \geq y \,,\, \forall\, r \in \mathcal{U}(2.445)$$

$$\Downarrow$$

$$r^T x \geq y \,,\, \forall\, r : r^T x \geq \min_{\bar{r} \in \mathcal{U}(2.445)} \bar{r}^T x$$



$r^*(x) = \arg \min_{z \in \mathcal{U}(2.445)} r^T x$

How RO approximates chance constraints

Let $m = 2$, $\epsilon = 5\%$, and $\Sigma = I$,

$\Phi^{-1}(1-\epsilon) = 1.645, \sqrt{F^{-1}_{\chi^2_2}(1-\epsilon)} = 2.445$

Let

$\mathcal{U}(\gamma) := \{r \mid \|\Sigma^{-1/2}(r-\mu)\|_2 \leq \gamma\}$

$$r^T x \geq y\,,\ \forall\, r \in \mathcal{U}(2.445)$$

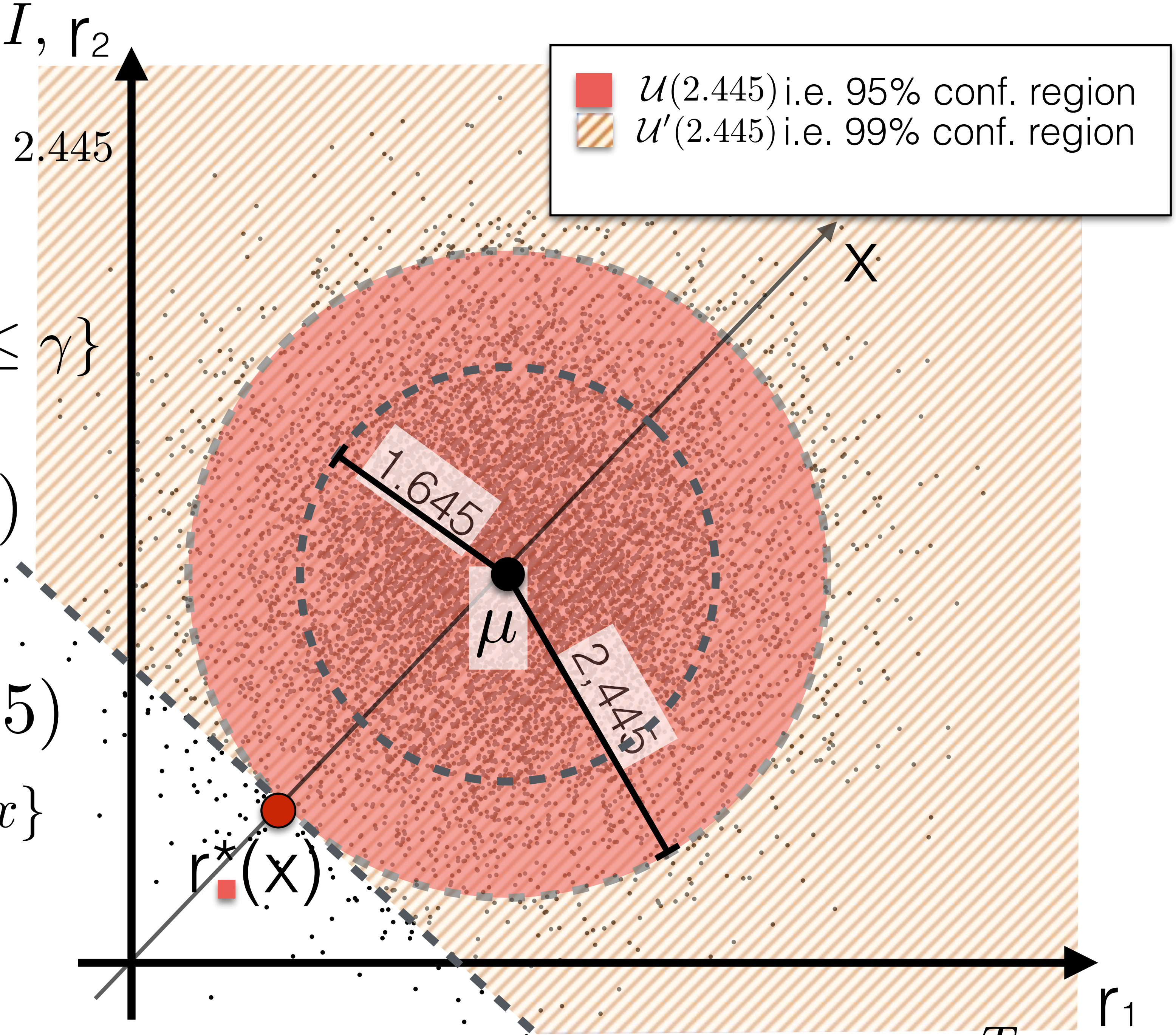
$$\Downarrow$$

$$r^T x \geq y\,,\ \forall\, r \in \mathcal{U}'(2.445)$$

$$\mathcal{U}'(\gamma) := \{r \mid r^T x \geq \min_{r \in \mathcal{U}(\gamma)} r^T x\}$$

$$\Downarrow$$

$$\boxed{P(r^T x \geq y) \geq 99\%}$$



$$r^*(x) = \arg \min_{z \in \mathcal{U}(2.445)} r^T x$$

How RO approximates chance constraints

Let $m = 2$, $\epsilon = 5\%$, and $\Sigma = I$,

$\Phi^{-1}(1-\epsilon) = 1.645, \sqrt{F^{-1}_{\chi^2_2}(1-\epsilon)} = 2.445$

Let

$\mathcal{U}(\gamma) := \{r \mid \|\Sigma^{-1/2}(r-\mu)\|_2 \leq \gamma\}$

$$r^T x \geq y \,,\, \forall\, r \in \mathcal{U}(\, 1.645\,)$$

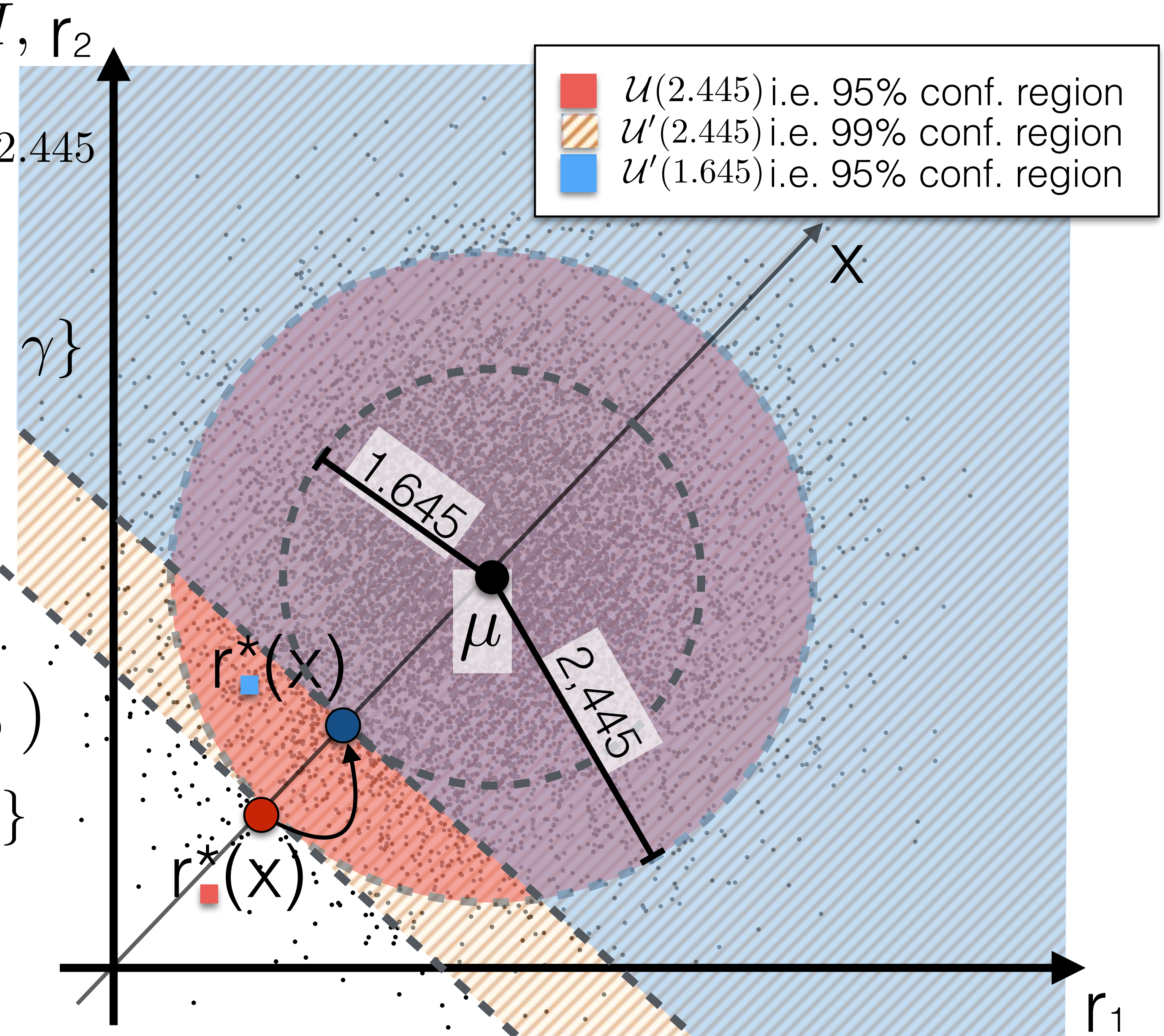
$$\Downarrow$$

$$r^T x \geq y \,,\, \forall\, r \in \mathcal{U}'(\, 1.645\,)$$

$$\mathcal{U}'(\gamma) := \{r \mid r^T x \geq \min_{r \in \mathcal{U}(\gamma)} r^T x\}$$

$$\Downarrow$$

$$P(r^T x \geq y) \geq 95\%$$



$\text{r}^*(x) = \arg \min_{r \in \mathcal{U}(1.645)} r^T x$

Implication for LP-RC

Corollary 3.3. : *Given some $\epsilon > 0$ and some random vector Z distributed according to F, let $\mathcal{Z}$ be a set such that*

$$\mathbb{P}(Z \in \mathcal{Z}) \geq 1 - \epsilon \,,$$

then the LP-RC optimization problem (2.1) is a conservative approximation of the stochastic program

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & VaR_{1-\epsilon}(Z^T P_0^T x + q_0^T Z + p_0^T x + r_0) \\ \text{subject to} \quad & \mathbb{P}(Z^T P_j^T x + p_j^T x \leq q_j^T Z + r_j) \geq 1 - \epsilon \,, \forall\, j = 1, ..., J \,, \end{aligned}$$

where $VaR_{1-\epsilon}(\cdot)$ is as defined in definition 3.1. Specifically, by conservative approximation we mean that an optimal solution to the LP-RC problem will be feasible according to the above stochastic program where it will achieve an objective value that is lower than what was established by the LP-RC optimization model.

Example: Portfolio with minimum VaR

You are given a set of historical monthly returns of 10 stocks for year 2000 - 2009, and are asked to approximate the following "value-at-risk" problem:

$$\begin{array}{ll} \underset{x,y}{\text{minimize}} & -y \\ \text{subject to} & \mathbb{P}(r^T x \geq y) \geq 1 - \epsilon \qquad \sum_{i=1}^{n} x_i = 1 \qquad x \geq 0\,, \end{array}$$

where $\epsilon = 5\%$ and the distribution of r is considered as the empirical distribution of the monthly stock returns over the whole period of 2000-2009, in other words, any monthly return vector observed in this period is as likely to occur.

Our answer: Let's consider the following approximation to the value-at-risk problem described above:

$$\begin{array}{ll} \underset{x,y}{\text{minimize}} & -y \\ \text{subject to} & r^T x \geq y\,,\, \forall\, r \in \mathcal{U} \qquad \sum_{i=1}^{n} x_i = 1 \qquad x \geq 0\,, \end{array}$$

where we will use the uncertainty set:

$$\mathcal{U}(r_0, \gamma) := \{ r \in \mathbb{R}^{10} \,|\, \|r - r_0\|_2 \leq \gamma \}\,,$$

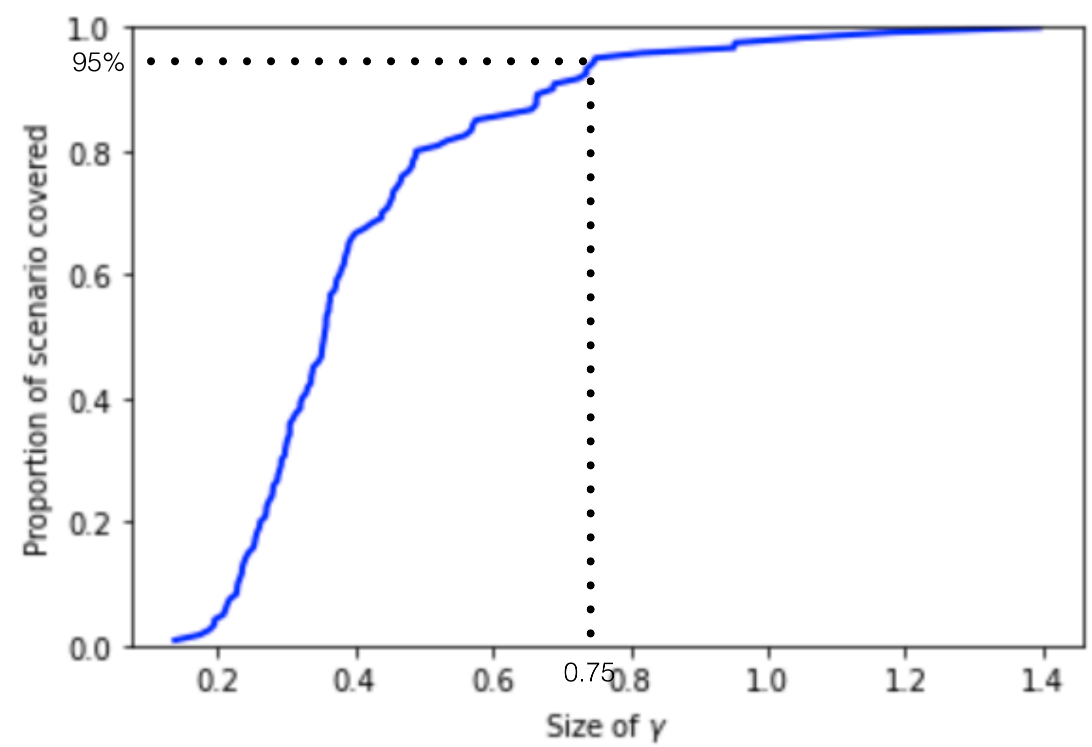
How would you calibrate « r_0 » and γ ?

Example: Portfolio with minimum VaR (Google Colab)

- We center the set at the mean

```
[ ] r0=np.mean(Rs,axis=1)
```

- We choose a radius such that the ball includes 95% of the samples



- In this example, the radius ends up being 0,75 while with a normal distribution we would use 0,24

Example: Portfolio with minimum VaR (Google Colab)

- The robust solution offers a bound on 95%-VaR of 21%
- In-sample, the VaR is 6.3%
- Out-of-sample, the VaR is 6.6%

```
[ ]  n = Rs.shape[0]

     #Create model
     model = ro.Model('RobustPorfolioVaR')
     # Define variables
     x=model.dvar(n)
     y=model.dvar(1)
     # Define uncertain parameters
     r=model.rvar(n)

     UncertaintySet=(rso.norm(r-r0,2)<=gamma)

     model.min(-y)
     model.st((r@x>=y).forall(UncertaintySet))
     model.st(sum(x)==1)
     model.st(x<=1)
     model.st(x>=0)

     model.solve(my_solver)
```

Risk-return tradeoff approximation

- In practice a decision maker is interested in the possible tradeoffs between risk and return

- In portfolio selection problem, this can be done with stochastic prog. or robust optimization

Stochastic Prog.

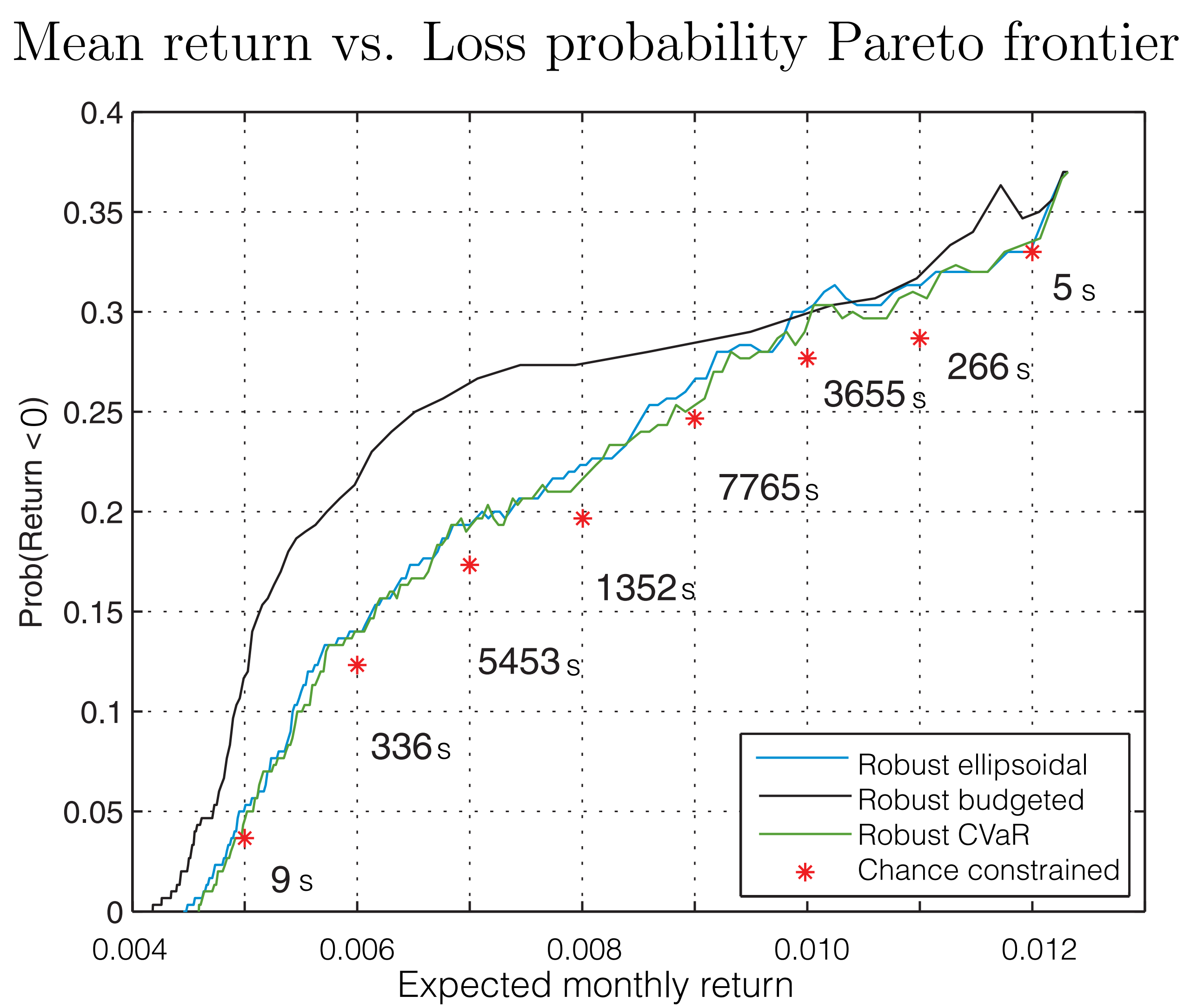
$$
\begin{aligned}
\text{maximize} \quad & \mathbb{E}\left[r^T x\right] \\
\text{subject to} \quad & \mathbb{P}(r^T x \geq 0) \geq 1 - \epsilon \\
& \sum_{i=1}^{n} x_i = 1 \\
& x \geq 0 \,.
\end{aligned}
$$

Robust optimization

$$
\begin{aligned}
\text{maximize} \quad & \mu^T x \\
\text{subject to} \quad & r^T x \geq 0 \,, \forall\, r \in \mathcal{U}(\Gamma) \\
& \sum_{i=1}^{n} x_i = 1 \\
& x \geq 0 \,,
\end{aligned}
$$

Risk-return tradeoff approximation

- In Bertsimas et al. 2011, the authors show that while RO only requires a fraction of computations needed by SP, it identifies solutions that are nearly optimal w.r.t. SP efficient frontier



RO as approximation to ambiguous chance constraints

Assumption 3.4. : Let $Z \in \mathbb{R}^m$ be a random vector for which the distribution is not known, yet what is known of the random vector is that all Z_i's are independent from each other and that each of them is symmetrically distributed on the interval $[-1, 1]$.

Theorem 3.5. : *Given some $\epsilon > 0$ and some random vector Z that satisfies assumption 3.4, one has the guarantee that any x satisfying the robust constraint*

$$a(z)^T x \leq b(z) \, , \forall \, z \in \mathcal{Z}_{ell}(\gamma) \ ,$$

where

$$\mathcal{Z}_{ell}(\gamma) := \{z \in \mathbb{R}^m \mid \|z\|_2 \leq \gamma\}$$

and $\gamma := \sqrt{2\ln(1/\epsilon)}$ is guaranteed to satisfy the following chance constraint

$$\mathbb{P}(a(Z)^T x \leq b(Z)) \geq 1 - \epsilon \ .$$

even though the distribution of Z is not known.

A corollary result for the budgeted uncertainty set

Corollary 3.8. : *Given some* $\epsilon > 0$ *and some random vector* Z *that satisfies assumption 3.4, one has the guarantee that any* x *satisfying the robust constraint*

$$a(z)^T x \leq b(z)\,,\,\forall\, z \in \mathcal{Z}_{budg}(\Gamma)\ ,$$

where

$$\mathcal{Z}_{budg}(\Gamma) := \{z \in \mathbb{R}^m \,|\, z_i \in [-1,\,1],\, \|z\|_1 \leq \Gamma\}$$

and $\Gamma := \sqrt{2m\ln(1/\epsilon)}$ *is guaranteed to satisfy the following chance constraint*

$$\mathbb{P}(a(Z)^T x \leq b(Z)) \geq 1 - \epsilon\ .$$

even though the distribution of Z *is not known.*

Definition of Coherent Risk Measures

In [2], Artzner et al. introduce for the first time the notion of a family of risk measures that are rational to employ. He indicates that such measures ρ should satisfy the following properties when defined in terms of an uncertain income:

- Translation invariance : the risk of a position to which we add a guaranteed income is reduced by the amount of the income, i.e. $\rho(Y + c) = \rho(Y) - c$ when c is certain
- Subadditivity: the risk of the sum of risky positions should be lower than the sum of the risks, i.e. $\rho(X + Y) \leq \rho(X) + \rho(Y)$
- Positive homogeneity : if the consequences of a risky position are scaled by the same positive amount $\lambda \geq 0$, then the risk should be scaled by the same amount, i.e. $\rho(\lambda Y) = \lambda\rho(Y)$
- Monotonicity: A risky position that is guaranteed to return larger income than another risky position should be considered less risky, i.e. $X \geq Y \Rightarrow \rho(X) \leq \rho(Y)$.
- Relevance : if a risky position has the potential of leading to a loss, then the risk should be strictly positive, i.e. $X \leq 0 \& X \neq 0 \Rightarrow \rho(X) > 0$.

Based on these five axioms, the authors are able to demonstrate that the risk measure must be representable in the following form:

$$\boxed{\rho(Y) := \sup_{F \in \mathcal{D}} \mathbb{E}_F[-Y] \,,}$$

RO as imposing a bound on a coherent risk measure

Theorem 3.9. : *Given a coherent risk measure $\rho(\cdot)$, there always exists a convex uncertainty set $\mathcal{Z}$ such that the no risk constraint*

$$\rho(b(Z) - a(Z)^T x) \leq 0$$

is equivalent to imposing the robust constraint

$$a(z)^T x \leq b(z)\,,\ \forall\, z \in \mathcal{Z};\,.$$

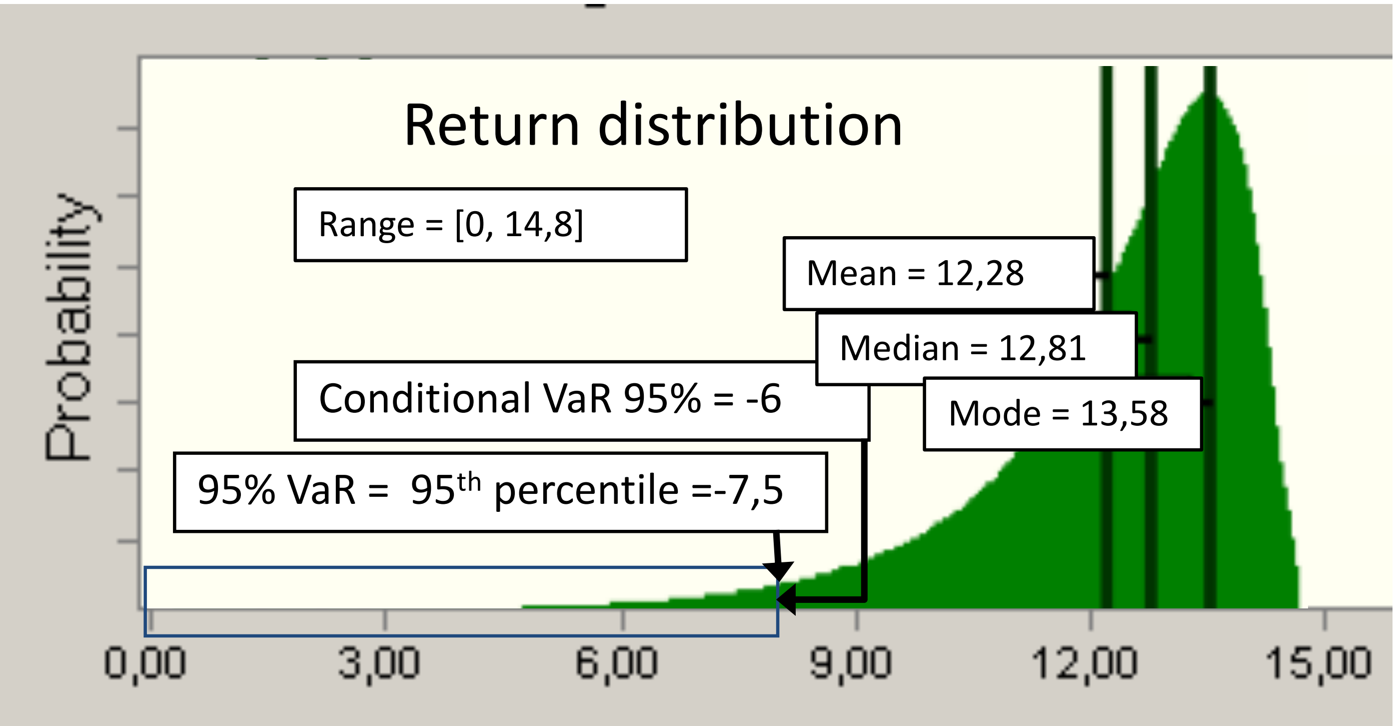
The converse is also true.

The Case of Conditional Value-at-Risk

Mathematically, the most popular representation for the CVaR measure appeared in [34] and takes the following form when the random variable Y represents an uncertain revenue

$$\mathrm{CVaR}_{1-\epsilon}(Y) := \inf_t t + \frac{1}{\epsilon}\mathbb{E}\left[\max(0, -Y - t)\right].$$

Intuitively, it is worth knowing that at optimum the value t^* will captures the value at risk for the given uncertain revenue so that



The Case of Conditional Value-at-Risk

Use Theorem 3.9 to show that when the distribution of z is

$$\mathbb{P}(Z = \bar{z}_i) = p_i, \forall\, i = 1, \ldots, K$$

the bounded CVaR constraint

$$CVaR_{1-\epsilon}(b(Z) - a(Z)^T x) \leq 0$$

can be equivalently reformulated as the following robust constraint

$$a(z)^T x \leq b(z), \forall\, z \in \mathcal{Z}_{\mathrm{CVaR}}(\epsilon)$$

where

$$\mathcal{Z}_{\mathrm{CVaR}}(\epsilon) = \{z \in \mathbb{R}^m \,|\, \exists q \in \mathbb{R}^K,\ q \geq 0,\ q_i \leq p_i/\epsilon,\ \sum_{i=1}^{K} q_i = 1,\ z = \sum_{i=1}^{K} \bar{z}_i q_i\}$$