ME 234(b): Convex approximations of chance constraints

Anushri Dixit

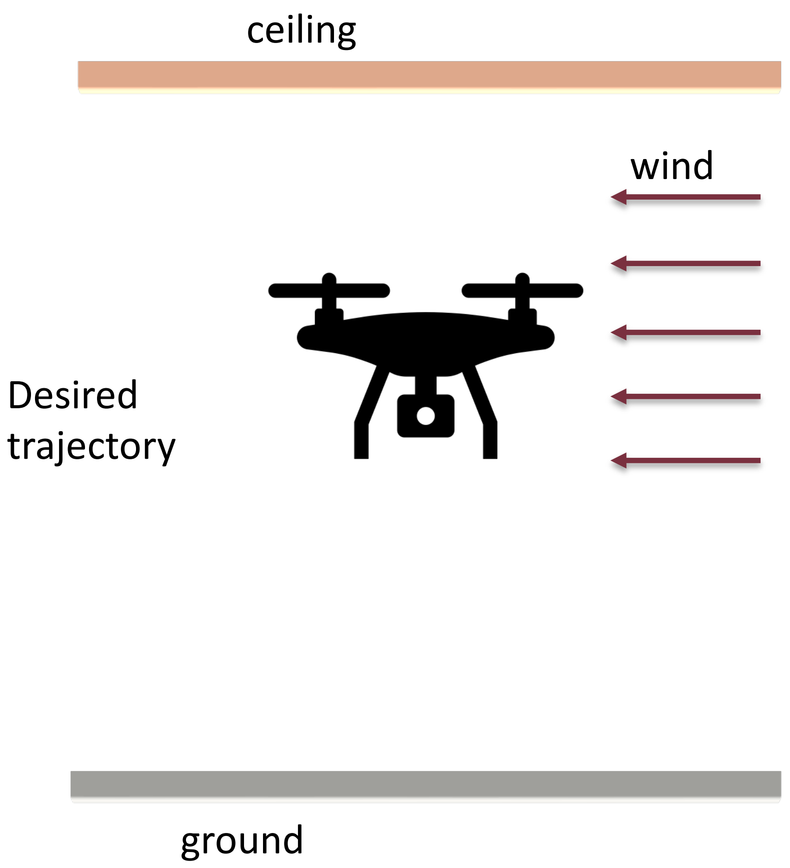
Spring 2022

Slides adapted from Convex Optimization by S. Boyd and Convex Approximations of Chance Constrained Programs by Nemirovski and Shapiro

Caltech

What kind of disturbances?

1. Gaussian distribution
2. **Bounded uncertainty**
3. Mean and variance
4. Discrete distribution



Bounded disturbance

Assume that the i.i.d. disturbances are bounded and lie in a convex polytope, i.e., $w(k) \in \mathcal{W}$ where,

$$\mathcal{W} = \{w \in \mathbb{R}^{n_w} : F_w w \leq b_w\}$$

Since every disturbance from this set is equally likely, we would like to **account for all possible disturbances.**

Caltech

Maxima over convex polytopes

Lemma:

Let $g : \mathbb{R}^{n_x} \times \mathbb{R}^{n_w} \to \mathbb{R}^{n_g}$ be a function of (x, w) convex in w for each x. Assume that $w \in \mathcal{W}$ where $\mathcal{W}$ is a polytope with vertices $\{\bar{w}_i\}_{i=1}^{n_\mathcal{W}}$. Then the constraint,

$$g(x, w) \leq 0 \quad \forall w \in \mathcal{W}$$

is satisfied iff

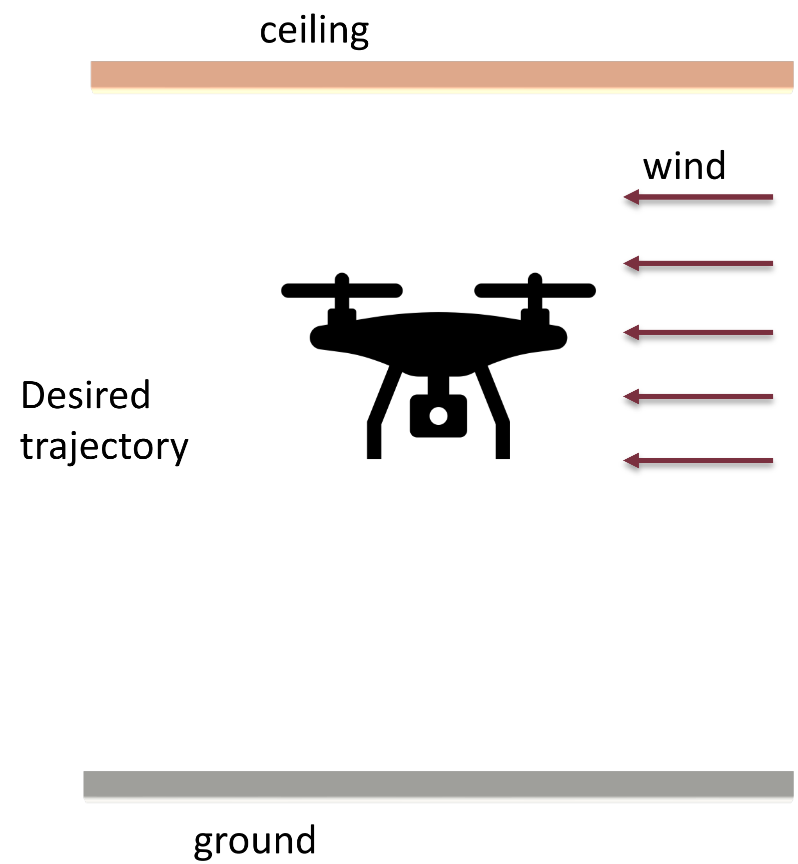
$$g(x, \bar{w}_i) \leq 0, \forall i \in \{1, \ldots, n_\mathcal{W}\}$$

This means that the maximum of a function convex in w, always occurs on the vertices of the polytope we are maximizing over.

See Lemma 10.1 in the textbook

What kind of disturbances?

1. **Gaussian distribution**
2. Bounded uncertainty
3. Mean and variance
4. Discrete distribution



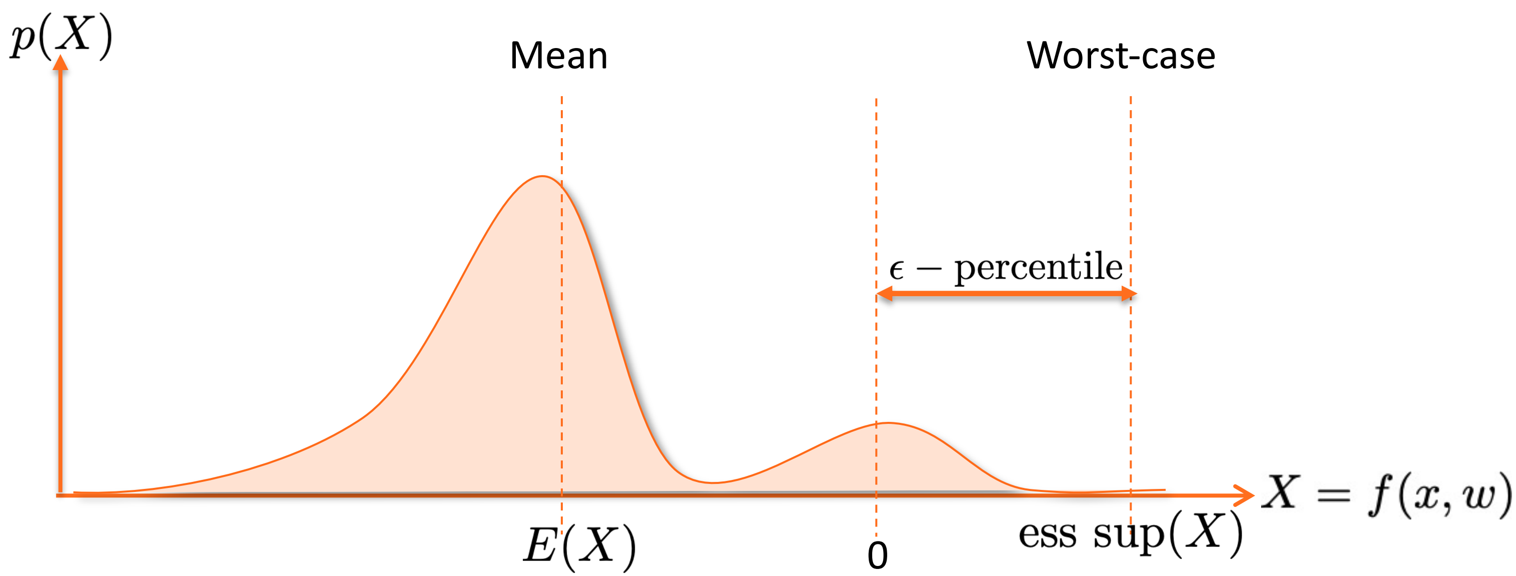
Caltech

Chance constraints

For a random variable w, a chance constraint satisfies the inequality

$$\text{Prob}(f(x,w) > 0) \leq \epsilon$$

This means a constraints of the form $f(x,w) \leq 0$ holds with at least $1-\epsilon$ confidence.



Caltech

Chance Constraints for Gaussian disturbances

Chance Constraints for Linear Inequality:

For a normal random variable $w \sim \mathcal{N}(0, \Sigma)$, we want to satisfy

$$\text{Prob}(a^T w + x > 0) \leq \epsilon, \quad a, w \in \mathbb{R}n_w, \; x \in \mathbb{R}$$

It follows that $a^T w + x \sim \mathcal{N}(x, a^T \Sigma a)$

Hence, $$\text{Prob}(a^T w + x > 0) = 1 - \text{Prob}(a^T w + x \leq 0)$$
$$= 1 - \underbrace{\Phi\left(\frac{-x}{\sqrt{a^t \Sigma a}}\right)}_{\text{cdf of } \mathcal{N}(0,1)}$$

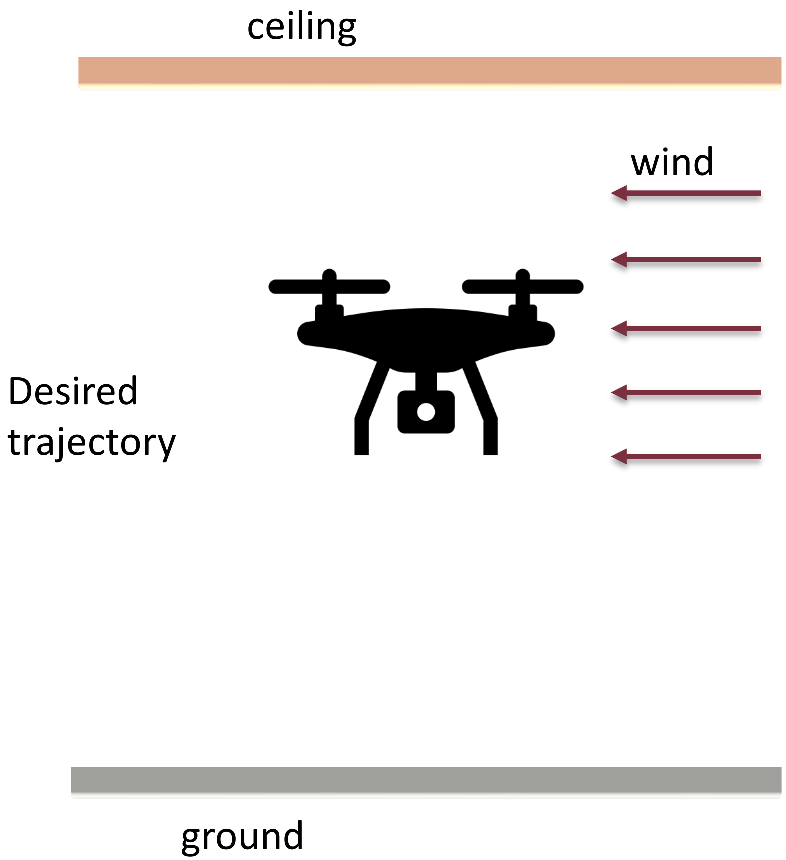
Or we can equivalently write the chance constraint as

$$x + \|\Sigma^{1/2} a\|_2 \Phi^{-1}(1 - \epsilon) \leq 0$$

Caltech

What kind of disturbances?

1. Gaussian distribution
2. Bounded uncertainty
3. **Mean and variance**
4. Discrete distribution



Caltech

Chance constraints for disturbances with known mean and variance

We no longer have ***exact reformulations*** of the chance constraint like we did for the Gaussian case:

$$\mathrm{Prob}(f(x, w) > 0) \leq \epsilon$$

But what are some ways in which we can find over-approximations of the chance constraint?

Idea: Find an upper bound for the probability of constraint violations.

Caltech

Chance constraints for disturbances with known mean and variance

We no longer have *exact reformulations* of the chance constraint like we did for the Gaussian case:

$$\mathrm{Prob}(f(x,w) > 0) \leq \epsilon$$

But what are some ways in which we can find over-approximations of the chance constraint?

Idea: Find an upper bound for the probability of constraint violations.

Solution #1: **Markov's inequality**

For a non-negative random variable, Z, we have

$$\mathrm{Prob}(Z > a) \leq \frac{\mathbb{E}[Z]}{a} \quad \forall a > 0$$

Caltech

Chance constraints for disturbances with known mean and variance

We no longer have *exact reformulations* of the chance constraint like we did for the Gaussian case:

$$\mathrm{Prob}(f(x,w) > 0) \leq \epsilon$$

But what are some ways in which we can find over-approximations of the chance constraint?

Idea: Find an upper bound for the probability of constraint violations.

Solution #1: **Markov's inequality**

For a non-negative random variable, Z, we have

$$\mathrm{Prob}(Z > a) \leq \frac{\mathbb{E}[Z]}{a} \quad \forall a > 0$$

Problem: Our random variable (the LHS of constraints on states with noise) may not be nonnegative.

Caltech

Chance constraints for disturbances with known mean and variance

Solution #1: **Markov's inequality**

For a non-negative random variable, Z, we have

$$\mathrm{Prob}(Z > a) \leq \frac{\mathbb{E}[Z]}{a} \quad \forall a > 0$$

Can we massage our problem into this form?

Extended version of Markov's inequality:

For a non-negative, nondecreasing function $\phi : \mathbb{R} \to \mathbb{R}$, it holds,

$$\mathrm{Prob}(Z > a) \leq \mathrm{Prob}(\phi(Z) \geq \phi(a)) \leq \frac{\mathbb{E}[\phi(Z)]}{\phi(a)}$$

Caltech

Chance constraints for disturbances with known mean and variance

Solution #1: **Markov's inequality**

For a non-negative random variable, Z, we have

$$\mathrm{Prob}(Z > a) \leq \frac{\mathbb{E}[Z]}{a} \quad \forall a > 0$$

Can we massage our problem into this form?

Extended version of Markov's inequality:

For a non-negative, nondecreasing function $\phi : \mathbb{R} \to \mathbb{R}$, it holds,

$$\mathrm{Prob}(Z > a) \leq \mathrm{Prob}(\phi(Z) \geq \phi(a)) \leq \frac{\mathbb{E}[\phi(Z)]}{\phi(a)}$$

Now, if $\phi(0) = 1$, and for any $\alpha > 0$,

$$\mathrm{Prob}(f(x,w) > 0) \leq \mathrm{Prob}\left(\phi\left(\frac{f(x,w)}{\alpha}\right) \geq \phi(0)\right) \leq \mathbb{E}\left[\phi\left(\frac{f(x,w)}{\alpha}\right)\right]$$

Caltech

Chance constraints for disturbances with known mean and variance

Solution #1.1**: Extension of Markov's inequality (Chebyshev's ineq)**

For a non-negative, nondecreasing function $\phi : \mathbb{R} \to \mathbb{R}$, and when $\phi(0) = 1$ with $\alpha > 0$, it holds that

$$\mathrm{Prob}(f(x,w) > 0) \leq \mathrm{Prob}\left(\phi\left(\frac{f(x,w)}{\alpha}\right) \geq \phi(0)\right) \leq \mathbb{E}\left[\phi\left(\frac{f(x,w)}{\alpha}\right)\right]$$

Let,

$$\phi(u) = (u+1)_+^2 = \begin{cases} 0 & u+1 < 0 \\ (u+1)^2 & u+1 \geq 0 \end{cases}$$

Caltech

Chance constraints for disturbances with known mean and variance

Solution #1.1**: Extension of Markov's inequality (Chebyshev's ineq)**

For a non-negative, nondecreasing function $\phi : \mathbb{R} \to \mathbb{R}$, and when $\phi(0) = 1$ with $\alpha > 0$, it holds that

$$\mathrm{Prob}(f(x,w) > 0) \le \mathrm{Prob}\left(\phi\left(\frac{f(x,w)}{\alpha}\right) \ge \phi(0)\right) \le \mathbb{E}\left[\phi\left(\frac{f(x,w)}{\alpha}\right)\right]$$

Let,

$$\phi(u) = (u+1)_+^2 = \begin{cases} 0 & u+1 < 0 \\ (u+1)^2 & u+1 \ge 0 \end{cases}$$

Plugging in $u = f(x,w)$, minimizing w.r.t α (with some conservative approximations) we show that $\mathrm{Prob}(f(x,w) > 0) \le \epsilon$ holds if,

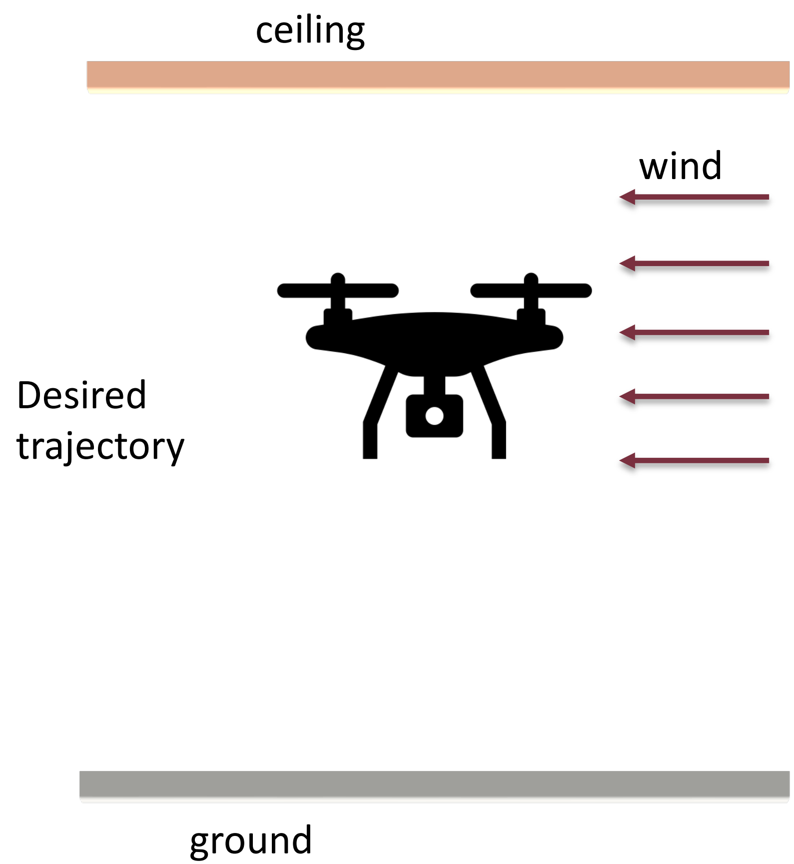
$$\mathbb{E}[f(x,w)] + \left((1-\epsilon)\mathbb{E}[f(x,w)^2]\right) \le 0$$

where, $f(x,w) = F_x x_{k|t} - b_x$ (the next steps are the same as before: write the inequalities in batch form and evaluate the expectations).

Caltech

What kind of disturbances?

1. Gaussian distribution
2. Bounded uncertainty
3. Mean and variance
4. **Discrete distributions and beyond**



Caltech

Chance constraints for general distributions

Solution #1.2**: Conditional Value-at-Risk**

For a non-negative, nondecreasing function $\phi : \mathbb{R} \to \mathbb{R}$, and when $\phi(0) = 1$ with $\alpha > 0$, it holds that,

$$\mathrm{Prob}(f(x,w) > 0) \leq \mathrm{Prob}\left(\phi\Big(\frac{f(x,w)}{\alpha}\Big) \geq \phi(0)\right) \leq \mathbb{E}\left[\phi\Big(\frac{f(x,w)}{\alpha}\Big)\right]$$

Let,

$$\phi(u) = (u+1)_+ = \begin{cases} 0 & u+1 < 0 \\ u+1 & u+1 \geq 0 \end{cases}$$

Plugging in $u = \dfrac{f(x,w)}{\alpha}$,

Caltech

Chance constraints for general distributions

Solution #1.2**: Conditional Value-at-Risk**

For a non-negative, nondecreasing function $\phi : \mathbb{R} \to \mathbb{R}$, and when $\phi(0) = 1$ with $\alpha > 0$, it holds that,

$$\mathrm{Prob}(f(x,w) > 0) \leq \mathrm{Prob}\left(\phi\left(\frac{f(x,w)}{\alpha}\right) \geq \phi(0)\right) \leq \mathbb{E}\left[\phi\left(\frac{f(x,w)}{\alpha}\right)\right]$$

Let,

$$\phi(u) = (u+1)_+ = \begin{cases} 0 & u+1 < 0 \\ u+1 & u+1 \geq 0 \end{cases}$$

Plugging in $u = \dfrac{f(x,w)}{\alpha}$,

$$\mathrm{Prob}(f(x,w) > 0) \leq \mathbb{E}\left[\left(\frac{f(x,w)}{\alpha} + 1\right)_+\right]$$

$$= \mathbb{E}\left[\alpha^{-1}(f(x,w) + \alpha)_+\right]$$

Hence, we need to satisfy the following constraint to satisfy our original chance constraint $\mathrm{Prob}(f(x,w) > 0) \leq \epsilon$

$$\mathbb{E}\left[\alpha^{-1}(f(x,w) + \alpha)_+\right] \leq \epsilon$$

Caltech

Chance constraints for general distributions

Solution #1.2**: Conditional Value-at-Risk**

We have,

$$\mathbb{E}\big[\alpha^{-1}(f(x,w)+\alpha)_+\big] \leq \epsilon$$

This can be rewritten as

$$\inf_{\alpha>0} \mathbb{E}\Big[\big((f(x,w)+\alpha)_+ - \alpha\epsilon\big)\Big] \leq 0$$

We do not affect the above result by replacing the $\inf\limits_{\alpha>0}$ with $\inf\limits_{\alpha\in\mathbb{R}}$, i.e,

$$\inf_{\alpha\in\mathbb{R}} \mathbb{E}\Big[(f(x,w)+\alpha)_+ - \alpha\epsilon\Big] \leq 0$$

Caltech

Chance constraints for general distributions

Solution #1.2**: Conditional Value-at-Risk**

We have,

$$\mathbb{E}\left[\alpha^{-1}(f(x,w)+\alpha)_+\right] \leq \epsilon$$

This can be rewritten as

$$\inf_{\alpha>0} \mathbb{E}\Big[\big((f(x,w)+\alpha)_+ - \alpha\epsilon\big)\Big] \leq 0$$

We do not affect the above result by replacing the $\inf\limits_{\alpha>0}$ with $\inf\limits_{\alpha\in\mathbb{R}}$, i.e,

$$\inf_{\alpha\in\mathbb{R}} \mathbb{E}\Big[(f(x,w)+\alpha)_+ - \alpha\epsilon\Big] \leq 0$$

The above constraint is equivalent to finding the Conditional Value-at-Risk (CVaR) of the function $f(x,w)$

$$\mathrm{CVaR}_{1-\epsilon}(f(x,w)) := \inf_{\tau\in\mathbb{R}} \left[\tau + \frac{1}{\epsilon}\mathbb{E}[f(x,w)-\tau]_+\right] \leq 0$$

Caltech

Conditional Value-at-Risk

We define the Conditional Value-at-Risk as,

$$\mathrm{CVaR}_{1-\epsilon}(Z) := \inf_{\tau \in \mathbb{R}} \left[\tau + \frac{1}{\epsilon} \mathbb{E}[Z - \tau]_+ \right], \quad \epsilon \in (0,1)$$

What does it mean?

Caltech

Conditional Value-at-Risk

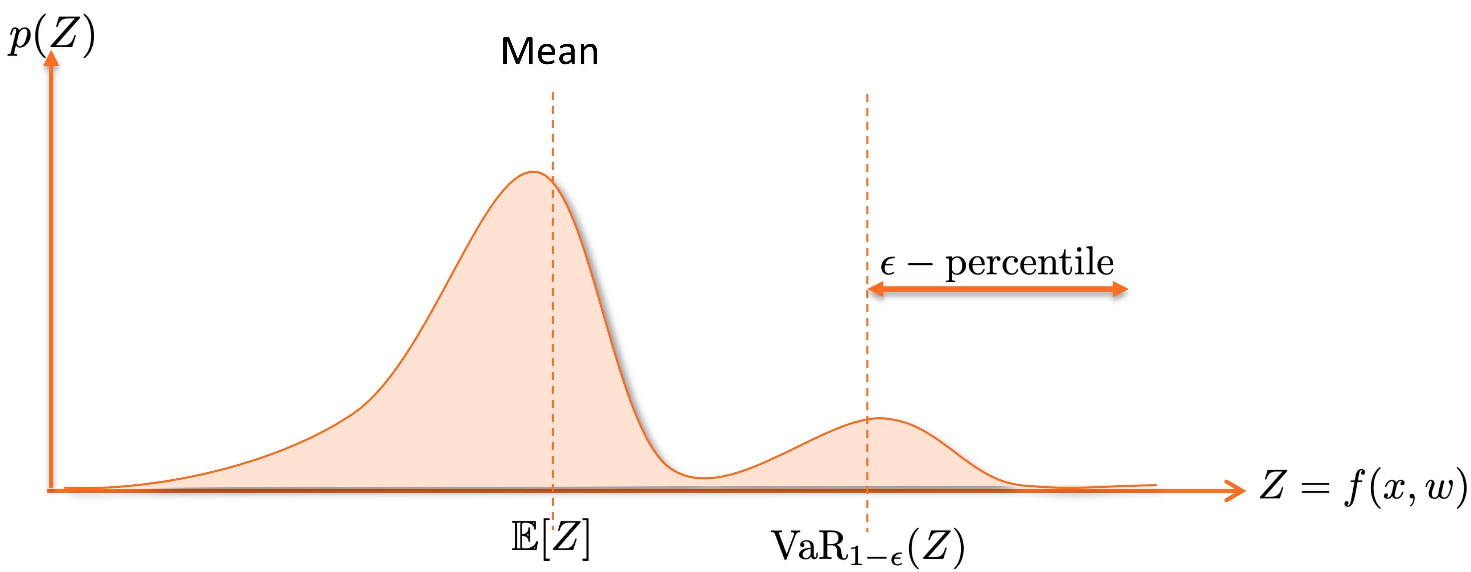
We define the Conditional Value-at-Risk as,

$$\mathrm{CVaR}_{1-\epsilon}(Z) := \inf_{\tau \in \mathbb{R}} \left[\tau + \frac{1}{\epsilon}\mathbb{E}[Z-\tau]_+\right], \quad \epsilon \in (0,1)$$

What does it mean?

To understand CVaR, let's first define the *Value-at-Risk* as,

$$\mathrm{VaR}_{1-\epsilon}(Z) := \inf\{z \in \mathbb{R} : \mathrm{Prob}(Z > z) \leq \epsilon\}$$



Caltech

Conditional Value-at-Risk

We define the Conditional Value-at-Risk as,

$$\mathrm{CVaR}_{1-\epsilon}(Z) := \inf_{\tau \in \mathbb{R}} \left[\tau + \frac{1}{\epsilon}\mathbb{E}[Z-\tau]_+\right], \quad \epsilon \in (0,1)$$

What does it mean?

To understand CVaR, let's first define the *Value-at-Risk* as,

$$\mathrm{VaR}_{1-\epsilon}(Z) := \inf\{z \in \mathbb{R} : \mathrm{Prob}(Z > z) \leq \epsilon\}$$

We can show that the optimal τ^* for the infimum in CVaR, is the Value-at-Risk, i.e., (for continuous distributions)

$$\begin{aligned} 0 &= \frac{d}{d\tau}\left(\tau + \frac{1}{\epsilon}\mathbb{E}\big[(Z-\tau)_+\big]\right) \\ &= 1 - \frac{1}{\epsilon}\mathrm{Prob}(Z > \tau) \end{aligned}$$

$$\implies \mathrm{Prob}(Z > \tau) = \epsilon$$

$$\implies \tau^* = \mathrm{VaR}_{1-\epsilon}(Z)$$

Caltech

Conditional Value-at-Risk

We define the Conditional Value-at-Risk as,

$$\mathrm{CVaR}_{1-\epsilon}(Z) := \inf_{\tau\in\mathbb{R}} \left[\tau + \frac{1}{\epsilon}\mathbb{E}[Z-\tau]_+\right], \quad \epsilon \in (0,1)$$

Properties:

1. CVaR is convex in the random variable Z (VaR is generally non-convex).
2. It provides an upper bound for VaR. It is in fact the conditional tail expectation.
3. It gives us a really nice way to switch between expectation-based and robust-planning by choosing ϵ.
4. For Gaussian distributions, it can be computed in closed form.
5. Is a part of a class of risk measures called *Coherent Risk Measures.*

Caltech

How to compute CVaR for discrete distributions

We have defined the Conditional Value-at-Risk for a random variable Z

$$\mathrm{CVaR}_{1-\epsilon}(Z) := \inf_{\tau \in \mathbb{R}} \left[\tau + \frac{1}{\epsilon}\mathbb{E}[Z-\tau]_+\right], \quad \epsilon \in (0,1)$$

Now, if Z is defined by a discrete distribution of size J, we denote the j[th] occurrence of the random variable by Z^j with probability of occurrence $p(j)$. Hence CVaR can be reformulated as,

$$\begin{aligned} \mathrm{CVaR}_{1-\epsilon}(Z) = \inf_{z \in \mathbb{R}, s \in \mathbb{R}^J} & \sum_{j=1}^{J} p(j)\left(z + \frac{s(j)}{\epsilon}\right) \\ \text{s.t.} \quad & s(j) \geq 0, \\ & s(j) + z \geq Z^j \end{aligned}$$

Caltech

MPC reformulation using CVaR

We started with the following MPC problem,

$$
\begin{aligned}
J_t^*(x(t)) = \min_{U_t} \quad & \mathbb{E}\Big[x_{t+N|t}^T P x_{t+N|t} + \sum_{k=t}^{t+N-1} \left(x_{k|t}^T Q x_{k|t} + u_{k|t}^T R u_{k|t}\right)\Big] \\
\text{s.t.} \quad & x_{k+1|t} = A x_{k|t} + B u_{k|t} + D w_{k|t}, \quad \forall k \in \{t, t+1, \dots t+N-1\} \\
& \text{Prob}(x_{k|t} \notin \mathcal{X}) \leq \epsilon, \ u_{k|t} \in \mathcal{U}, \quad \forall k \in \{t, t+1, \dots t+N-1\} \\
& x_{t|t} = x(t)
\end{aligned}
$$

Using CVaR, we can get a “middle ground” between robust planning and stochastic MPC (that uses an expectation-based cost and chance constraints).

$$
\begin{aligned}
J_t^*(x(t)) = \min_{U_t} \quad & \text{CVaR}_{1-\tilde{\epsilon}}\Big[x_{t+N|t}^T P x_{t+N|t} + \sum_{k=t}^{t+N-1} \left(x_{k|t}^T Q x_{k|t} + u_{k|t}^T R u_{k|t}\right)\Big] \\
\text{s.t.} \quad & x_{k+1|t} = A x_{k|t} + B u_{k|t} + D w_{k|t}, \quad \forall k \in \{t, t+1, \dots t+N-1\} \\
& \text{CVaR}_{1-\epsilon}(F_x x_{k|t} - b_x) \leq 0, \ u_{k|t} \in \mathcal{U}, \quad \forall k \in \{t, t+1, \dots t+N-1\} \\
& x_{t|t} = x(t)
\end{aligned}
$$

Caltech