

# Alternatives to Expectation

John E. Mitchell

Department of Mathematical Sciences  
RPI, Troy, NY 12180 USA

April 2019

# Stochastic two stage MIPs

We will focus on **stochastic two stage mixed integer programs with recourse**. The general formulation can be written

$$\begin{aligned} \min_x \quad & c^T x + E_{\xi} [Q(x, \xi)] \\ \text{subject to} \quad & Ax = b \\ & x \in \mathcal{X} \end{aligned} \tag{1}$$

where the first stage decisions are  $x \in \mathcal{X} \subseteq \mathbf{R}^{n_1}$ , the constraint matrix  $A$  is  $m_1 \times n_1$ ,  $b \in \mathbf{R}^{m_1}$ ,  $c \in \mathbf{R}^{n_1}$ ,  $\xi$  is the uncertainty, and  $Q(x, \xi)$  is the cost of the recourse decision when the first stage decision is  $x$  and the uncertainty is  $\xi$ .

Thus,  $Q(x, \xi)$  is the second stage cost.

We take the expectation of the second stage cost over all scenarios  $\xi$ .

## Second stage problem

The second stage cost

$$Q(x, \xi) = \min_y \quad q^T y \quad (2)$$

subject to  $Wy = h(\xi) - T(\xi)x$   
 $y \in \mathcal{Y}$

where  $y \in \mathcal{Y} \subseteq \mathbf{R}^{n_2}$ ,  $W$  is a fixed  $m_2 \times n_2$  matrix, the right hand side  $h(\xi) \in \mathbf{R}^{m_2}$  depends on the uncertainty  $\xi$ , and the  $m_2 \times n_2$  technology matrix  $T(\xi)$  also depends on  $\xi$ .

Note that the second stage optimization is over  $y$ , with  $x$  taken as a parameter.

# Risk measures

Sometimes the expectation does not capture the risk sufficiently.

There are several alternatives.

# Robust optimization

Replace expectation by worst possible outcome:

$$\begin{aligned} \min_{x,v} \quad & c^T x + v \\ \text{subject to} \quad & Ax = b \\ & v \geq Q(x, \xi) \quad \forall \xi \\ & x \in \mathcal{X} \end{aligned} \tag{3}$$

This doesn't require knowledge of the density function for  $\xi$ .

This is conservative.

## More on robust optimization

A slightly less conservative choice is to only consider  $\xi$  within some set:

$$\begin{aligned} \min_x \quad & c^T x + v \\ \text{subject to} \quad & Ax = b \\ & v \geq Q(x, \xi) \quad \forall \xi \in \Xi \\ & x \in \mathcal{X} \end{aligned} \tag{4}$$

Typical choices for  $\Xi$  are ellipsoids or boxes.

# Chance constrained formulations

Require that  $Q(x, \xi)$  be below some threshold  $\beta$  with probability  $1 - \alpha$ :

$$\begin{array}{ll} \min_x & c^T x \\ \text{subject to} & Ax = b \\ & \Pr(Q(x, \xi) > \beta) \leq \alpha \\ & x \in \mathcal{X} \end{array} \quad (5)$$

# Value-at-Risk

A special case is minimizing the **Value-at-Risk (VaR)**:

$$\begin{array}{ll} \min_{x, \eta} & \eta \\ \text{subject to} & Ax = b \\ & \Pr(Q(x, \xi) > \eta) \leq \alpha \\ & x \in \mathcal{X} \end{array} \quad (6)$$

With  $\alpha = 0.1$ , this is choosing  $x$  so that 90th percentile outcome is as good as possible.

When the uncertainty  $\xi$  consists of a finite number of scenarios, these problems can be modeled as equivalent **integer programs**.



## Conditional Value-at-Risk (CVaR)

Instead of taking the expectation of all possible outcomes, take the expectation of the worst  $\alpha$  outcomes.

Perhaps, take the expectation of the worst 10% of outcomes.

$$\begin{array}{ll} \min_x & c^T x + E_{\xi} [Q(x, \xi) \mid Q(x, \xi) \text{ is in worst } \alpha\% \text{ of outcomes}] \\ \text{subject to} & Ax = b \\ & x \in \mathcal{X} \end{array} \quad (7)$$

## CVaR as an MIP

With a finite set of scenarios, this can be modeled as a mixed integer program:

$$\begin{array}{llll} \min_{x,v,\eta} & c^T x & + & \eta & + & \frac{1}{\alpha} \sum_s p_s v_s \\ \text{subject to} & Ax & & = & b \\ & Q(x, \xi_s) & - & \eta & \leq & v_s \quad \forall s \\ & & & v_s & \geq & 0 \quad \forall s \\ & x & & \in & \mathcal{X} \end{array} \quad (8)$$

This LP is only slightly more complicated than the original stochastic program (1).

The values  $v_s$  give the excess of  $Q(x, \xi_s)$  over  $\eta$ .

The value of  $\eta$  can fall in a range: anywhere between the best  $(1 - \alpha)$  outcomes and the worst  $\alpha$  outcomes.

# Coherent risk measures

Can combine minimizing the expected value with minimizing some measure of risk.

There is a theory of **coherent risk measures** which possess various nice properties:

*convexity, monotonicity, translation equivalence, positive homogeneity*

# Variance

**Variance not a coherent risk measure.** It fails monotonicity, which requires:

*if the outcome is always worse then the risk measure should be higher*

For example, assume we have two strategies each with two possible outcomes:

- Strategy A: in scenario  $\xi^1$ , value of outcome is 10; in scenario  $\xi^2$ , value of outcome is 20.
- Strategy B: in scenario  $\xi^1$ , value of outcome is 25; in scenario  $\xi^2$ , value of outcome is 25.

So Strategy A always outperforms Strategy B, so we should choose strategy A. The variance of the outcomes with Strategy A is positive, whereas it is zero with Strategy B. So if we were seeking to minimize variance, we would choose Strategy B.

# Variance

**Variance not a coherent risk measure.** It fails monotonicity, which requires:

*if the outcome is always worse then the risk measure should be higher*

For example, assume we have two strategies each with two possible outcomes:

- Strategy A: in scenario  $\xi^1$ , value of outcome is 10; in scenario  $\xi^2$ , value of outcome is 20.
- Strategy B: in scenario  $\xi^1$ , value of outcome is 25; in scenario  $\xi^2$ , value of outcome is 25.

So Strategy A always outperforms Strategy B, so we should choose strategy A. The variance of the outcomes with Strategy A is positive, whereas it is zero with Strategy B. So if we were seeking to minimize variance, we would choose Strategy B.

## Variance

**Variance not a coherent risk measure.** It fails monotonicity, which requires:

*if the outcome is always worse then the risk measure should be higher*

For example, assume we have two strategies each with two possible outcomes:

- Strategy A: in scenario  $\xi^1$ , value of outcome is 10; in scenario  $\xi^2$ , value of outcome is 20.
- Strategy B: in scenario  $\xi^1$ , value of outcome is 25; in scenario  $\xi^2$ , value of outcome is 25.

So Strategy A always outperforms Strategy B, so we should choose strategy A. The variance of the outcomes with Strategy A is positive, whereas it is zero with Strategy B. So if we were seeking to minimize variance, we would choose Strategy B.

## Other risk measures

**CVaR** is coherent.

**VaR** is not coherent: it fails convexity.