# Alternatives to Expectation

John E. Mitchell

Department of Mathematical Sciences
RPI, Troy, NY 12180 USA

April 2019

# Stochastic two stage ~~LPs~~ MIPs

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

$$\begin{array}{llcl} \min_x & c^T x & + & E_\xi\left[Q(x,\xi)\right] \\ \text{subject to} & Ax & = & b \\ & x & \in & \mathcal{X} \end{array} \qquad (1)$$

where the first stage decisions are $x \in \mathcal{X} \subseteq \boldsymbol{R}^{n_1}$, the constraint matrix $A$ is $m_1 \times n_1$, $b \in \boldsymbol{R}^{m_1}$, $c \in \boldsymbol{R}^{n_1}$, $\xi$ 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

$$Q(x,\xi) = \begin{array}{lrcl} \min_y & q^T y & & \\ \text{subject to} & Wy & = & h(\xi) - T(\xi)x \\ & y & \in & \mathcal{Y} \end{array} \tag{2}$$

where $y \in \mathcal{Y} \subseteq \boldsymbol{R}^{n_2}$, $W$ is a fixed $m_2 \times n_2$ matrix, the right hand side $h(\xi) \in \boldsymbol{R}^{m_2}$ 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:

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

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:

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

Typical choices for $\Xi$ are ellipsoids or boxes. (Bertsimas & Sim)

Eg: demands $d$ are stochastic.
Only a few demands differ from predicted means.

$d_i^s \in [d_i^{sl}, d_i^{su}]$

$d_s^i \leq \bar{d}_i^s + \alpha_i^s z_i$
$d_s^i \geq \bar{d}_i^s - \alpha_i^s z_i$

$\sum_i z_i \leq K$
$z_i$ binary

# Chance constrained formulations

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

$$\begin{array}{lrcl} \min_x & c^T x & & \\ \text{subject to} & Ax & = & b \\ & \Pr_{\xi}(Q(x,\xi) \geq \beta) & \leq & \alpha \\ & x & \in & \mathcal{X} \end{array} \tag{5}$$

$>$

Eg: 20 equally likely scenarios. $\alpha = 0.1$

So at most two scenarios have $Q(x,\xi) > \beta$

$Q(x,\xi) - t_s \leq \beta \qquad t_s \leq M z_s \qquad z_s$ binary, $t_s \geq 0$

$\sum_s z_s \leq 2$

# Value-at-Risk

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

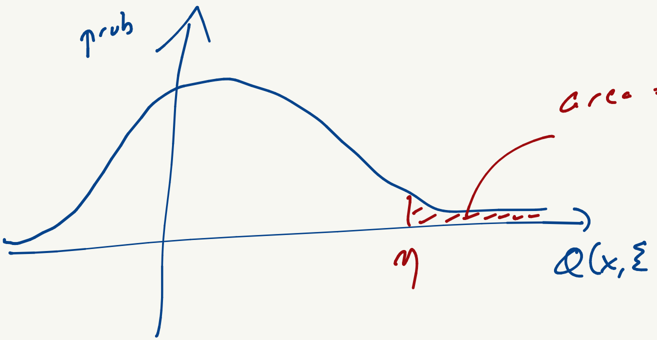
$$\begin{array}{lrcl} \min_{x,\eta} & \eta & & \\ \text{subject to} & Ax & = & b \\ & \Pr(Q(x,\xi) \geq \eta) & \leq & \alpha \\ & x & \in & \mathcal{X} \end{array} \tag{6}$$

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.**

Given x, have graph of $Q(x, \xi)$

bell chart

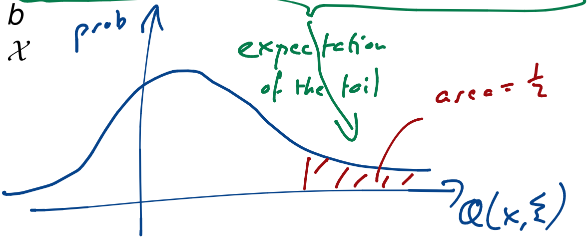


# Conditional Value-at-Risk (CVaR)

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.

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



# CVaR as an LP

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

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

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.