SDDP.jl

Stochastic dual dynamic programming in Julia

Oscar Dowson

Department of Industrial Engineering and Management Sciences
Northwestern

October 23, 2018

Outline

Some background

Risk

SDDP.jl tutorial

Outline

Northwestern ENGINEERING

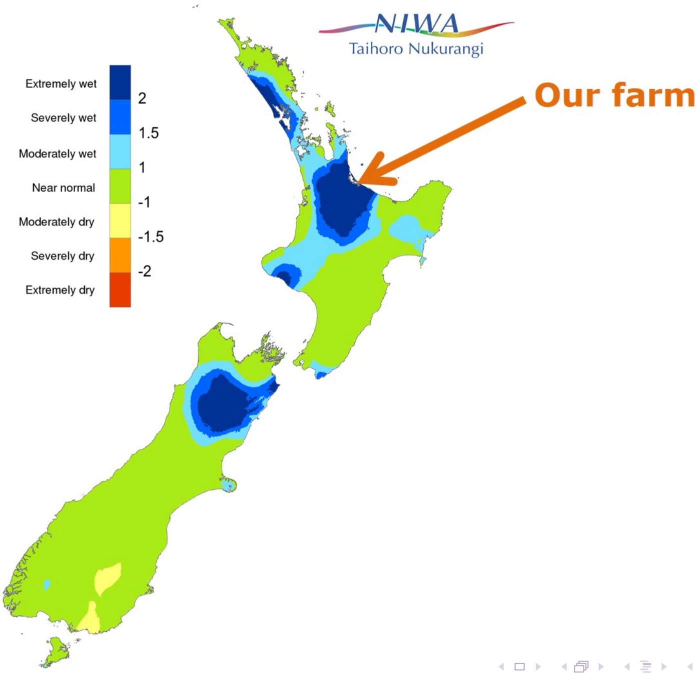
Some background

Risk

SDDP.jl tutorial


In my opinion, all palm oil should be banned.

SPI Drought Index for 9am 27/08/2017 to 9am 26/09/2017



What went wrong?

- A high milk price was forecast.

What went wrong?

Northwestern ENGINEERING

- A high milk price was forecast.
  - Farmers don't get paid for milk on delivery

What went wrong?

- A high milk price was forecast.
  - Farmers don't get paid for milk on delivery
  - After each year, they are back-paid an *end-of-season* milk price

What went wrong?

Northwestern ENGINEERING

- A high milk price was forecast.
  - Farmers don't get paid for milk on delivery
  - After each year, they are back-paid an *end-of-season* milk price
  - We can *forecast* the end-of-season price during the year

What went wrong?

Northwestern ENGINEERING

- A high milk price was forecast.
  - Farmers don't get paid for milk on delivery
  - After each year, they are back-paid an *end-of-season* milk price
  - We can *forecast* the end-of-season price during the year
  - But with high uncertainty

What went wrong?

Northwestern ENGINEERING

- A high milk price was forecast.
  - Farmers don't get paid for milk on delivery
  - After each year, they are back-paid an *end-of-season* milk price
  - We can *forecast* the end-of-season price during the year
  - But with high uncertainty
- So the farmer kept a high number of cows,

What went wrong?

Northwestern ENGINEERING

- A high milk price was forecast.
  - Farmers don't get paid for milk on delivery
  - After each year, they are back-paid an *end-of-season* milk price
  - We can *forecast* the end-of-season price during the year
  - But with high uncertainty
- So the farmer kept a high number of cows,
- and they had plenty of feed so they didn't buy more.

What went wrong?

Northwestern ENGINEERING

- A high milk price was forecast.
  - Farmers don't get paid for milk on delivery
  - After each year, they are back-paid an *end-of-season* milk price
  - We can *forecast* the end-of-season price during the year
  - But with high uncertainty
- So the farmer kept a high number of cows,
- and they had plenty of feed so they didn't buy more.
- But, a wet year had left paddocks damaged.

What went wrong?

- A high milk price was forecast.
  - Farmers don't get paid for milk on delivery
  - After each year, they are back-paid an *end-of-season* milk price
  - We can *forecast* the end-of-season price during the year
  - But with high uncertainty
- So the farmer kept a high number of cows,
- and they had plenty of feed so they didn't buy more.
- But, a wet year had left paddocks damaged.
- 30 years of experience said: we can't have a bad Summer, Autumn, Winter, AND Spring

What went wrong?

- A high milk price was forecast.
  - Farmers don't get paid for milk on delivery
  - After each year, they are back-paid an *end-of-season* milk price
  - We can *forecast* the end-of-season price during the year
  - But with high uncertainty
- So the farmer kept a high number of cows,
- and they had plenty of feed so they didn't buy more.
- But, a wet year had left paddocks damaged.
- 30 years of experience said: we can't have a bad Summer, Autumn, Winter, AND Spring right?

The dairy farmer problem

Northwestern ENGINEERING

maximise:	revenue from milk production less operating costs
by deciding:	the number of cows to farm
	the quantity of grass to feed
	the quantity of supplement to feed
	when to dry-off the herd
subject to:	obtaining a high Body Condition Score at the end of the season
	uncertainty in grass growth
	uncertainty in the milk price

To learn more about this, come to my talk

Wednesday, October 31, 2018 @ 2:00 p.m. Room 274 Animal Sciences Bldg.

Outline

Why care about risk?

If the tail matters more than the average.

- For a farmer, bad years mean cows starve or you go bankrupt and lose your farm
- For electricity generators, bad years mean blackouts in cities
- In finance, bad years mean losing all your money

Definition

A *risk measure* $\mathbb{F}$ is a function that maps a random variable to a real number.

Definition

A *risk measure* $\mathbb{F}$ is a function that maps a random variable to a real number.

Math

We restrict our attention to random variables with a finite sample space $\Omega := \{z_1, z_2, \ldots, z_K\}$ equipped with a sigma algebra of all subsets of Ω and respective (strictly positive) probabilities $\{p_1, p_2, \ldots, p_K\}$.

We denote the random variable with the uppercase Z.

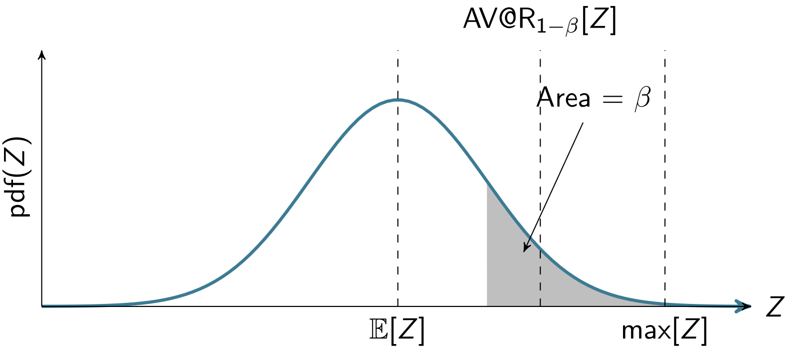
Definition

The *Average Value-at-Risk* at the β quantile ($\mathrm{AV@R}_{1-\beta}$) is:

$$\mathrm{AV@R}_{1-\beta}[Z] = \inf_{\zeta} \left\{ \zeta + \frac{1}{\beta} \sum_{k=1}^{K} p_k (z_k - \zeta)_+ \right\},$$

where $(x)_+ = \max\{0, x\}$. (Rockafellar and Uryasev 2002)

Note that when $\beta = 1$, $\mathrm{AV@R}_{1-\beta}[Z] = \mathbb{E}[Z]$, and $\lim_{\beta \to 0} \mathrm{AV@R}_{1-\beta}[Z] = \max[Z]$.



$\mathrm{AV@R}_{1-\beta}[Z]$
Area $= \beta$
pdf(Z)
$\mathbb{E}[Z]$
$\max[Z]$
Z

Northwestern ENGINEERING

Definition

A *coherent* risk measure is a risk measure $\mathbb{F}$ that satisfies the axioms of Artzner et al. 1999. For two discrete random variables Z_1 and Z_2, each with drawn from a sample space with K elements, the axioms are:

- **Monotonicity**: If $Z_1 \leq Z_2$, then $\mathbb{F}[Z_1] \leq \mathbb{F}[Z_2]$.
- **Sub-additivity**: For Z_1, Z_2, then $\mathbb{F}[Z_1 + Z_2] \leq \mathbb{F}[Z_1] + \mathbb{F}[Z_2]$.
- **Positive homogeneity**: If $\lambda \geq 0$ then $\mathbb{F}[\lambda Z] = \lambda \mathbb{F}[Z]$.
- **Translation equivariance**: If $a \in \mathbb{R}$ then $\mathbb{F}[Z + a] = \mathbb{F}[Z] + a$.

Positive homogeneity and sub-additivity give:

- **Convexity**: For $\lambda \in [0, 1]$,
  $\mathbb{F}[\lambda Z_1 + (1 - \lambda) Z_2] \leq \lambda \mathbb{F}[Z_1] + (1 - \lambda) \mathbb{F}[Z_2]$.

We can also define coherent risk measures in terms of *risk sets*. That is, a coherent risk measure $\mathbb{F}$ has a dual representation that can be viewed as taking the expectation of the random variable with respect to the worst probability distribution within some set $\mathfrak{A}$ of possible distributions:

$$\mathbb{F}[Z] = \sup_{\xi \in \mathfrak{A}} \mathbb{E}_{\xi}[Z] = \sup_{\xi \in \mathfrak{A}} \sum_{k=1}^{K} \xi_k z_k, \tag{1}$$

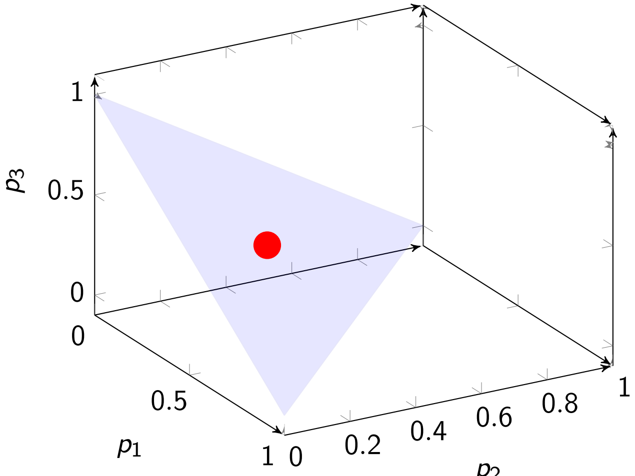
where $\mathfrak{A}$ is a convex subset of:

$$\mathfrak{P} = \left\{ \xi \in \mathbb{R}^K \ : \ \sum_{k=1}^{K} \xi_k = 1, \ \xi \geq 0 \right\}.$$

Static Risk Measures

Northwestern ENGINEERING

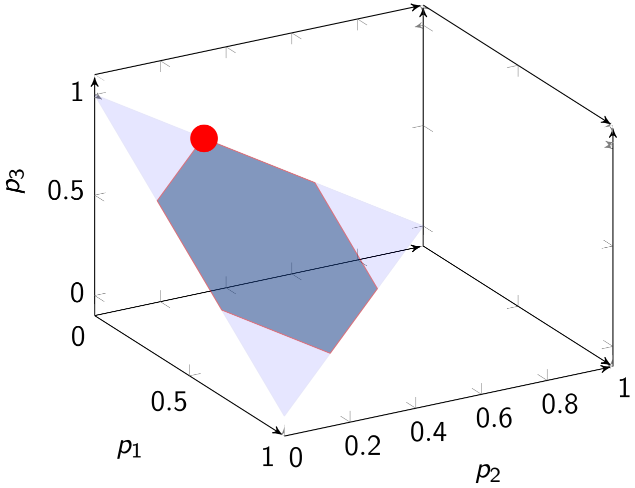
If $\mathfrak{A}$ is a singleton, containing only the original probability distribution, then the risk measure $\mathbb{F}$ is equivalent to the expectation operator.



Static Risk Measures

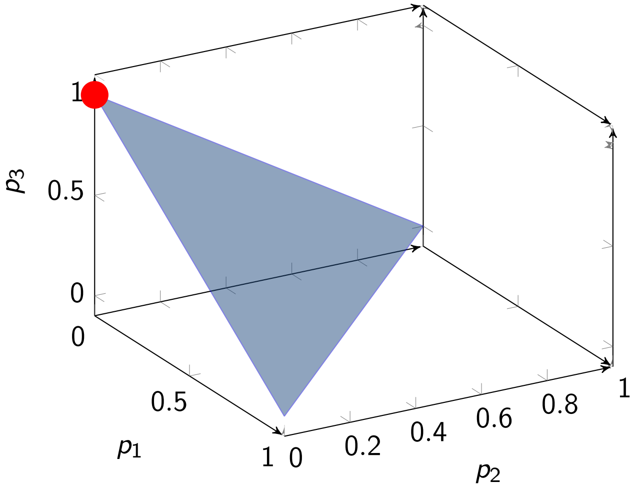
Northwestern ENGINEERING

If $\mathfrak{A} = \left\{ \xi \in \mathfrak{P} \mid \xi_k \leq \frac{p_k}{\beta},\ k = 1, 2, \ldots, K \right\}$, then the risk measure $\mathbb{F}$ is equivalent to AV@R$_{1-\beta}$.



Static Risk Measures

If $\mathfrak{A} = \mathfrak{P}$, then $\mathbb{F}$ is the Worst-case risk measure.



Dynamic Risk Measures

Northwestern ENGINEERING

From static to dynamic risk measures. There are three main views:

End-of-horizon: most natural

See, e.g., Pflug and Pichler 2016; Baucke, Downward, and Zakeri 2018

$$\mathbb{F}[Z_1, Z_2, Z_3] = \mathbb{F}_{\omega_1,\omega_2,\omega_3}[Z_1 + Z_2 + Z_3]$$

Nested: easiest to compute

See, e.g., Ruszczyński 2010; Philpott, de Matos, and Finardi 2013

$$\mathbb{F}[Z_1, Z_2, Z_3] = \mathbb{F}_{\omega_1}[Z_1 + \mathbb{F}_{\omega_2|\omega_1}[Z_2 + \mathbb{F}_{\omega_3|\omega_1,\omega_2}[Z_3]\,]\,]$$

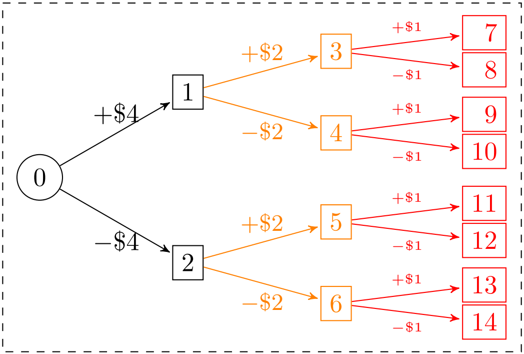
Expected conditional: a compromise

See, e.g., Homem-de-Mello and Pagnoncelli 2016

$$\mathbb{F}[Z_1, Z_2, Z_3] = \mathbb{F}_{\omega_1}[Z_1] + \mathbb{E}_{\omega_1}[\mathbb{F}_{\omega_2|\omega_1}[Z_2] + \mathbb{E}_{\omega_2|\omega_1}[\mathbb{F}_{\omega_3|\omega_1,\omega_2}[Z_3]]]$$

End-of-horizon risk measure

$$\mathbb{F}[Z_1, Z_2, Z_3] = \mathbb{F}_{\omega_1,\omega_2,\omega_3}[Z_1 + Z_2 + Z_3]$$

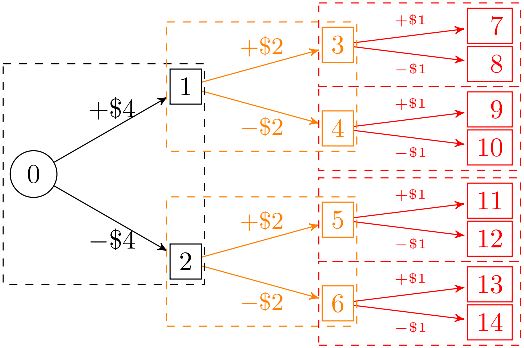


Dynamic Risk Measures

Northwestern ENGINEERING

Nested risk measure

$$\mathbb{F}[Z_1, Z_2, Z_3] = \mathbb{F}_{\omega_1}[Z_1 + \mathbb{F}_{\omega_2|\omega_1}[Z_2 + \mathbb{F}_{\omega_3|\omega_1,\omega_2}[Z_3]\,]\,]$$



Risk in SDDP

Northwestern ENGINEERING

Recall our favourite dynamic programming recursion:

$$\begin{aligned} V_t(x_t, \omega_t) = \min_{u_t} \quad & C_t(x_t, u_t, \omega_t) + \mathop{\mathbb{E}}_{\omega_{t+1} \in \Omega_{t+1}} \left[V_{t+1}(x_{t+1}, \omega_{t+1}) \right] \\ \text{s.t.} \quad & x_{t+1} = T_t(x_t, u_t, \omega_t) \\ & u_t \in U_t(x_t, \omega_t), \end{aligned}$$

where the decision-rule $\pi_t(x_t, \omega_t)$ takes the value of u_t in the optimal solution.

Risk in SDDP

Northwestern ENGINEERING

Recall our favourite dynamic programming recursion:

$$\begin{aligned} V_t(x_t, \omega_t) = \min_{u_t} \quad & C_t(x_t, u_t, \omega_t) + \underset{\omega_{t+1} \in \Omega_{t+1}}{\mathbb{F}} \left[V_{t+1}(x_{t+1}, \omega_{t+1}) \right] \\ \text{s.t.} \quad & x_{t+1} = T_t(x_t, u_t, \omega_t) \\ & u_t \in U_t(x_t, \omega_t), \end{aligned}$$

where the decision-rule $\pi_t(x_t, \omega_t)$ takes the value of u_t in the optimal solution.

Northwestern ENGINEERING

Recall

Given an original probability distribution $\{p_1, p_2, \ldots, p_K\}$ and a coherent risk measure $\mathbb{F}$, there exists a *changed* probability distribution $\{\xi_1, \xi_2, \ldots, \xi_K\}$ such that $\mathbb{F}[Z] = \mathbb{E}_\xi[Z]$.

The meat of the matter. Consider the following (re-stated) proposition from Philpott, de Matos, and Finardi 2013:

Proposition

Suppose for each $\omega \in \Omega$, that $\lambda(\bar{x}, \omega)$ is a subgradient of $V(x, \omega)$ at $\bar{x}$. Then, given $\mathbb{F}[V(\bar{x}, \omega)] = \mathbb{E}_{\xi}[V(\bar{x}, \omega)]$, $\mathbb{E}_{\xi}[\lambda(\bar{x}, \omega)]$ is a subgradient of $\mathbb{F}[V(x, \omega)]$ at $\bar{x}$.

The meat of the matter. Consider the following (re-stated) proposition from Philpott, de Matos, and Finardi 2013:

Proposition

Suppose for each $\omega \in \Omega$, that $\lambda(\bar{x}, \omega)$ is a subgradient of $V(x, \omega)$ at $\bar{x}$. Then, given $\mathbb{F}[V(\bar{x}, \omega)] = \mathbb{E}_\xi[V(\bar{x}, \omega)]$, $\mathbb{E}_\xi[\lambda(\bar{x}, \omega)]$ is a subgradient of $\mathbb{F}[V(x, \omega)]$ at $\bar{x}$.

The meat of the matter. Consider the following (re-stated) proposition from Philpott, de Matos, and Finardi 2013:

Proposition

Suppose for each $\omega \in \Omega$, that $\lambda(\bar{x}, \omega)$ is a subgradient of $V(x, \omega)$ at $\bar{x}$. Then, given $\mathbb{F}[V(\bar{x}, \omega)] = \mathbb{E}_{\xi}[V(\bar{x}, \omega)]$, $\mathbb{E}_{\xi}[\lambda(\bar{x}, \omega)]$ is a subgradient of $\mathbb{F}[V(x, \omega)]$ at $\bar{x}$.

Northwestern ENGINEERING

The meat of the matter. Consider the following (re-stated) proposition from Philpott, de Matos, and Finardi 2013:

Proposition

Suppose for each $\omega \in \Omega$, that $\lambda(\bar{x}, \omega)$ is a subgradient of $V(x, \omega)$ at $\bar{x}$. Then, given $\mathbb{F}[V(\bar{x}, \omega)] = \mathbb{E}_\xi[V(\bar{x}, \omega)]$, $\mathbb{E}_\xi[\lambda(\bar{x}, \omega)]$ is a subgradient of $\mathbb{F}[V(x, \omega)]$ at $\bar{x}$.

Northwestern ENGINEERING

So what is this saying?

To obtain a cut for $\mathbb{F}_{\omega_{t+1}\in\Omega_{t+1}}[V_{t+1}(x_{t+1},\omega_{t+1})]$

- We can go and solve the $t+1$ stage problems to obtain an objective value $\bar{\theta}_{\omega_{t+1}}$ and a dual vector $\lambda_{\omega_{t+1}}$ for each realization of ω_{t+1}.
- Normally, we take the expectation of these to get the cut

$$\theta_{t+1} \geq \mathbb{E}[\bar{\theta}_{\omega_{t+1}}] + \mathbb{E}[\lambda_{\omega_{t+1}}]^\top(x_{t+1} - \bar{x}_{t+1})$$

- Instead, we compute ξ according to $\bar{\theta}_{\omega_{t+1}}$ and then take the risk-adjusted expectation to get the cut

$$\theta_{t+1} \geq \mathbb{E}_{\xi}[\bar{\theta}_{\omega_{t+1}}] + \mathbb{E}_{\xi}[\lambda_{\omega_{t+1}}]^\top(x_{t+1} - \bar{x}_{t+1})$$

Outline

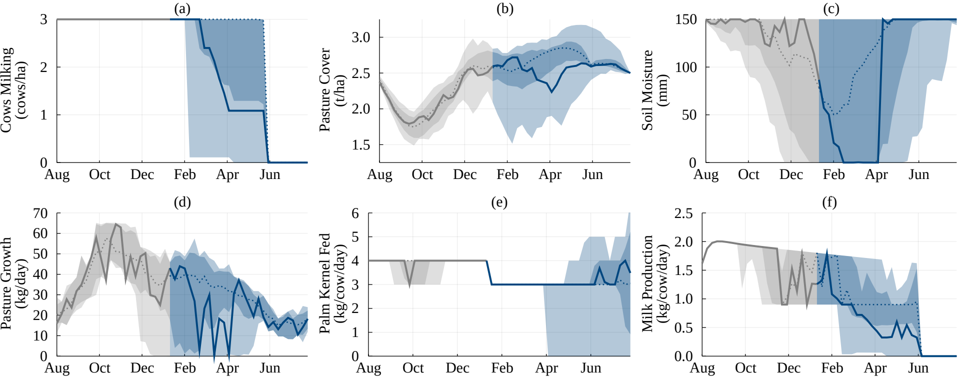
Some background

Risk

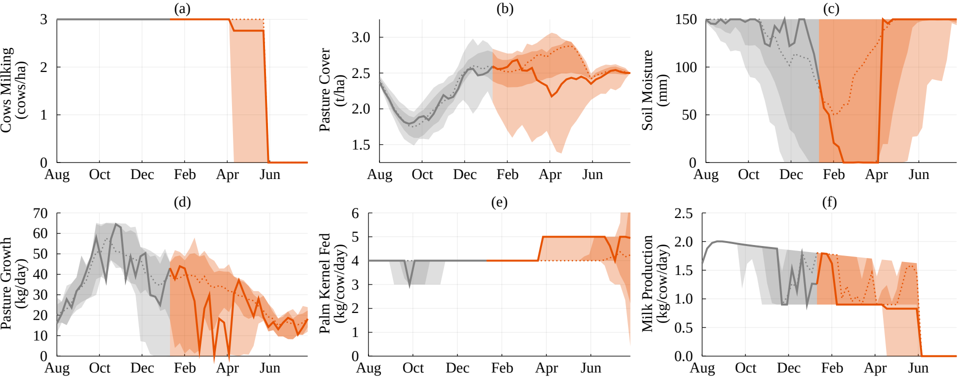
SDDP.jl tutorial

POWDER

Northwestern ENGINEERING



POWDER



(a)
Cows Milking (cows/ha)
3
2
1
0
Aug Oct Dec Feb Apr Jun
(b)
Pasture Cover (t/ha)
3.0
2.5
2.0
1.5
Aug Oct Dec Feb Apr Jun
(c)
Soil Moisture (mm)
150
100
50
0
Aug Oct Dec Feb Apr Jun
(d)
Pasture Growth (kg/day)
70
60
50
40
30
20
10
0
Aug Oct Dec Feb Apr Jun
(e)
Palm Kernel Fed (kg/cow/day)
6
5
4
3
2
1
0
Aug Oct Dec Feb Apr Jun
(f)
Milk Production (kg/cow/day)
2.5
2.0
1.5
1.0
0.5
0.0
Aug Oct Dec Feb Apr Jun

But Wait! Danger ahead!

Northwestern ENGINEERING

Do these nested risk measures make sense?

Remember how the *end-of-horizon* risk measure made the most sense:

$$\mathbb{F}[X_1, X_2, X_3] = \mathbb{F}[X_1 + X_2 + X_3]$$

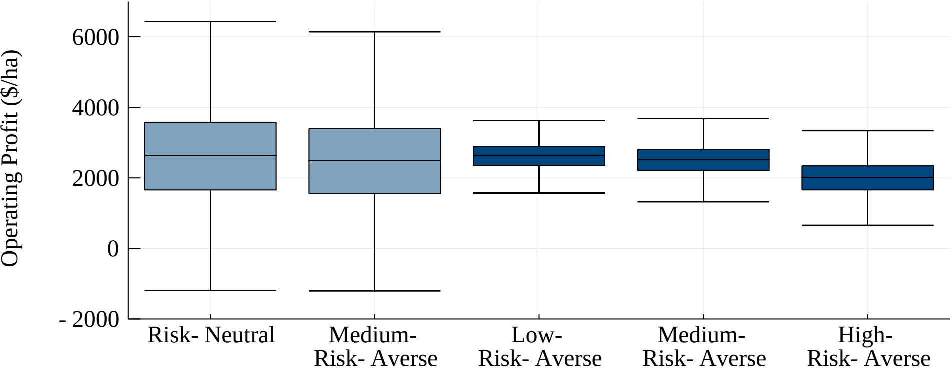
But we actually used the *nested* risk measure:

$$\mathbb{F}[X_1, X_2, X_3] = \mathbb{F}[X_1 + \mathbb{F}[X_2 + \mathbb{F}[X_3 \mid X_1, X_2] \mid X_1]]$$

What is the interpretation of a nested risk measure? This can lead to perverse, counter-intuitive results!

POWDER

Northwestern ENGINEERING



References I

- Philippe Artzner et al. "Coherent Measures of Risk". In: *Mathematical Finance* 9.3 (1999), pp. 203–228.
- Regan Baucke, Anthony Downward, and Golbon Zakeri. "A Deterministic Algorithm for Solving Multistage Stochastic Minimax Dynamic Programmes". In: *Optimization Online* (2018). URL: http://www.optimization-online.org/DB_FILE/2018/02/6449.pdf.
- Tito Homem-de-Mello and Bernardo K. Pagnoncelli. "Risk Aversion in Multistage Stochastic Programming: A Modeling and Algorithmic Perspective". In: *European Journal of Operational Research* 249.1 (2016), pp. 188–199.
- Georg Ch. Pflug and Alois Pichler. "Time-Inconsistent Multistage Stochastic Programs: Martingale Bounds". In: *European Journal of Operational Research* 249.1 (2016), pp. 155–163.

References II

Northwestern ENGINEERING

Andy Philpott, Vitor de Matos, and Erlon Finardi. "On Solving Multistage Stochastic Programs with Coherent Risk Measures". In: *Operations Research* 61.4 (2013), pp. 957–970.

Tyrrell R. Rockafellar and Stanislav P. Uryasev. "Conditional Value-at-Risk for General Loss Distributions". In: *Journal of Banking and Finance* 26 (2002), pp. 1443–1471.

Andrzej Ruszczyński. "Risk-Averse Dynamic Programming for Markov Decision Processes". In: *Mathematical Programming* 125.2 (2010), pp. 235–261.