SDDP.jl

Stochastic dual dynamic programming in Julia

Oscar Dowson

Department of Industrial Enginnering and Management Sciences
Northwestern

October 22, 2018

Outline

Some background

Risk

SDDP.jl tutorial

Outline

Northwestern ENGINEERING

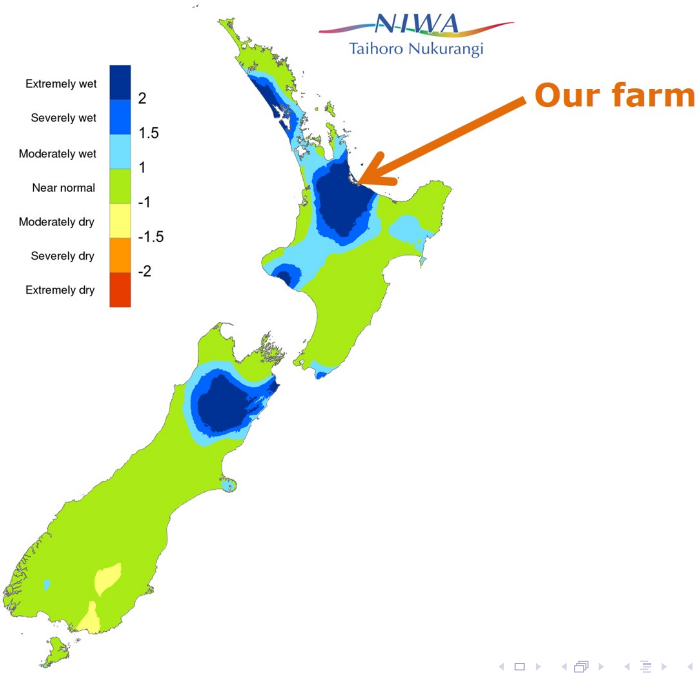
Some background

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?

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

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

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?

- 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?

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.
- 30 years of experience said: we can't have a bad Summer, Autumn, Winter, AND Spring

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.
- 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

Some background

Risk

SDDP.jl tutorial

Northwestern ENGINEERING

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

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

Northwestern ENGINEERING

Definition

Risk measure 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.

Northwestern ENGINEERING

Two simple examples that you already know

Given $\Omega = \{1, 2, 3\}$ with equal probability.

- Expectation: $\mathbb{F}[Z] = \mathbb{E}[Z] = (1 + 2 + 3)/3 = 2$
- Worst-cast: $\mathbb{F}[Z] = \max\{1, 2, 3\} = 3$

Definition

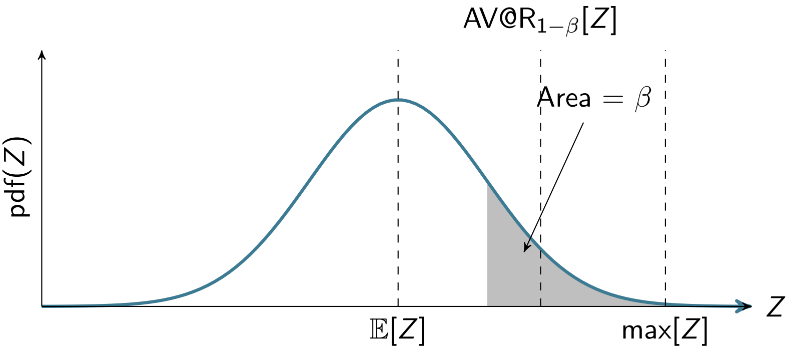
The *Average Value-at-Risk* at the β quantile ($\text{AV@R}_{1-\beta}$) is:

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

where $(x)_+ = \max\{0, x\}$.

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

Northwestern ENGINEERING



The primal formulation of AV@R

$$\min_{\xi,x} \quad \xi + \frac{1}{\beta}\frac{1}{K}\sum_{k=1}^{K} x_k \tag{1}$$

$$\text{s.t.} \quad x_k \geq z_k - \xi \tag{2}$$

$$x_k \geq 0 \tag{3}$$

AV@R Examples

Given $\Omega = \{1, 2, 3\}$ with equal probability.

- $\text{AV@R}_{1-{}^{2}/_{3}}$: $\mathbb{F}[Z] = 0.5 \times 3 + 0.5 \times 2 = 2.5$
- $\text{AV@R}_{1-0.5}$: $\mathbb{F}[Z] = 2/3 \times 3 + 1/3 \times 2 = 2\,{}^{2}/_{3}$

Northwestern ENGINEERING

Definition

A *coherent* risk measure is a risk measure $\mathbb{F}$ that satisfies the axioms of Artzner (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:

$$\mathbb{F}[Z] = \sup_{\xi \in \mathfrak{A}} \mathbb{E}_{\xi}[Z] = \sup_{\xi \in \mathfrak{A}} \sum_{k=1}^{K} \xi_k z_k, \tag{4}$$

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

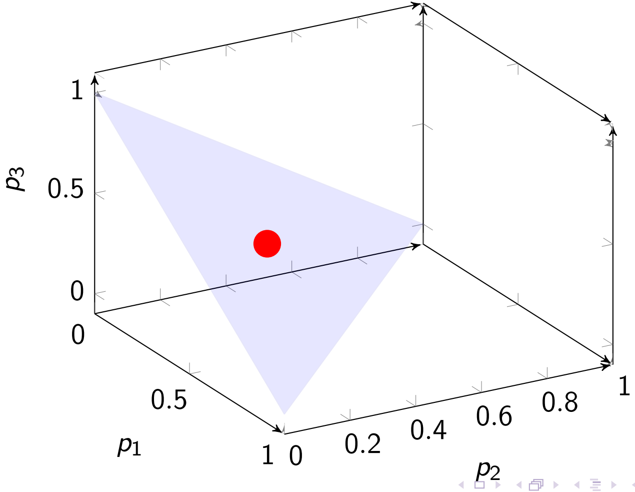
$$\mathfrak{P} = \left\{ \xi \in \mathbb{R}^K \ : \ \sum_{k=1}^{K} \xi_k = 1, \ \xi \geq 0 \right\}.$$

z	$\mathbb{E}$	max	$\text{AV@R}_{0.5}$
1	$1/3$	0	0
2	$1/3$	0	$1/3$
3	$1/3$	1	$2/3$
$\mathbb{F}[Z]$	2	3	2.67

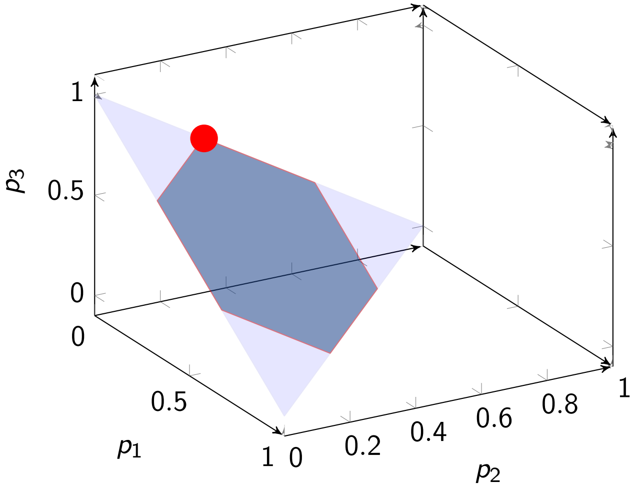
Table: Caption

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.

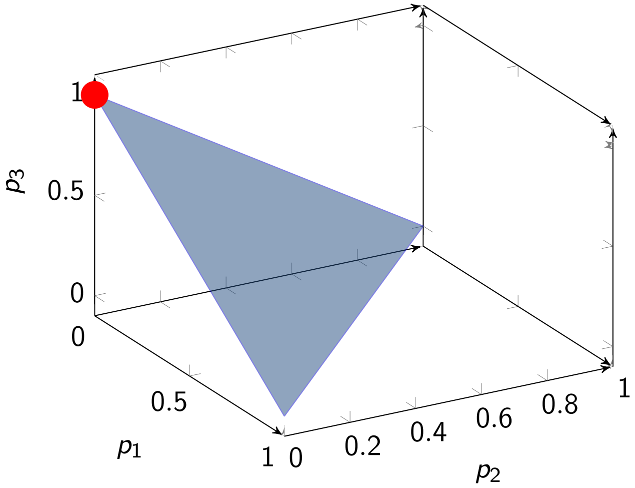


If $\mathfrak{A} = \left\{ \xi \in \mathfrak{P} \mid \xi_k \leq \frac{p_k}{\beta},\ k = 1, 2, \ldots, K \right\}$, then the risk measure $\mathbb{F}$ is equivalent to $\mathsf{AV@R}_{1-\beta}$.



Risk

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



Northwestern ENGINEERING

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

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

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

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

Definition

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 $\mathbb{F}[Z] = \mathbb{E}_\xi[Z]$.

Definition

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

So what is this saying?

- We can go and solve the next 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 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

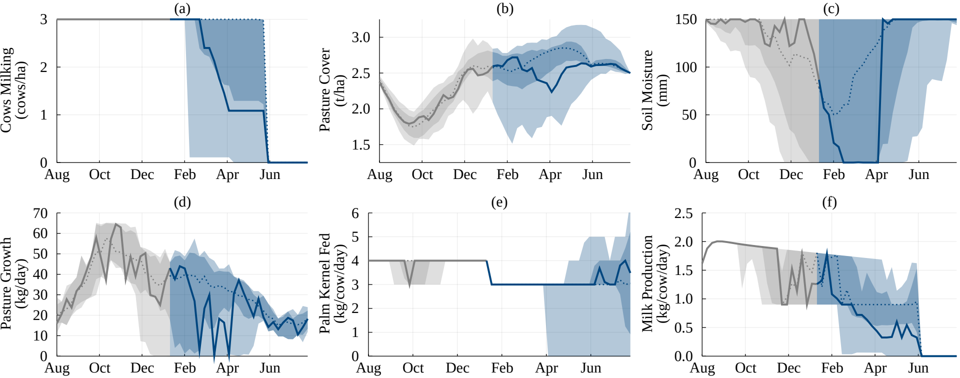
Some background

Risk

SDDP.jl tutorial

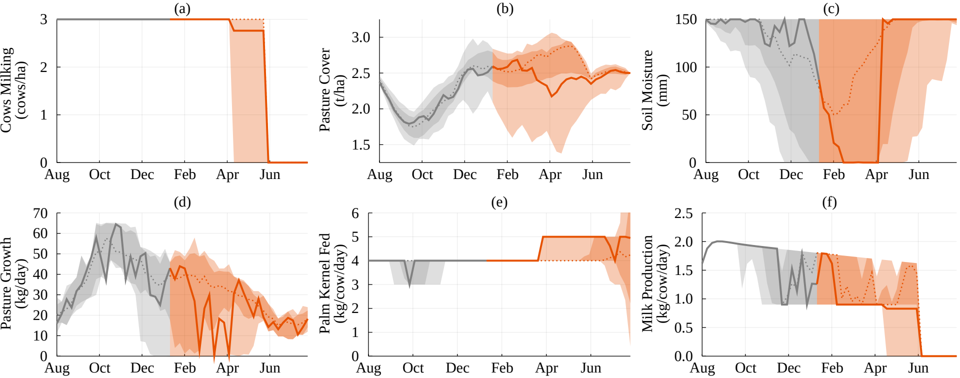
POWDER

Northwestern ENGINEERING



POWDER

Northwestern ENGINEERING



But Wait! Danger ahead!

Northwestern ENGINEERING

Do these nested risk measures make sense?

Often, we want to calculate the *end-of-horizon* value:

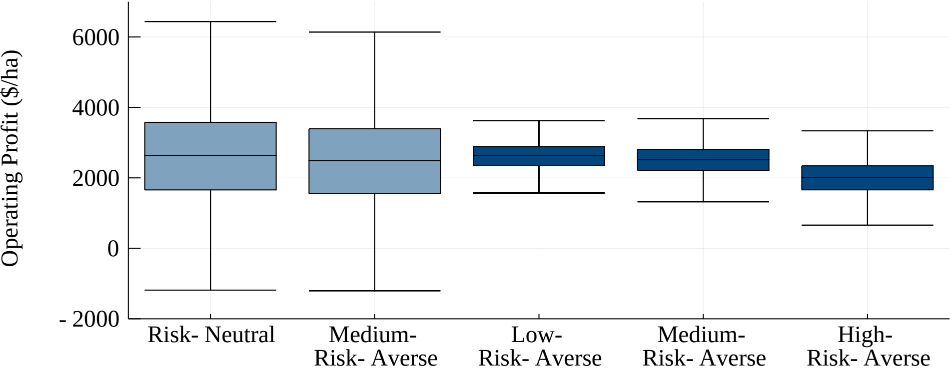
$$\mathbb{F}[X_1, X_2, X_3] = \mathbb{F}[X_1 + X_2 + X_3]$$

But we are actually calculating the *nested* value:

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

This can lead to perverse, counter-intuitive results!

POWDER



Operating Profit ($/ha)
6000
4000
2000
0
- 2000
Risk- Neutral
Medium-
Risk- Averse
Low-
Risk- Averse
Medium-
Risk- Averse
High-
Risk- Averse