Introduction to Risk Measures for Stochastic Optimization

Risk is in the eyes of the beholder

Michel De Lara
Cermics, École des Ponts ParisTech
France

École des Ponts ParisTech

April 14, 2022

Outline of the presentation

Minimizing costs with risk

Worst-case, expectation criteria and variations

Axiomatics for risk measures

What makes optimization under uncertainty specific

- Optimization set is made of random variables
- Criterion generally derives from aggregating uncertainties, mostly by a mathematical expectation, or by a risk measure
- Constraints
  - generally include measurability constraints, like the nonanticipativity constraints,
  - and may also include probability constraints, or robust constraints

Here are the ingredients for a general abstract optimization problem under uncertainty

- A set $\mathbb{U}$ of decisions
- A set Ω of scenarios
- An optimization set $\mathbb{V} \subset \mathbb{U}^{\Omega}$ containing random variables $\mathrm{V} : \Omega \to \mathbb{U}$
- A criterion $J : \mathbb{V} \to \mathbb{R} \cup \{+\infty\}$
- Constraints of the form $\mathrm{V} \in \mathbb{V}^{ad} \subset \mathbb{V}$

$$\inf_{\mathrm{V} \in \mathbb{V}^{ad}} J(\mathrm{V})$$

Here is the most common framework for robust and stochastic optimization

- A set $\mathbb{U}$ of decisions
- A set Ω of scenarios, or states of Nature, possibly equipped with a σ-algebra
- An optimization set $\mathbb{V} \subset \mathbb{U}^{\Omega}$ containing random variables $\mathsf{V} : \Omega \to \mathbb{U}$
- A risk measure $\mathbb{F} : \mathbb{V} \to \mathbb{R} \cup \{+\infty\}$
- A function $j : \mathbb{U} \times \Omega \to \mathbb{R} \cup \{+\infty\}$ (say, the "deterministic" criterion)
- Constraints of the form $\mathsf{V} \in \mathbb{V}^{ad} \subset \mathbb{V}$

$$\inf_{\mathsf{V} \in \mathbb{V}^{ad}} J(\mathsf{V}) = \mathbb{F}\big[j(\mathsf{V}(\cdot), \cdot)\big]$$

where the notation means that the risk measure $\mathbb{F}$ has for argument the random variable

$$j(\mathsf{V}(\cdot), \cdot) : \Omega \to \mathbb{R} \cup \{+\infty\} \ , \ \ \omega \mapsto j(\mathsf{V}(\omega), \omega)$$

Examples of classes of robust and stochastic optimization problems

$$\mathbb{U}^{ad} \subset \mathbb{U}$$

- Stochastic optimization "à la" gradient stochastique
  - The risk measure $\mathbb{F}$ is a mathematical expectation $\mathbb{E}$
  - Measurability constraints make that the random variables $\mathsf{V} \in \mathbb{V}^{ad}$ are constant, that is, are deterministic decision variables

$$\inf_{u \in \mathbb{U}^{ad}} \mathbb{E}_{\mathbb{P}}\big[j(u, \cdot)\big]$$

- Robust optimization
  - The risk measure $\mathbb{F}$ is the fear operator/worst case $\sup_{\omega \in \overline{\Omega}}$, where $\overline{\Omega} \subset \Omega$
  - Measurability constraints make that the random variables $\mathsf{V} \in \mathbb{V}^{ad}$ are constant, that is, are deterministic decision variables

$$\inf_{u \in \mathbb{U}^{ad}} \sup_{\omega \in \overline{\Omega}} j(u, \omega)$$

Multistage optimization examples

- A set $\mathbb{U}$
  $\mathbb{U} = \mathbb{U}_0 \times \mathbb{U}_1$ in two stage programming
- A set Ω of scenarios
  Ω finite, $\Omega = \mathbb{N} \times \mathbb{W}^{\mathbb{N}}$ for discrete time stochastic processes
- An optimization set $\mathbb{V} \subset \mathbb{U}^{\Omega}$ containing random variables $\mathsf{V} : \Omega \to \mathbb{U}$
- A risk measure $\mathbb{F} : \mathbb{V} \to \mathbb{R} \cup \{+\infty\}$
  most often a mathematical expectation $\mathbb{E}$,
  but can be $\sup_{\omega \in \overline{\Omega}}$ in the robust case, with $\overline{\Omega} \subset \Omega$
- A function $j : \mathbb{U} \times \Omega \to \mathbb{R} \cup \{+\infty\}$
- Constraints of the form $\mathsf{V} \in \mathbb{V}^{ad} \subset \mathbb{V}$
  - Measurability constraints,
    like the nonanticipativity constraints
  - Pointwise constraints,
    like probability constraints and robust constraints

Most common constraints in robust and stochastic optimization problems

- Measurability constraints

$$\mathsf{V} \in \text{linear subspace of } \mathbb{U}^{\Omega}$$

like the nonanticipativity constraints

  - $\mathsf{V} = (\mathsf{V}_0, \mathsf{V}_1)$, V_0 is $\mathcal{F}_0$-measurable, V_1 is $\mathcal{F}_1$-measurable, with $\{\emptyset, \Omega\} = \mathcal{F}_0 \subset \mathcal{F}_1 = \mathcal{F}$
  - $\mathsf{V} = \{\mathsf{V}_t\}_{t=0,\dots,T}$, $\sigma(\mathsf{V}_t) \subset \mathcal{F}_t$ for $t = 0, \dots, T$ with filtration $\{\mathcal{F}_t\}_{t=0,\dots,T}$

- Pointwise constraints, with $\mathbb{U}^{ad} : \Omega \rightrightarrows \mathbb{U}$
  - probability constraints

$$\mathbb{P}\big(\mathsf{V} \in \mathbb{U}^{ad}\big) \geq 1 - \epsilon$$

  - robust constraints

$$\mathsf{V}(\omega) \in \mathbb{U}^{ad}(\omega) \; , \quad \forall \omega \in \overline{\Omega} \subset \Omega$$

Savage's minimal regret criterion... "Had I known"

The regret performs an additive normalization of the function $j : \mathbb{U} \times \Omega \to \mathbb{R} \cup \{+\infty\}$

Regret

For $u \in \mathbb{U}$ and $\omega \in \Omega$, the regret is

$$r(u, \omega) = j(u, \omega) - \min_{u' \in \mathbb{U}} j(u', \omega)$$

Then, take any risk measure $\mathbb{F}$ and solve

$$\min_{\mathsf{V} \in \mathbb{V}^{ad}} \mathbb{F}\big[r(\mathsf{V}, \cdot)\big] = \min_{\mathsf{V} \in \mathbb{V}^{ad}} \mathbb{F}\big[j(\mathsf{V}(\omega), \omega) - \min_{u \in \mathbb{U}} j(u, \omega)\big]$$

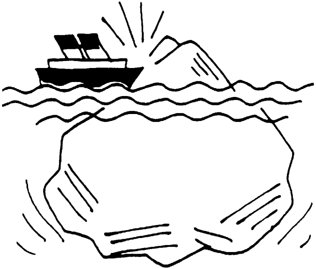
so that one can have minimal worse regret, minimal expected regret, etc.

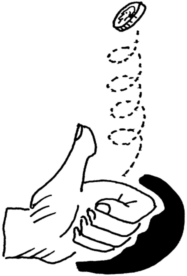
Where have we gone till now? And what comes next

- We have layed out the ingredients to set up problems of optimization under uncertainty
- We have overviewed the constraints
- Now, we detail how we can build an array of criteria from an array of risk measures
- Robust and stochastic optimization will appear as special cases

Outline of the presentation

For you, Nature is rather random or hostile?





Outline of the presentation

The $(+, \times)$ algebra of probability theory

Probability space

- The set Ω is equipped with a σ-field $\mathcal{F}$ ($(\Omega, \mathcal{F})$ measurable space), and the elements of $\mathcal{F} \subset 2^{\Omega}$ are called events
- One speaks of a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ when the measurable space $(\Omega, \mathcal{F})$ is equipped with a probability $\mathbb{P}$ (supposed, when needed and for the sake of simplicity, to have a density p w.r.t. a reference measure, thus covering the finite case)
- The probability $\mathbb{P} : \mathcal{F} \to [0, 1]$ has the properties
  - normalization
    $$\mathbb{P}(\emptyset) = 0 \ , \ \ \mathbb{P}(\Omega) = 1$$
  - additivity
    $$\mathbb{P}(\bigcup_{n \in \mathcal{N}} A_n) = \sum_{n \in \mathcal{N}} \mathbb{P}(A_n)$$
    for any countable set $\mathcal{N}$, $A_n \in \mathcal{F}$ for all $n \in \mathcal{N}$, such that $m \neq n \implies A_m \cap A_m = \emptyset$

Expected value

- A random variable is a measurable mapping $X : (\Omega, \mathcal{F}) \to (\mathbb{X}, \mathcal{X})$ (between measurable spaces)
- The expected value of a nonnegative random variable $X : \Omega \to \mathbb{R}_+ \cup \{+\infty\}$ is

$$\mathbb{E}[X] = \int_\Omega X(\omega)\,\mathrm{d}\mathbb{P}(\omega) \quad \Big(\int_\Omega X(\omega)p(\omega)\,\mathrm{d}(\omega)\Big) \quad \Big(\sum_{\omega\in\Omega} \mathbb{P}\{\omega\}X(\omega)\Big)$$

- The notation $\mathbb{E}$ (or $\mathbb{E}_{\mathbb{P}}$ or $\mathbb{E}^{\mathbb{P}}$) refers to the mathematical expectation (operator) over Ω under probability $\mathbb{P}$, extended to integrable real-valued random variables
- The expectation operator $\mathbb{E}$ enjoys linearity in the $(+, \times)$ algebra

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$

- The random variables X, Y are independent under $\mathbb{P}$ when their joint distribution $\mathbb{P}_{(X,Y)}$ can be decomposed as a product

$$\mathbb{P}_{(X,Y)} = \mathbb{P}_X \otimes \mathbb{P}_Y$$

The $(\max, +)$ algebra of decision/robust/plausability theory

Decision space, cost measure, plausability are the robust counterparts of probability space

- The set Ω is equipped with a σ-field $\mathcal{F}$ ($(\Omega, \mathcal{F})$ measurable space)
- One speaks of a decision space $(\Omega, \mathcal{F}, \mathbb{K})$ when the measurable space $(\Omega, \mathcal{F})$ is equipped with a cost measure $\mathbb{K}$ (supposed, when needed, to have a density κ, thus covering the finite case)
- The cost measure (plausibility) $\mathbb{K} : \mathcal{F} \to [-\infty, 0]$ has the properties
  - normalization
    $$\mathbb{K}(\emptyset) = -\infty \ , \ \ \mathbb{K}(\Omega) = 0$$
  - $(\max, +)$ "additivity"
    $$\mathbb{K}(\bigcup_{n \in \mathcal{N}} A_n) = \sup_{n \in \mathcal{N}} \mathbb{K}(A_n)$$

    for any countable set $\mathcal{N}$, $A_n \in \mathcal{F}$ for all $n \in \mathcal{N}$, such that $m \neq n \implies A_m \cap A_m = \emptyset$

Cost density, plausibility function

- The function $\kappa : \Omega \to [-\infty, 0]$ is a cost density of the cost measure $\mathbb{K}$ if

$$\mathbb{K}(A) = \sup_{\omega \in A} \kappa(\omega) \ , \ \ \forall A \in \mathcal{F}$$

- A function $\kappa : \Omega \to \mathbb{R} \cup \{-\infty\} = [-\infty, 0]$, such that $\sup_{\omega \in \Omega} \kappa(\omega) = 0$, is a cost density, also called plausibility function

The fear operator [Bernhard, 1995]

The Moreau lower addition extends the usual addition with

$$(+\infty) \dot{+} (-\infty) = (-\infty) \dot{+} (+\infty) = -\infty$$

- A decision variable is a mapping $(\Omega, \mathcal{F}) \to (\mathbb{T}, \mathcal{T})$ (with codomain a topological space)
- The feared value of a function $\psi : \Omega \to [-\infty, +\infty]$ (real-valued decision variable) is defined by

$$\mathbb{F}(\psi) = \sup_{\omega \in \Omega} \left[\psi(\omega) \dot{+} \kappa(\omega) \right]$$

- The fear operator $\mathbb{F}$ enjoys linearity in the $(\max, +)$ algebra

$$\mathbb{F}\big(\max\{\psi, \phi\}\big) = \max\{\mathbb{F}(\psi), \mathbb{F}(\phi)\}$$

- Independence

$$\mathbb{K}_{(\psi,\phi)} = \mathbb{K}_{\psi} + \mathbb{K}_{\phi}$$

More or less implausible events

- For any subset $\Omega'' \subset \Omega$, we have that

$$\mathbb{K}(\emptyset) = -\infty \leq \mathbb{K}(\Omega'') \leq \mathbb{K}(\Omega) = 0$$

- The higher (closest to zero from below), the more plausible, whereas totally implausible outcomes in Ω'' are such that $\mathbb{K}(\Omega'') = -\infty$
- With any subset $\Omega' \subset \Omega$, we associate the characteristic function $\delta_{\Omega'} : \Omega \to \{0, +\infty\}$

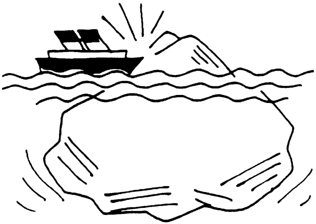
$$\delta_{\Omega'}(\omega) = \begin{cases} 0 & \text{if } \omega \in \Omega' \\ +\infty & \text{if } \omega \notin \Omega' \end{cases}$$

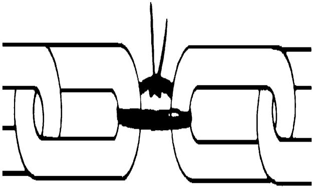
- The cost measure associated with the uniform density $-\delta_{\Omega'}$ satisfies

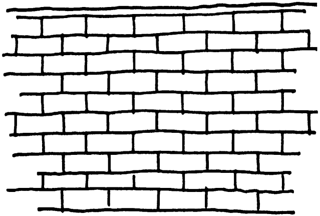
$$\forall \Omega'' \subset \Omega \,, \;\; \mathbb{K}(\Omega \setminus \Omega'') = \begin{cases} -\infty & \text{if } \Omega'' \subset \Omega' \quad (\Omega \setminus \Omega'' \text{ implausible }) \\ 0 & \text{if } \Omega'' \cap \Omega' \neq \emptyset \end{cases}$$

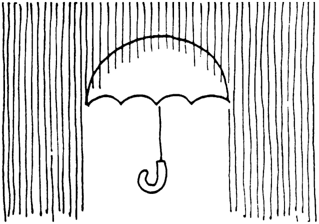
Outline of the presentation

In the robust or pessimistic approach,
Nature is supposed to be malevolent,
and the DM aims at protection against all odds









In the robust or pessimistic approach, Nature is supposed to be malevolent

- In the robust approach, the DM considers the worst cost

$$\underbrace{\sup_{\omega \in \Omega} j(\mathsf{V}(\omega), \omega)}_{\text{worst cost}}$$

- Nature is supposed to be malevolent,
  and specifically selects the worst outcome:
  the DM plays after Nature has played, and minimizes the worst cost

$$\min_{\mathsf{V} \in \mathbb{V}^{ad}} \sup_{\omega \in \Omega} j(\mathsf{V}(\omega), \omega)$$

- Robust, pessimistic, worst case, minimax, maximin (for payoffs)

The robust approach can be softened with plausibility weighting

- Let $\kappa : \Omega \to \mathbb{R} \cup \{-\infty\} = [-\infty, 0]$, such that $\sup_{\omega \in \Omega} \kappa(\omega) = 0$, be a plausibility function
- The higher (closest to zero from below), the more plausible, whereas totally implausible outcomes ω are those for which $\kappa(\omega) = -\infty$
- Nature is malevolent, and specifically selects the worst outcome, but weighs it according to the plausibility function κ
- The DM plays after Nature has played, and solves

$$\min_{\mathsf{V} \in \mathbb{V}^{ad}} \left[\sup_{\omega \in \Omega} \left(j(\mathsf{V}(\omega), \omega) + \underbrace{\kappa(\omega)}_{\text{plausibility}} \right) \right]$$

In the optimistic approach, Nature is supposed to benevolent

Future. That period of time in which our affairs prosper, our friends are true and our happiness is assured.

Ambrose Bierce

- Instead of minimizing the worst cost as in a robust approach, the optimistic focuses on the most favorable cost

$$\underbrace{\inf_{\omega \in \Omega} j(\mathsf{V}(\omega), \omega)}_{\text{lowest cost}}$$

- Nature is supposed to benevolent, and specifically selects the best outcome: the DM plays after Nature has played, and solves

$$\min_{\mathsf{V} \in \mathbb{V}^{ad}} \inf_{\omega \in \Omega} j(\mathsf{V}(\omega), \omega)$$

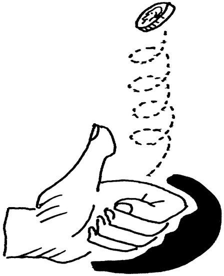
The Hurwicz criterion reflects an intermediate attitude between optimistic and pessimistic approaches

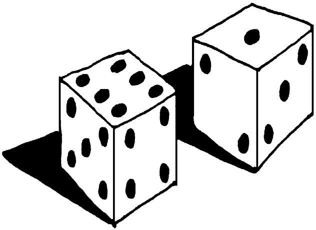
A proportion $\alpha \in [0,1]$ graduates the level of prudence

$$\min_{\mathsf{V} \in \mathbb{V}^{ad}} \left\{ \alpha \overbrace{\sup_{\omega \in \Omega} j(\mathsf{V}(\omega), \omega)}^{\text{pessimistic}} + (1-\alpha) \underbrace{\inf_{\omega \in \Omega} j(\mathsf{V}(\omega), \omega)}_{\text{optimistic}} \right\}$$

Outline of the presentation

In the stochastic or expected approach, Nature is supposed to play stochastically





In the stochastic or expected approach, Nature is supposed to play stochastically

- $(\Omega, \mathcal{F})$ equipped with a probability $\mathbb{P}$
- The expected cost is

$$\overbrace{\mathbb{E}^{\mathbb{P}}\Big[j\big(\mathsf{V}(\cdot),\cdot\big)\Big]}^{\text{mean cost}} = \int_{\Omega} j(\mathsf{V}(\omega),\omega)\, \mathrm{d}\mathbb{P}(\omega)$$

- Nature is supposed to play stochastically, according to distribution $\mathbb{P}$: the DM plays after Nature has played, and solves

$$\min_{\mathsf{V}\in\mathbb{V}^{ad}} \mathbb{E}^{\mathbb{P}}\Big[j\big(\mathsf{V}(\cdot),\cdot\big)\Big]$$

The expected disutility approach distorts costs before taking the expectation

- We consider a disutility function L to assess the disutility of the costs (for instance minus a CARA exponential utility function)
- The expected disutility is

$$\underbrace{\mathbb{E}\Big[L\Big(j\big(\mathsf{V}(\cdot),\cdot\big)\Big)\Big]}_{\text{expected disutility}} = \int_{\Omega} L\Big(j\big(\mathsf{V}(\omega),\omega\big)\Big)\,\mathrm{d}\mathbb{P}(\omega)$$

- The expected disutility minimizer solves

$$\min_{\mathsf{V}\in\mathbb{V}^{ad}} \mathbb{E}\Big[L\Big(j\big(\mathsf{V}(\cdot),\cdot\big)\Big)\Big]$$

The ambiguity or multi-prior approach combines robust and expected criterion

- Set $\Delta(\Omega)$ of probabilities on Ω
- Different probabilities $\mathbb{P}$ (often termed as beliefs or priors) belonging to a subset $\mathcal{P} \subset \Delta(\Omega)$ of admissible probabilities on Ω
- The multi-prior approach combines robust and expected criterion by taking the worst beliefs in terms of expected cost

$$\min_{\mathsf{V} \in \mathbb{V}^{ad}} \underbrace{\sup_{\mathbb{P} \in \mathcal{P}} \overbrace{\mathbb{E}^{\mathbb{P}}\Big[j\big(\mathsf{V}(\cdot), \cdot\big)\Big]}^{\text{mean cost}}}_{\text{pessimistic over probabilities}}$$

Extension of the ambiguity or multi-prior approach

- Subset $\mathcal{P} \subset \Delta(\Omega)$ of admissible probabilities on Ω
- To each probability $\mathbb{P}$ is attached a plausibility $\kappa(\mathbb{P}) \in [-\infty, 0]$

$$\min_{\mathsf{V} \in \mathbb{V}^{ad}} \underbrace{\sup_{\mathbb{P} \in \mathcal{P}} \Big(\overbrace{\mathbb{E}^{\mathbb{P}}\Big[j\big(\mathsf{V}(\cdot), \cdot\big)\Big]}^{\text{mean cost}} + \overbrace{\kappa(\mathbb{P})}^{\text{plausibility}} \Big)}_{\text{pessimistic over probabilities}}$$

What is the special case where $\mathcal{P}$ is
either the set of Dirac measures or the whole set of probabilities?

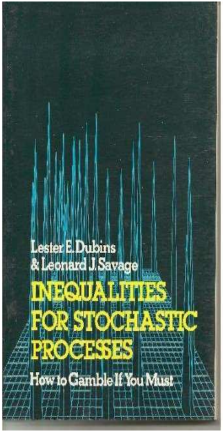
Convex risk measures cover a wide range of risk criteria (more on convex risk measures later)

- Plausibility function $\kappa : \Delta(\Omega) \to [-\infty, 0]$

$$\underbrace{\sup_{\mathbb{P} \in \Delta(\Omega)} \Big(\overbrace{\mathbb{E}^{\mathbb{P}}\big[j(\mathsf{V}(\cdot), \cdot)\big]}^{\text{mean cost}} + \overbrace{\kappa(\mathbb{P})}^{\text{plausibility}} \Big)}_{\text{pessimistic over probabilities}}$$

- Expected: $\kappa(\bar{\mathbb{P}}) = 0$, and $\kappa(\mathbb{P}) = -\infty$ for any $\mathbb{P} \neq \bar{\mathbb{P}}$
- Robust: $\kappa(\mathbb{P}) = 0$ for any $\mathbb{P} \in \Delta(\Omega)$
- Multi-prior: $\kappa(\mathbb{P}) = 0$ for any $\mathbb{P} \in \mathcal{P} \subset \Delta(\Omega)$, and $\kappa(\mathbb{P}) = -\infty$ for any $\mathbb{P} \notin \mathcal{P}$

Non convex risk measures can lead to non diversification



How to gamble if you must, L.E. Dubbins and L.J. Savage, 1965

Imagine yourself at a casino with $1,000. For some reason, you desperately need $10,000 by morning; anything less is worth nothing for your purpose.

The only thing possible is to gamble away your last cent, if need be, in an attempt to reach the target sum of $10,000.

- The question is how to play, not whether. What ought you do? How should you play?
  - Diversify, by playing 1 $ at a time?
  - Play boldly and concentrate, by playing 1,000 $ only one time?
- What is your decision criterion?

Outline of the presentation

Minimizing costs with risk

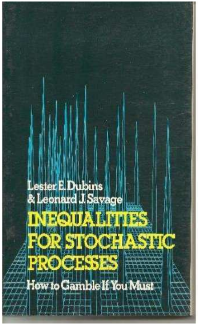
Worst-case, expectation criteria and variations

Axiomatics for risk measures

Outline of the presentation

Here are some guiding principles for any risk formalization

- The essence of risk is asymmetry between bad and good odds
- If a prospect makes better than another in every state of nature, it should be less risky
- Diversification should not increase risk
(this point should be discussed)



Recalls on measurable spaces, σ-fields, probability

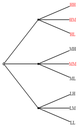
- Finite set Ω, random variables are mappings $X : \Omega \to \mathbb{R}$
- The law of large numbers requires an infinite (product) probability space able to sustain an infinite number of distinct random variables
- It is not possible to define a probability $\mathbb{P}$ on all subsets of an infinite sample space Ω
- This is why we restrict events to elements of a σ-field $\mathcal{F}$
- $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space
- Random variables are measurable mappings $X : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, namely $X^{-1}(\mathcal{B}_{\mathbb{R}}) \subset \mathcal{F}$
- With a random variable X we associate the image probability $\mathbb{P}_X = \mathbb{P} \circ X^{-1}$ on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$

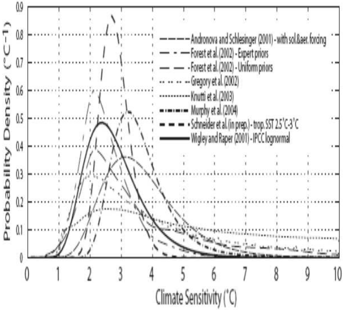
"There are things we do not know we don't know"

"Reports that say that something hasn't happened are always interesting to me, because as we know, there are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns – the ones we don't know we don't know. And if one looks throughout the history of our country and other free countries, it is the latter category that tend to be the difficult ones."

Donald Rumsfeld, former United States Secretary of Defense. From Department of Defense news briefing, February 12, 2002.

Uncertainty, risk, ambiguity





- *Uncertainty*: set Ω
- $\omega \in \Omega$: *outcome*, *scenario*, *state of nature*, etc.
- *Information*: σ-field $\mathcal{F}$ of *events* $F \subset \mathcal{F}$
- *Risk*: $(\Omega, \mathcal{F})$ carries a *probability* $\mathbb{P}$
- *Ambiguity*: family $\mathcal{Q}$ of probabilities $\mathbb{P} \in \mathcal{Q}$

Variance and standart deviation fail the test as risk measures: they are measures of dispersion and variability

$$\text{var}(X) = \mathbb{E}\left[\left(X - \mathbb{E}[X] \right)^2 \right]$$

- The variance is not measured in the same units than X, since $\text{var}(\theta X) = \theta^2 \text{var}(X)$
- This can be corrected by using the standart deviation $\sigma(X) = \sqrt{\text{var}(X)}$
- The variance is not monotonous: $X \geq Y \not\Rightarrow \text{var}(X) \leq \text{var}(Y)$ (take $Y = 0$ and any $X \geq 0$ which is not constant)
- The variance weighs symmetrically what is above and what is below the mean, whereas the essence of risk is asymmetry between bad and good odds: $\to \mathbb{E}\left[\left(X - \mathbb{E}[X] \right)_- \right]$

Outline of the presentation

Going from home to the airport with a safety margin

- When you go from home to the airport, you consider possible transportation delay (road accident, bus delay), represented by a (stochastic) transport time X
- You take a safety margin, and add some (deterministic) extra time $\rho(X)$
- This extra time $\rho(X)$ depends
  - on the randomness (X) that affects transportation
  - on how you perceive (ρ) the importance of being "just in time"
- This deterministic extra time is an example of (gauge) risk measure

$$\underbrace{\overbrace{\rho(X)}^{\text{deterministic extra time}} + \overbrace{X}^{\text{stochastic transportation time}}}_{\text{acceptable stochastic time from home to airport}} \in \mathcal{A}$$

Risk measures as capital requirement

A measure of risk associates to each cost X

- the minimum extra capital $\rho(X)$, a deterministic number,
- required to make it "acceptable" to a regulator
- that is, such that when you substract $\rho(X)$ from the cost X, the shifted cost $X - \rho(X)$ becomes acceptable

The lower $\rho(X)$, the better (the less risk)

Mathematical ingredients

- We consider a measurable space $(\Omega, \mathcal{F})$
- We denote by $\mathbb{L}(\Omega, \mathcal{F}, \mathbb{R})$ the set of random variables (measurable functions $Z : \Omega \to \mathbb{R}$)
- When Ω is finite (scenarios), the set of random variables coincides with $\mathbb{R}^{\Omega}$, the set of functions from Ω to $\mathbb{R}$

Interpreting the mathematical expectation as a gauge risk measure

- Define the following set of acceptable random variables

$$\mathcal{A} = \{Z \in \mathbb{L}(\Omega, \mathcal{F}, \mathbb{R}) \mid \mathbb{E}[Z] \leq 0\}$$

  When the random variable Z is interpreted as a cost,
  a cost with negative mean is acceptable

- The mathematical expectation $\mathbb{E}[X]$ of a random cost X is the smallest amount x you can substract to X to make $X - x$ acceptable

$$\mathbb{E}[X] = \inf\{x \in \mathbb{R} \mid X - x \in \mathcal{A}\} = \inf\{x \in \mathbb{R} \mid \mathbb{E}[X - x] \leq 0\}$$

Outline of the presentation

Axiomatics of risk measures

A risk measure is a mapping $\rho : \mathbb{L}(\Omega, \mathcal{F}, \mathbb{R}) \to \mathbb{R}$ (or $\mathbb{R} \cup \{-\infty\}$)

Risk measures

A risk measure ρ is

- (T) invariant by translation if $\rho(X + x) = \rho(X) + x$, for all $x \in \mathbb{R}$
- (M) monotonous whenever $X \geq Y \Rightarrow \rho(X) \geq \rho(Y)$
- (C) convex if $\rho(\theta X + (1-\theta)Y) \leq \theta\rho(X) + (1-\theta)\rho(Y)$
- (PH) positively homogeneous if $\rho(\theta X) = \theta\rho(X)$ when $\theta > 0$
- (S) subadditive if $\rho(X + Y) \leq \rho(X) + \rho(Y)$

One says that ρ is a monetary risk measure if it is monotonous (M) and invariant by translation (T)

Acceptance set of a translation invariant risk measure

$\rho \mapsto \mathcal{A}_\rho$

With any translation invariant risk measure ρ, we associate the subset $\mathcal{A}_\rho$ of random variables, called acceptance set of ρ

$$\mathcal{A}_\rho = \{\mathsf{X} \in \mathbb{L}(\Omega, \mathcal{F}, \mathbb{R}) \mid \rho(\mathsf{X}) \leq 0\}$$

Denoting by 1 the constant random variable having value 1, we have that

$$\mathcal{A}_\rho + \mathbb{R}_+ 1 \subset \mathcal{A}_\rho$$

since ρ is translation invariant

Translation invariant risk measure associated with a given acceptance set

$\mathcal{A} \mapsto \rho_{\mathcal{A}}$

Letting $\mathcal{A} \subset \mathbb{L}(\Omega, \mathcal{F}, \mathbb{R})$ be such that $\mathcal{A} + \mathbb{R}_+ 1 \subset \mathcal{A}$, we define

$$\rho_{\mathcal{A}}(X) = \inf\{x \in \mathbb{R} \mid X - x \in \mathcal{A}\}$$

which is a translation invariant risk measure

Axiomatics of acceptance sets

Acceptance sets

An acceptance set $\mathcal{A}$ is

(M) monotonous whenever $X \geq Y$ and $Y \in \mathcal{A} \Rightarrow X \in \mathcal{A}$

(C) convex if $\theta X + (1-\theta)Y \in \mathcal{A}$ whenever $X \in \mathcal{A}$ and $Y \in \mathcal{A}$

(PH) a cone if $\theta\mathcal{A} \subset \mathcal{A}$ for any $\theta > 0$, that is,
if $\theta X \in \mathcal{A}$ whenever $X \in \mathcal{A}$ and $\theta > 0$

(S) stable by addition if $X + Y \in \mathcal{A}$
whenever $X \in \mathcal{A}$ and $Y \in \mathcal{A}$

(SS) star-shaped if $\theta\mathcal{A} \subset \mathcal{A}$ for any $\theta \in]0,1[$

Worst case risk measure

The worst case risk measure is

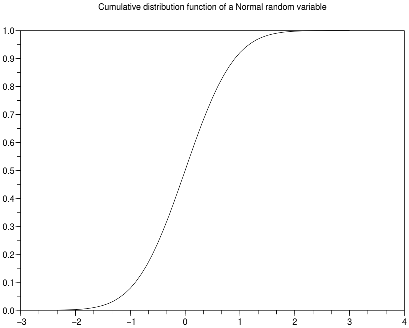
$$\rho_{\max}(X) = \sup_{\omega \in \Omega} X(\omega)$$

with acceptance set

$$\mathcal{A} = \{Z \in \mathbb{L}(\Omega, \mathcal{F}, \mathbb{R}) \mid Z \leq 0\}$$

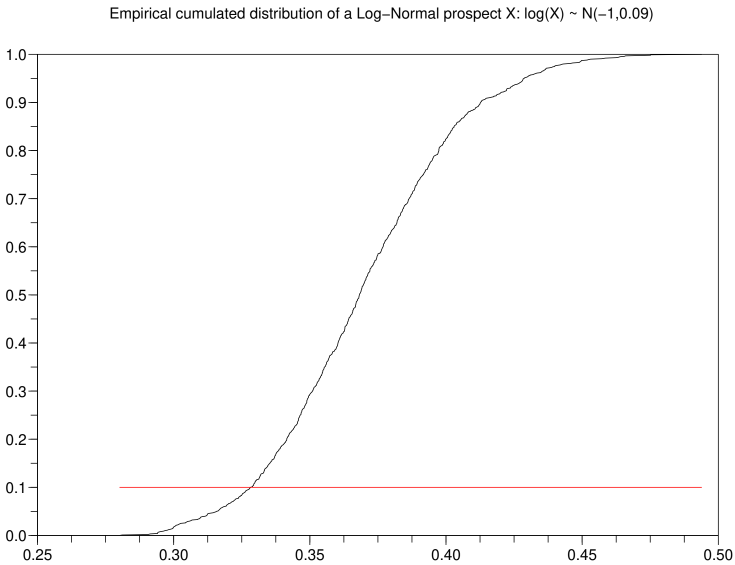
Outline of the presentation

Cumulative distribution function



- $(\Omega, \mathcal{F})$ equipped with a probability $\mathbb{P}$
- $\mathsf{X} : \Omega \to \mathbb{R}$ real random variable
- $\psi_{\mathsf{X}}(x) = \mathbb{P}(\mathsf{X} \leq x)$
- $x \mapsto \psi_{\mathsf{X}}(x)$ is increasing and *right-continuous*

Continuous case example



Quantiles in the continuous case

- Let $\lambda \in]0,1[$, that plays the role of a risk level
- In the case of a random variable X with positive density, the c.d.f. ψ_X is stricly increasing, and there is a unique λ-quantile given by

$$Q_\lambda^-(X) = Q_\lambda^+(X) = \psi_X^{-1}(\lambda)$$

- Some quantiles of a Normal random variable are

x	-3.33	-2.57	-2.33	-1.96	-1.64
$\psi_{\mathcal{N}(0,1)}(x)$	0.0005	0.005	0.01	0.025	0.05

Table: Quantiles of a Normal random variable $\mathcal{N}(0,1)$

Discrete case example

X	$\mathbb{P}$
12	0.2
10	0.3
6	0.4
-3	0.1

for $x \leq 12$	$\psi_X(x) = 1$
for $10 \leq x < 12$	$\psi_X(x) = 0.1 + 0.4 + 0.3$
for $6 \leq x < 10$	$\psi_X(x) = 0.1 + 0.4$
for $-3 \leq x < 6$	$\psi_X(x) = 0.1$
for $x < -3$	$\psi_X(x) = 0$

Cumulative distribution function left-continuous inverse

In the general case, we introduce

- left-continuous inverse of ψ_X

$$\psi_X^{-1}(\lambda) = \inf\{x \in \mathbb{R} \mid \psi_X(x) \geq \lambda\}$$

- characterized by the equivalence

$$\psi_X^{-1}(\lambda) \leq x \iff \lambda \leq \psi_X(x)$$

Quantiles

- Let $\lambda \in]0, 1[$, that plays the role of a risk level
- A λ-quantile of the real random variable X is any real number q such that

$$\mathbb{P}(X \leq q) \geq \underbrace{\lambda}_{\text{risk level}} \geq \mathbb{P}(X < q)$$

- The set of λ-quantiles is an interval $[Q_\lambda^-(X), Q_\lambda^+(X)]$ where

$$\begin{aligned} Q_\lambda^-(X) &= \inf\{q \mid \mathbb{P}(X \leq q) \geq \lambda\} \\ Q_\lambda^+(X) &= \inf\{q \mid \mathbb{P}(X \leq q) > \lambda\} \end{aligned}$$

- $Q_\lambda^-(X) = \inf\{q \mid \psi_X(q) \geq \lambda\} = \psi_X^{-1}(\lambda)$

Quantiles in the discrete case

- Let $X \sim \mathcal{B}(1/2; 1)$ follow a Bernoulli distribution

$$\mathbb{P}(X < 0) = 0 < \mathbb{P}(X \leq 0) = 1/2 = \mathbb{P}(X < 1) < \mathbb{P}(X \leq 1) = 1$$

- The quantiles are

	$0 < \lambda < 1/2$	$\lambda = 1/2$	$1/2 < \lambda < 1$
$Q^{-}_{\lambda}(X)$	0	0	1
$Q^{+}_{\lambda}(X)$	0	1	1

Table: Quantiles of a Bernoulli $\mathcal{B}(1/2; 1)$ random variable

Quantiles in the discrete case

- Discrete random variable X
  - having values $x_1 < x_2 < \cdots < x_n$
  - with probabilities $\mathbb{P}(X = x_i) = p_i > 0$, $i = 1, \ldots, n$
- Letting $q_1 = p_1,\ q_2 = p_1 + p_2,\ q_{n-1} = p_1 + \cdots + p_{n-1}$
- the quantiles are given by

	$0 < \lambda < q_1$	$\lambda = q_1$	$p_1 < \lambda < q_2$	$\lambda = q_2$	$\cdots$	$\lambda = q_{n-1}$	$q_{n-1} < \lambda < 1$
$Q_\lambda^-(X)$	x_1	x_1	x_2	x_2	$\cdots$	x_{n-1}	x_n
$Q_\lambda^+(X)$	x_1	x_2	x_2	x_3	$\cdots$	x_n	x_n

Table: Quantiles of a discrete random variable

The Value at Risk (quantile)

Let $\lambda \in]0,1[$, that plays the role of a risk level

Value at Risk

The Value at Risk of the cost X at level $\lambda \in]0,1[$ is

$$VaR_{\lambda}(X) = \inf\{x \in \mathbb{R} \mid \mathbb{P}(X > x) < \lambda\}$$

with acceptance set

$$\mathcal{A} = \{Z \in \mathbb{L}(\Omega, \mathcal{F}, \mathbb{R}) \mid \mathbb{P}(Z \geq 0) < \lambda\}$$

Value at Risk

- Intuitively, saying that the $VaR_{5\%}$ of a portfolio is 100 means that the loss will be more than 100 with probability at most 5%
- $VaR_{5\%}$ is the maximum loss in the 95% of the cases
- However, $VaR_{5\%}$ does not inform on the size of the loss
- If $\mathbb{P}(X > 0) = 0$, then $VaR_\lambda(X) \geq 0$, meaning that money could be added to the cost, and it would still be acceptable

Value at Risk and quantiles

$$\begin{aligned} VaR_\lambda(X) &= \inf\{x \in \mathbb{R} \mid \mathbb{P}(X > x) \leq \lambda\} \\ &= Q^-_{1-\lambda}(X) = -Q^+_\lambda(-X) \end{aligned}$$

where

$$\begin{aligned} Q^-_\lambda(X) &= \inf\{x \mid \mathbb{P}(X \leq x) \geq \lambda\} \\ &= \inf\{x \in \mathbb{R} \mid \psi_X(x) \geq \lambda\} = \psi_X^{-1}(\lambda) \end{aligned}$$

$$\begin{aligned} Q^+_\lambda(X) &= \inf\{x \mid \mathbb{P}(X \leq x) > \lambda\} \\ &= \sup\{x \mid \mathbb{P}(X < x) \leq \lambda\} \end{aligned}$$

Properties of the Value at Risk

The Value at Risk of a cost X is measured in the same units than X, and is

- invariant by translation

$$VaR_\lambda(X + x) = VaR_\lambda(X) + x \ , \ \ \forall x \in \mathbb{R}$$

- monotonous

$$X \geq Y \Rightarrow VaR_\lambda(X) \geq VaR_\lambda(Y)$$

- positively homogeneous

$$VaR_\lambda(\theta X) = \theta VaR_\lambda(X) \ , \ \ \forall \theta > 0$$

But. . . diversification does not always decrease risk!

Value at Risk and diversification

Beware: here X and Y are minus costs!

ω	X	$\mathbb{P}$
1	-100	4%
2	0	4%
3	0	4%
⋮	⋮	⋮
25	0	4%

ω	Y	$\mathbb{P}$
1	0	4%
2	-100	4%
3	0	4%
⋮	⋮	⋮
25	0	4%

ω	$0.5X + 0.5Y$	$\mathbb{P}$
1	-50	4%
2	-50	4%
3	0	4%
⋮	⋮	⋮
25	0	4%

- The minimum m to be added to X in such a way that $\mathbb{P}(X + m < 0) \leq 5\%$ is $m = 0$ since $\mathbb{P}(X - \epsilon < 0) = 100\% > 5\%$ for all $\epsilon > 0$.
- Hence $VaR_{5\%}(X) = VaR_{5\%}(Y) = 0$.
- And...

$$VaR_{5\%}(X) = VaR_{5\%}(Y) = 0 < 50 = VaR_{5\%}(0.5X + 0.5Y)$$

Outline of the presentation

The Tail Value at Risk (superquantile)

Let $\lambda \in]0,1[$, that plays the role of a risk level

Tail Value at Risk

The Tail Value at Risk of the cost X at level $\lambda \in]0,1[$ is

$$TVaR_{\lambda}(X) = \frac{1}{1-\lambda} \int_{\lambda}^{1} VaR_{\lambda'}(X) d\lambda'$$

Properties of the Tail Value at Risk

The Tail Value at Risk of a cost X is measured in the same units than X, and is

- invariant by translation

$$TVaR_\lambda(\mathrm{X}+x) = TVaR_\lambda(\mathrm{X}) + x \ , \quad \forall x \in \mathbb{R}$$

- monotonous

$$\mathrm{X} \geq \mathrm{Y} \Rightarrow TVaR_\lambda(\mathrm{X}) \geq TVaR_\lambda(\mathrm{Y})$$

- positively homogeneous

$$TVaR_\lambda(\theta \mathrm{X}) = \theta \, TVaR_\lambda(\mathrm{X}) \ , \quad \forall \theta > 0$$

- convex, hence favors diversification :-)

[Rockafellar and Uryasev, 2000]

$$TVaR_\lambda[X] = \inf_{s\in\mathbb{R}} \left\{ \frac{\mathbb{E}[(X-s)^+]}{1-\lambda} + s \right\} \ , \quad \lambda \in [0,1[$$

Limit cases

$$TVaR_0[X] = \mathbb{E}[X]$$
$$TVaR_1[X] = \lim_{\lambda\to 1} TVaR_\lambda[X] = \sup_{\omega\in\Omega} X(\omega)$$

More on the Tail Value at Risk

- The Average Value at Risk or Tail Value at Risk

$$TVaR_\lambda(X) = \frac{1}{\lambda}\int_0^\lambda VaR_{\lambda'}(X)d\lambda'$$

- The Worst Conditional Expectation

$$\sup\{\mathbb{E}[X \mid A] \, , \; A \in \mathcal{F} \, , \; \mathbb{P}(A) < \lambda\}$$

  are the worst costs conditioned over events
  of probability less than the risk level $\lambda \in]0,1[$

- If $\mathbb{P}\left\{X \leq Q^{-}_{1-\lambda}(X)\right\} = \lambda$,

$$TVaR_\lambda(X) = \mathbb{E}[X \mid \overbrace{X \geq VaR_\lambda(X)}^{\text{costs greater than VaR}}]$$

  is the average of costs greater than the Value at Risk

Outline of the presentation

Affine risk measures

The cost average under probability $\mathbb{Q}$ is

$$\rho(\mathsf{X}) = \mathbb{E}_{\mathbb{Q}}(\mathsf{X})$$

whereas the shifted cost average under probability $\mathbb{Q}$ is

$$\rho(\mathsf{X}) = \mathbb{E}_{\mathbb{Q}}(\mathsf{X}) - \gamma$$

Convex risk measures

Given

- a subset $\mathcal{Q}$ of probabilities on Ω, representing different priors about the randomness
- a function $\gamma : \mathcal{Q} \to \mathbb{R}$, with $\sup_{\mathbb{Q}\in\mathcal{Q}} \gamma(\mathbb{Q}) < +\infty$, representing cost shifts

we define

$$\rho(X) = \sup_{\mathbb{Q}\in\mathcal{Q}} \Big(\mathbb{E}_{\mathbb{Q}}(X) - \gamma(\mathbb{Q}) \Big)$$

which expresses

- first, an average of the cost X over different outcomes ponderations $\mathbb{Q} \in \mathcal{Q}$, each being *penalized* by $\gamma(\mathbb{Q})$
- second, a conservative attitude by taking the largest with the sup operation over priors

Bibliography

P. Bernhard. A separation theorem for expected value and feared value discrete time control. Technical report, INRIA, Projet Miaou, Sophia Antipolis, Décembre 1995.

H. Föllmer and A. Schied. *Stochastic Finance. An Introduction in Discrete Time*. Walter de Gruyter, Berlin, 2002.

R. T. Rockafellar and S. Uryasev. Optimization of Conditional Value-at-Risk. *Journal of Risk*, 2:21–41, 2000.