Uncertainty Quantification

Risk measures and their role in UQ

Alois Pichler

UQ and risk measures
Dortmund, DE
March 14, 2018

Mathematik!
TU Chemnitz

Outline

Outline

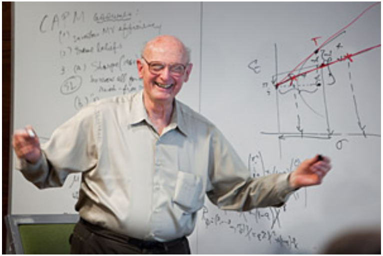


Figure: Harry Markowitz (1927) explains the CAPM and the mean variance plot. Nobel Memorial Prize in Economic Sciences (1990)

Markowitz considers the problem

$$\begin{aligned} \text{minimize (in } x \in \mathbb{R}^J) \ & \operatorname{var}\left(\xi^\top x\right) \\ \text{subject to } & \mathbb{E}\xi^\top x \geq \mu, \\ & \mathbb{1}^\top x \leq 1\text{k Euro}, \\ & (x \geq 0) \end{aligned}$$

Some statistics

Are $\mathcal{R}(\cdot) = \mathbb{E}(\cdot)$ and $\mathcal{R}(\cdot) = \operatorname{var}(\cdot)$ appropriate?

Capital Asset Pricing Model
CAPM

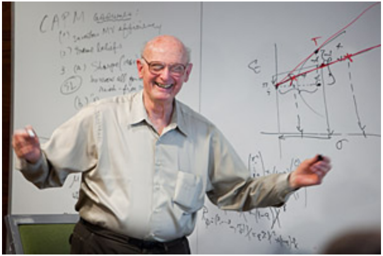


Figure: Harry Markowitz (1927) explains the CAPM and the mean variance plot. Nobel Memorial Prize in Economic Sciences (1990)

Markowitz considers the problem

$$
\begin{aligned}
\text{minimize (in } x \in \mathbb{R}^J) \ \operatorname{var}\left(\xi^\top x\right) \\
\text{subject to } \mathbb{E}\,\xi^\top x \geq \mu, \\
\mathbb{1}^\top x \leq 1\text{k Euro}, \\
(x \geq 0)
\end{aligned}
$$

Some statistics

Are $\mathcal{R}(\cdot) = \mathbb{E}(\cdot)$ and $\mathcal{R}(\cdot) = \operatorname{var}(\cdot)$ appropriate?

Example

$$\mathcal{R}(\cdot) = \mathbb{E}(\cdot).$$

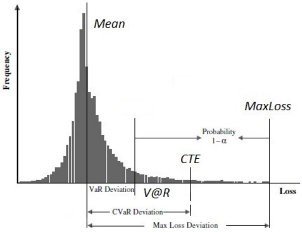
Proposition (Axioms, cf. Artzner et al. (1999))

Coherent measure of risk,

$$\mathcal{R}\colon \mathcal{Y} \to \mathbb{R} \cup \{\pm\infty\}.$$

1. *Monotonicity:* $X \le Y$ *a.e., then* $\mathcal{R}(X) \le \mathcal{R}(Y)$,
2. *Translation equivariance:* $\mathcal{R}(Y + y) \le \mathcal{R}(Y) + y$ *for* $Y \in \mathcal{Y}$ *and* $y \in \mathbb{R}$,
3. *Convexity:* $\mathcal{R}(X + Y) \le \mathcal{R}(X) + \mathcal{R}(Y)$ *for* $X, Y \in \mathcal{Y}$,
4. *Positive homogeneity:* $\mathcal{R}(\lambda Y) = \lambda \cdot \mathcal{R}(Y)$ *for* $\lambda > 0$.

Outline



Definition (V@R, AV@R, CTE)

The Value-at-Risk at risk level $\alpha \in (0,1)$ is

$$\mathrm{V@R}_\alpha(Y) = F_Y^{-1}(\alpha) = \inf\{y\colon P(Y \le y) \ge \alpha\},$$

the Average Value-at-Risk is

$$\begin{aligned}\mathrm{AV@R}_\alpha(Y) &= \frac{1}{1-\alpha}\int_\alpha^1 F_Y^{-1}(p)\mathrm{d}p,\\ &= \min_{t\in\mathbb{R}} t + \frac{1}{1-\alpha}\mathbb{E}(Y-t)_+.\end{aligned}$$

Fact

The name Conditional Value-at-Risk is suggested by the formula $\mathrm{AV@R}_\alpha(Y) = \mathbb{E}[Y \mid Y \ge \mathrm{V@R}_\alpha(Y)]$.

Axioms

$$\mathcal{R}(Y_1) \le \mathcal{R}(Y_2) \text{ if } Y_1 \le Y_2,$$

$$\mathcal{R}(Y+c) = \mathcal{R}(Y) + c$$

$$\mathcal{R}(Y_1+Y_2) \le \mathcal{R}(Y_1) + \mathcal{R}(Y_2)$$

$$\mathcal{R}(\lambda Y) = \lambda \mathcal{R}(Y)$$

Theorem (Fenchel–Moreau, cf. Rockafellar (1970))

It holds that

$$\mathcal{R}(Y) = \sup_{Z \in \mathcal{Y}^*} \mathbb{E}\, YZ - \mathcal{R}^*(Z),$$

where

$$\mathcal{R}^*(Z) = \sup_{Y \in \mathcal{Y}} \mathbb{E}\, YZ - \mathcal{R}(Y)$$

is the convex conjugate (dual) function.

Law invariant risk functional
Kusuoka representation

The distortion risk measure (spectral risk measure),

$$\mathcal{R}_\sigma(Y) := \int_0^1 \sigma(u) F_Y^{-1}(u)\,\mathrm{d}u$$
$$= \int_0^1 \mathsf{AV@R}_\alpha(Y)\,\mu_\sigma(\mathrm{d}\alpha).$$

楠岡 成雄 (くすおか しげお)
KUSUOKA, Shigeo

Theorem (From Fenchel Moreau–Theorem (cf. Kusuoka's representation))

If $\mathcal{R}$ is version independent (law-invariant), then, for some class S,

$$\mathcal{R}(Y) = \sup_{\sigma\in S} \mathcal{R}_\sigma(Y) = \sup_{\sigma\in\mathcal{S}} \int_0^1 F_Y^{-1}(\alpha)\sigma(\alpha)\,\mathrm{d}\alpha$$

Outline

A laaarge deviation example
Heavy tails

Chernoff bound

$$P(Y \geq \eta) = P\left(e^{tY} \geq e^{t\eta}\right) \leq \frac{\mathbb{E}\, e^{tY}}{e^{t\eta}},$$

or equivalently, for $t > 0$,

$$\eta + \frac{1}{t} \log P(Y \geq \eta) \leq \frac{1}{t} \log \mathbb{E}\, e^{tY}.$$

It is a further attempt to consider

$$\eta + \frac{1}{t} \log \frac{P(Y \geq \eta)}{1-\alpha} \leq \frac{1}{t} \log \frac{1}{1-\alpha} \mathbb{E}\, e^{tY}.$$

Particularly, if we choose $\eta := \mathrm{V@R}_\alpha(Y)$, then

$$\mathrm{V@R}_\alpha(Y) \leq \inf_{t>0} \frac{1}{t} \log \frac{1}{1-\alpha} \mathbb{E}\, e^{tY} =: \mathrm{EV@R}_\alpha(Y).$$

A laaarge deviation example
Heavy tails

Chernoff bound

$$P(Y \geq \eta) = P\left(e^{tY} \geq e^{t\eta}\right) \leq \frac{\mathbb{E}\, e^{tY}}{e^{t\eta}},$$

or equivalently, for $t > 0$,

$$\eta + \frac{1}{t} \log P(Y \geq \eta) \leq \frac{1}{t} \log \mathbb{E}\, e^{tY}.$$

It is a further attempt to consider

$$\eta + \frac{1}{t} \log \frac{P(Y \geq \eta)}{1-\alpha} \leq \frac{1}{t} \log \frac{1}{1-\alpha} \mathbb{E}\, e^{tY}.$$

Particularly, if we choose $\eta := \mathrm{V@R}_\alpha(Y)$, then

$$\mathrm{V@R}_\alpha(Y) \leq \inf_{t>0} \frac{1}{t} \log \frac{1}{1-\alpha} \mathbb{E}\, e^{tY} =: \mathrm{EV@R}_\alpha(Y).$$

A laaarge deviation example

Heavy tails

Chernoff bound

$$P(Y \geq \eta) = P\left(e^{tY} \geq e^{t\eta}\right) \leq \frac{\mathbb{E}\, e^{tY}}{e^{t\eta}},$$

or equivalently, for $t > 0$,

$$\eta + \frac{1}{t} \log P(Y \geq \eta) \leq \frac{1}{t} \log \mathbb{E}\, e^{tY}.$$

It is a further attempt to consider

$$\eta + \frac{1}{t} \log \frac{P(Y \geq \eta)}{1-\alpha} \leq \frac{1}{t} \log \frac{1}{1-\alpha} \mathbb{E}\, e^{tY}.$$

Particularly, if we choose $\eta := \mathrm{V@R}_\alpha(Y)$, then

$$\mathrm{V@R}_\alpha(Y) \leq \inf_{t>0} \frac{1}{t} \log \frac{1}{1-\alpha} \mathbb{E}\, e^{tY} =: \mathrm{EV@R}_\alpha(Y).$$

Outline

Theorem (Fenchel–Moreau, Donsker–Varadhan variational formula cf. Ahmadi, P. (2017))

$$\mathrm{EV@R}_\alpha(Y) = \inf_{t>0} \frac{1}{t} \log \frac{1}{1-\alpha} \mathbb{E}\, e^{tY}$$
$$= \sup\left\{ \mathbb{E}\, YZ \colon\ \mathbb{E}\, Z = 1,\ Z \geq 0 \text{ and } \underbrace{\mathbb{E}\, Z \log Z}_{H(Z)} \leq \log \frac{1}{1-\alpha} \right\}.$$

and conversely,

$$\underbrace{\mathbb{E}\, Z \log Z}_{\text{entropy } H(Z)} = \sup\left\{ \mathbb{E}\, YZ - \log \mathbb{E}\, e^{Y} \colon\ Y \in \mathcal{Y} \right\}.$$

First law of thermodynamics:
$E = const$
Second law of thermodynamics:
$\Delta S \geq 0$

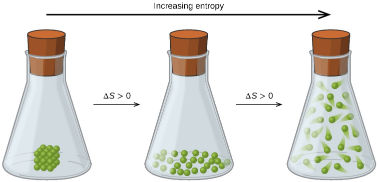


Figure: Entropy

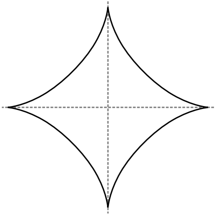
Proposition

Choose the density $\mathrm{d}Q = Z\mathrm{d}P$, *then*

$$
\begin{aligned}
\mathsf{EV@R}_\alpha(Y) &= \inf_{t>0} \frac{1}{t} \log \frac{1}{1-\alpha}\, \mathbb{E}\, e^{tY} \\
&= \sup\left\{ \mathbb{E}\, YZ \colon\ \mathbb{E}\, Z = 1,\ Z \geq 0 \textit{ and } \underbrace{\mathbb{E}\, Z \log Z}_{H(Z)} \leq \log \frac{1}{1-\alpha} \right\} \\
&= \sup\left\{ \mathbb{E}_Q\, Y \colon\ D_{KL}(Q\|P) \leq \log \frac{1}{1-\alpha} \right\}.
\end{aligned}
$$

Definition (Entropies)

1. Shannon Entropy: $H_1(Z) := \mathbb{E}\, Z \log Z$,
2. Rényi Entropy ($q \neq 1$):
   $H_q(Z) := \frac{1}{q-1} \log \mathbb{E}\, Z^q = \frac{q}{q-1} \log \|Z\|_q$.



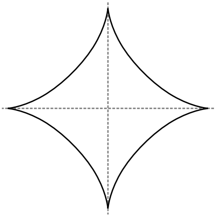
Proposition (Risk measure based on Rényi entropy, Dentcheva et al. (2010))

$$\mathrm{EV@R}_\alpha^p(Y) := \sup \left\{ \mathbb{E}\, YZ \;\middle|\; \begin{array}{l} Z \geq 0, \mathbb{E}\, Z = 1 \text{ and} \\ H_q(Z) \leq \log \frac{1}{1-\alpha} \end{array} \right\}$$

$$= \inf_{t \in \mathbb{R}} \left\{ t + \left(\frac{1}{1-\alpha} \right)^{1/p} \cdot \|(Y-t)_+\|_p \right\}.$$

Definition (Entropies)

1. Shannon Entropy: $H_1(Z) := \mathbb{E}\, Z \log Z$,
2. Rényi Entropy ($q \neq 1$):
   $H_q(Z) := \frac{1}{q-1} \log \mathbb{E}\, Z^q = \frac{q}{q-1} \log \|Z\|_q$.



Proposition (Risk measure based on Rényi entropy, Dentcheva et al. (2010))

$$\begin{aligned}\mathrm{EV@R}_\alpha^p(Y) :&= \sup\left\{ \mathbb{E}\, YZ \;\middle|\; \begin{array}{l} Z \geq 0, \mathbb{E}\, Z = 1 \text{ and} \\ H_q(Z) \leq \log \frac{1}{1-\alpha} \end{array} \right\} \\ &= \inf_{t\in\mathbb{R}} \left\{ t + \left(\frac{1}{1-\alpha} \right)^{1/p} \cdot \|(Y-t)_+\|_p \right\}.\end{aligned}$$

Outline

Definition (Wasserstein/ Kantorovich Distance)

The Kantorovich distance (also *Wasserstein distance*) *of order r* on a Polish space (Ξ, d)

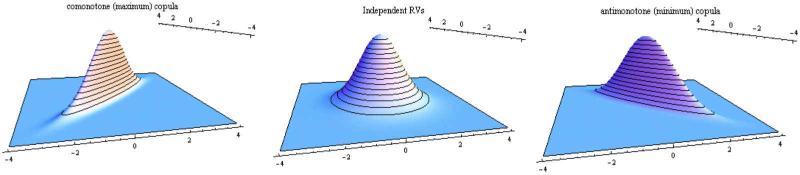
$$w_r(P, Q; d) := \left(\inf_\pi \iint_{\Xi \times \tilde{\Xi}} d(x, y)^r \, \pi(\mathrm{d}x, \mathrm{d}y) \right)^{\frac{1}{r}},$$

where the infimum is taken over all (bivariate) probability measures π on $\Xi \times \Xi$ which have respective marginals, that is

$$\pi(A \times \Xi) = P(A) \text{ and } \pi(\Xi \times B) = Q(B)$$

for all measurable sets $A \subseteq \Xi$ and $B \subseteq \Xi$.

Wasserstein Distance
Kantorovich, resp.



Why Wasserstein?

- Rachev lists 76 metrics for measures in his book...
- the empirical measures $\frac{1}{n}\sum_{i=1}^{n}\delta_{\xi_i}$ (and $\sum_{i=1}^{n} p_i\delta_{\omega_i}$) should be dense,
- $\int Y \mathrm{d}P_n \to \int Y \mathrm{d}P$,
- We do have (stochastic) optimization problems in mind.

Another link

There is a 1:1 relationship

A spectral risk measure *always* comes with Wasserstein.

Theorem (For Measures P and $\tilde{P}$ on the real line $\mathbb{R}$)

For a measure P on the real line $\mathbb{R}$ and $r = 2$, a random variable Y and U uniform, then

$$2 \cdot \mathcal{R}_\sigma(Y) = \|Y\|_{L^2}^2 + \|\sigma\|_{L^2}^2 - w_2\left(P^Y, P^{\sigma(U)}\right)^2$$

The Dual for the Wasserstein/ Kantorovich Distance

Theorem (Kantorovich Rubinstein)

The Dual of the Wasserstein problem reads $w(P,Q)$

$$\begin{array}{rl} \text{maximize (in } Y) & \mathbb{E}_P Y - \mathbb{E}_Q Y \\ \text{subject to} & Y(\xi) - Y\left(\tilde{\xi}\right) \le d\left(\xi, \tilde{\xi}\right). \end{array}$$

Change of measure
with Wasserstein

Proposition (A tight bound)

Let $\mathcal{R}_{\mathcal{S}}$ be a general risk functional. Suppose that the random variables Y, $\tilde{Y}: \Xi \to \mathbb{R}$ satisfy

$$Y(\xi) - \tilde{Y}(\tilde{\xi}) \leq L \cdot d(\xi, \tilde{\xi}).$$

Then

$$\mathcal{R}_{\mathcal{S};P}(Y) - \mathcal{R}_{\mathcal{S};Q}(\tilde{Y}) \leq L \cdot w_r(P, Q) \cdot \sup_{\sigma \in \mathcal{S}} \|\sigma\|_q \,,$$

where $q \in (1, \infty]$ is the Hölder conjugate exponent of r (the order of the Wasserstein metric), i.e., $\frac{1}{q} + \frac{1}{r} = 1$.

Corollary

Consider the problem

$$v(P) := \textit{minimize}\ \mathcal{R}_P(c(x,\xi)) \\ \textit{subject to}\ x \in \mathbb{X},$$

then

$$v(P) - v(\tilde{P}) \leq L \cdot w_r(P,\tilde{P}) \cdot \sup_{\sigma\in\mathcal{S}} \|\sigma\|_q,$$

Outline

Classical stochastic problem formulations include

Aleatoric risk

The three problems are

- parametric optimization
  - one parameter ξ: standard deterministic optimization
  - set of possible parameters $\xi \in \Xi$: robust optimization
  - the parameters are distributed $\xi \sim P$: stochastic optimization

$\min_{y\in Y} c(y)$

$\min_{y\in Y} \max_{\xi\in\Xi} c(y,\xi)$

$\min_{y\in Y} \mathcal{R}(c(y,\xi))$

TECHNISCHE UNIVERSITÄT CHEMNITZ

Problem

The typical formulation of the multistage problem reads

$$\text{minimize } \mathcal{R}(c(x,\xi)),$$
$$x \in \mathbb{X}$$

Problem (Probabilistic constraints)

$$\text{minimize}_{x\in\mathbb{X}} \mathcal{R}(c(x,\xi)),$$
$$\text{subject to } P(g(x,\xi) \le 0) \ge \alpha$$

Problem (Markowitz)

$$\text{minimize}_{x\in\mathbb{X}} \mathcal{R}(c(x,\xi)),$$
$$\text{subject to } \mathbb{E}\, g(x,\xi) \ge \mu$$

Problem (alternative Markowitz)

$$\text{maximize } \mathbb{E}\, g(x,\xi),$$
$$\text{subject to } \mathcal{R}(c(x,\xi)) \le c$$
$$x \in \mathbb{X}$$

Problem (Integrated risk mgmt)

$$\begin{aligned} &\text{minimize}\, \gamma \cdot \mathbb{E}\, g(x,\xi) + (1-\gamma)\cdot \mathcal{R}(c(x,\xi)),\\ &\text{subject to } x \in \mathbb{X} \end{aligned}$$

Problem (Integrated risk mgmt)

$$\begin{aligned} \text{minimize}_{\text{in } z(\cdot)}\ & \mathrm{AV@R}_\alpha\left(\int_0^1 \big(u(x,z(x))-1\big)^2 \mathrm{d}x\right) + \frac{\alpha}{2}\int_0^1 z(x)^2 \mathrm{d}x,\\ \text{subject to } & \nu(\xi) u_{xx}(\xi,x) + u(\xi,x)\cdot u_x = f(x) + z(x)\\ & u(0,\cdot) = d_0(\cdot) \text{ and } u(1,\cdot) = d_1(\cdot) \end{aligned}$$

Outline

Following Ellsberg (1961) we distinguish between the

- *uncertainty problem* (aleatoric), if the model is fully known, but the realizations of the random variables are unknown; and the
- *ambiguity problem* (epistemic), if the probability model itself is unknown. Another name for ambiguity is *Knightian uncertainty* (referring to F. Knight's 1921 book Knight (1921)).

optimization with randomly distributed parameters

- there is only one distribution $\mathcal{P}$: stochastic optimization
- set of possible distributions: $P \in \mathcal{P}$: distributionally robust optimization
- the parameters are random themselves, $P \sim \mu$: Bayesian ambiguous optimization

Example: $\mathcal{P} := \{Q \colon w(P,Q) \le \varepsilon\}$.

The ambiguity extension considers the new objective

$$\min_y \max_{P \in \mathcal{P}} \mathcal{R}_P\big(c(y,\xi)\big).$$

Ambiguity extension (cont.)

Epistemic risk

We consider the ambiguity problems (distributionally robust, epistemic) in

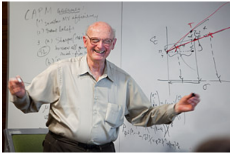
- optimization,
- 2 stage stochastic optimization
- multistage stochastic optimization and
- dynamic optimization.

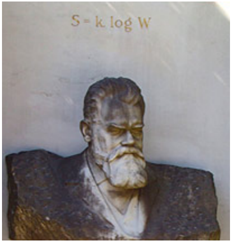
Outline

TECHNISCHE UNIVERSITÄT CHEMNITZ

Conclusion

$$\mathsf{EV@R}_\alpha(Y) := \inf_{t>0} \frac{1}{t} \log \frac{1}{1-\alpha}\, \mathbb{E}\, e^{tY}$$
$$= \sup\left\{ \mathbb{E}\, YZ \;\middle|\; \begin{array}{l} Z \geq 0,\, \mathbb{E}\, Z = 1 \text{ and} \\ H(Z) \leq \log \frac{1}{1-\alpha} \end{array} \right\}$$





1. Entropies: Boltzmann — Shannon — Rényi
2. Risk measures based on Entropies and relations to Wasserstein
3. Dual representation, even in non-convex situations
4. Empirical measure
5. Ambiguity

References

Mathematik!
TU Chemnitz

Ahmadi-Javid, A. and Pichler, A. (2017). An analytical study of norms and Banach spaces induced by the entropic value-at-risk. *Mathematics and Financial Economics*, 11(4):527–550.

Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent Measures of Risk. *Mathematical Finance*, 9:203–228.

Dentcheva, D., Penev, S., and Ruszczyński, A. (2010). Kusuoka representation of higher order dual risk measures. *Annals of Operations Research*, 181:325–335.

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *The Quarterly Journal of Economics*, 75(4):643–669.

Knight, F. H. (1921). *Risk, Uncertainty and Profit*. Library of Economics and Liberty.

Rockafellar, R. T. (1970). *Convex Analysis*. Princeton University Press.