STK4400 – Week 4

K. R. Dahl & A. B. Huseby

Department of Mathematics
University of Oslo, Norway

Today:

K. R. Dahl & A. B. Huseby

Dual characterization of convex risk measures.

Examples: Two measures of risk.

Recap: Convex risk measures

Definition (Convex risk measure)

A convex risk measure is a function $\rho : \mathbb{X} \to \mathbb{R}$ which satisfies the following for each $X, Y \in \mathbb{X}$:

(*i*) (Convexity) $\rho(\lambda X + (1-\lambda)Y) \leq \lambda\rho(X) + (1-\lambda)\rho(Y)$ for $0 \leq \lambda \leq 1$.

(*ii*) (Monotonicity) If $X \leq Y$, then $\rho(X) \geq \rho(Y)$.

(*iii*) (Translation invariance) If $m \in \mathbb{R}$, then $\rho(X + m\mathbf{1}) = \rho(X) - m$.

If $\rho(X) \leq 0$, X is acceptable since it does not have a positive risk. On the other hand, if $\rho(X) > 0$, X is unacceptable.

Recap: Interpretations of convex risk measures

Interpret ρ as a capital requirement, that is: $\rho(X)$ is the extra amount of money which should be added to the portfolio in a risk free way to make the position acceptable for an agent.

Interpretation of the conditions in Definition 1:

Convexity: Diversification reduces risk. The total risk of loss in two portfolios should be reduced when the two are weighed into a mixed portfolio.

Recap: Interpretations of convex risk measures, ctd.

K. R. Dahl & A. B. Huseby

Monotonicity: The risk of loss is reduced by choosing a portfolio that has a higher value in every possible state of the world.

Translation invariance: ρ is the amount of money one needs to add to the portfolio in order to make it acceptable for an agent. Hence, if one adds a risk free amount m to the portfolio, the capital requirement should be reduced by the same amount.

A dual characterisation of convex risk measures

K. R. Dahl & A. B. Huseby

Goal today: Derive dual characterisation of a convex risk measure ρ.

Result by Frittelli and Gianin.

Let V be a vector space paired with the vector space $\mathbb{X}$ of financial positions.

What is a pairing? Let's recall!

Conjugate functions in paired spaces

Definition (Pairing of spaces)

A pairing of two linear spaces X and V is a real valued bilinear form $\langle \cdot, \cdot \rangle$ on $X \times V$.

The pairing associates for each $v \in V$ a linear function $\langle \cdot, v \rangle : x \mapsto \langle x, v \rangle$ on X, and similarly for X.

Compatible topology

Definition (Compatible topology)

Assume there is a pairing between the spaces X and V. A topology on X is compatible with the pairing if it is a locally convex topology such that the linear function $\langle \cdot, v \rangle$ is continuous, and any continuous linear function on X can be written in this form for some $v \in V$. A compatible topology on V is defined similarly.

Definition (Paired spaces)

X and V are paired spaces if one has chosen a pairing between X and V, and the two spaces have compatible topologies with respect to the pairing.

Paired space with the space of financial positions

K. R. Dahl & A. B. Huseby

Let V be a vector space paired with the vector space $\mathbb{X}$ of financial positions.

Example: If $\mathbb{X}$ is given a Hausdorff topology, so it becomes a topological vector space (for definitions of these terms, see Pedersen), V can be the set of all continuous linear functionals from $\mathbb{X}$ into $\mathbb{R}$, as in Frittelli and Gianin.

Conjugate and biconjugate of a convex risk measure

K. R. Dahl & A. B. Huseby

Let ρ^* denote the (convex) conjugate of ρ.

Let ρ^{**} denote the biconjugate of ρ.

We let $\langle \cdot, \cdot \rangle$ be a pairing between V and the space of financial positions $\mathbb{X}$.

Recap: What is the conjugate and biconjugate?

Recall: Convex conjugate of a function

K. R. Dahl & A. B. Huseby

Definition (Convex conjugate of a function, f^*)

Let X and V be paired spaces. For a function $f : X \to \bar{\mathbb{R}}$, define the conjugate of f, denoted by $f^* : V \to \bar{\mathbb{R}}$, by

$$f^*(v) = \sup\{\langle x, v\rangle - f(x) : x \in X\}. \tag{1}$$

Finding f^* is called taking the conjugate of f in the convex sense.

Biconjugate of a function and the Fenchel transform

Definition (Biconjugate of a function, f^{**})

Let X and V be paired spaces. For a function $f : X \to \bar{\mathbb{R}}$, define *the biconjugate* of f, f^{**}, to be the conjugate of f^*, so $f^{**}(x) = \sup\{\langle x, v\rangle - f^*(v) : v \in V\}$.

Definition

The operation $f \mapsto f^*$ is called *the Fenchel transform*.

If $f : \mathbb{R}^n \to \bar{\mathbb{R}}$, then the operation $f \mapsto f^*$ is sometimes called the Legendre-Fenchel transform.

General duality result for convex risk measures

Theorem

Let $\rho : \mathbb{X} \to \mathbb{R}$ be a convex risk measure. Assume in addition that ρ is lower semi-continuous. Then $\rho = \rho^{**}$. Hence for each $X \in \mathbb{X}$

$$\begin{aligned} \rho(X) &= \sup\{\langle X, v\rangle - \rho^*(v) : v \in V\} \\ &= \sup\{\langle X, v\rangle - \rho^*(v) : v \in \mathrm{dom}(\rho^*)\} \end{aligned}$$

where $\langle \cdot, \cdot \rangle$ is a pairing between $\mathbb{X}$ and V.

Recall: A duality theorem

K. R. Dahl & A. B. Huseby

Before we do the proof, let's recall the following duality theorem:

Theorem

Let $f : X \to \bar{\mathbb{R}}$ be arbitrary. Then the conjugate f^* is a closed, convex function on V and $f^{**} = \mathrm{cl}(\mathrm{co}(f))$. Similarly if one starts with a function in V.

In particular, the Fenchel transform induces a one-to-one correspondence $f \mapsto h, h = f^*$ between the closed, convex functions on X and the closed, convex functions on V.

Proof:

Since ρ is a convex risk measure, it is a convex function (from the definition).

Hence, the convex hull of ρ is equal to ρ, i.e., $\mathrm{co}(\rho) = \rho$ (recall the def. of convex hull).

Also, since ρ is lower semi-continuous and always greater than $-\infty$, ρ is closed (see comment after the def. of closed functions, week 2).

This means that

$$\mathrm{cl}(\rho) = \rho.$$

Proof, ctd.:

Therefore

$$\mathrm{cl}(\mathrm{co}(\rho)) = \mathrm{cl}(\rho) = \rho.$$

But from the duality theorem recalled above, Theorem 9,

$$\rho^{**} = \mathrm{cl}(\mathrm{co}(\rho)),$$

hence $\rho = \rho^{**}$.

The second to last equation in the theorem follows directly from the definition of ρ^{**}.

The last equation follows because the supremum cannot be achieved when $\rho^* = +\infty$.

The finite dimensional case

Theorem 8 is in an abstract form.

If we choose a specific set of paired spaces, $\mathbb{X}$ and V, we get more specific results.

Now, a theorem by Föllmer and Schied:

Let $\mathbb{X} = \mathbb{R}^n$, $V = \mathbb{R}^n$ be paired spaces with the standard Euclidean inner product, denoted $\cdot$, as pairing.

Let $(\Omega, \mathcal{F})$ be a measurable space and let $\mathcal{P}$ denote the set of all probability measures over Ω.

Dual representation of convex risk measures: Finite Ω

Theorem

Assume that Ω is finite. Then, any convex risk measure $\rho : \mathbb{X} \to \mathbb{R}$ can be represented in the form

$$\rho(X) = \sup_{Q \in \mathcal{P}} \{E_Q[-X] - \alpha(Q)\} \tag{2}$$

where $E_Q[\cdot]$ denotes the expectation with respect to Q and $\alpha : \mathcal{P} \to (-\infty, \infty]$ is a "penalty function" which is convex and closed. Actually, $\alpha(Q) = \rho^*(-Q)$ for all $Q \in \mathcal{P}$.

Proof (Luthi and Doege)

To show that $\rho : \mathbb{X} \to \mathbb{R}$ (as in Theorem 10) is a convex risk measure, we check Definition 1:

Convexity:
Let $\lambda \in [0,1], m \in \mathbb{R}, X, Y \in \mathbb{X}$.

$$
\begin{aligned}
\rho(\lambda X + (1-\lambda)Y) &= \sup_{Q \in \mathcal{P}} \{E_Q[-(\lambda X + (1-\lambda)Y)] - \alpha(Q)\} \\
&= \sup_{Q \in \mathcal{P}} \{\lambda E_Q[-X] + (1-\lambda)E_Q[-Y] - \alpha(Q) \\
&\leq \lambda \sup_{Q \in \mathcal{P}} \{E_Q[-X] - \alpha(Q)\} \\
&\quad + (1-\lambda) \sup_{Q \in \mathcal{P}} \{E_Q[-Y] - \alpha(Q)\} \\
&= \lambda\rho(X) + (1-\lambda)\rho(Y).
\end{aligned}
$$

Proof, ctd.:

Monotonicity: Assume $X \leq Y$. Then $-X \geq -Y$, so

$$\begin{aligned} \rho(X) &= \sup_{Q \in \mathcal{P}} \{E_Q[-X] - \alpha(Q)\} \\ &\geq \sup_{Q \in \mathcal{P}} \{E_Q[-Y] - \alpha(Q)\} \\ &= \rho(Y). \end{aligned}$$

Translation invariance:

$$\begin{aligned} \rho(X + m1) &= \sup_{Q \in \mathcal{P}} \{E_Q[-(X + m1)] - \alpha(Q)\} \\ &= \sup_{Q \in \mathcal{P}} \{E_Q[-X] - mE_Q[1] - \alpha(Q)\} \\ &= \sup_{Q \in \mathcal{P}} \{E_Q[-X] - m - \alpha(Q)\} \\ &= \sup_{Q \in \mathcal{P}} \{E_Q[-X] - \alpha(Q)\} - m \end{aligned}$$

Proof, ctd.

To prove the converse: Assume that ρ is a convex risk measure.

The conjugate function of ρ, denoted ρ^*, is

$$\rho^*(v) = \sup_{X \in \mathbb{X}} \{v \cdot X - \rho(X)\}$$

(where $\cdot$ denotes Euclidean inner product) for all $v \in V = \mathbb{R}^n$.

Fix $X \in \mathbb{X}$ and consider $Y_m := X + m1 \in \mathbb{X}$ for an arbitrary $m \in \mathbb{R}$.

Proof, ctd.

Then

$$\rho^*(v) \geq \sup_{m \in \mathbb{R}} \{v \cdot Y_m - \rho(Y_m)\}$$

because $\{Y_m\}_{m \in \mathbb{R}} \subset \mathbb{X}$.

This means that

$$\begin{aligned} \rho^*(v) &\geq \sup_{m \in \mathbb{R}} \{v \cdot (X + m1) - \rho(X + m1)\} \\ &= \sup_{m \in \mathbb{R}} \{m(v \cdot 1 + 1)\} + v \cdot X - \rho(X) \end{aligned}$$

where the equality follows from the translation invariance of ρ (see Definition 1).

Proof, ctd.

The first term in the last expression is only finite if $v \cdot 1 + 1 = 0$, (where $1 = (1, 1, \ldots, 1) \in \mathbb{R}^n$).

That is, if $\sum_{i=1}^{n} v_i = -1$ (if not, one can make the first term go towards $+\infty$ by letting m go towards either $+\infty$ or $-\infty$).

Proof, ctd.

We have now proved that in order for $\rho^*(v) < +\infty$, $\sum_{i=1}^{n} v_i = -1$ must hold.

Again, consider an arbitrary, but fixed $X \in \mathbb{X}, X \geq 0$ (component-wise).

Then, for all $\lambda \geq 0$,

$$\lambda X \geq 0,$$

and $\lambda X \in \mathbb{X}$.

Hence, $\rho(\lambda X) \leq \rho(0)$ (from the monotonicity of ρ, see Definition 1).

Proof, ctd.

K. R. Dahl & A. B. Huseby

Therefore, by the same type of arguments as above

$$\rho^*(v) \geq \sup_{\lambda \geq 0}\{v \cdot \lambda X - \rho(\lambda X)\} \geq \sup_{\lambda \geq 0}\{v \cdot (\lambda X)\} - \rho(0).$$

Here, $\rho^*(v)$ is only finite if $v \cdot X \leq 0$ for all $X \geq 0$.

Hence $v \leq 0$.

The conjugate ρ^* is reduced to

$$\rho^*(v) = \begin{cases} \sup_{X \in \mathbb{X}}\{v \cdot X - \rho(X)\} \text{ where } v \cdot 1 = -1 \text{ and } v \leq 0 \\ +\infty \text{ otherwise }. \end{cases}$$

Proof, ctd.

Now, define $\alpha(Q) = \rho^*(-Q)$ for all $Q \in \mathcal{P}$.

From Theorem 8, $\rho = \rho^{**}$.

But

$$
\begin{aligned}
\rho^{**}(X) &= \sup_{v \in V}\{v \cdot X - \rho^*(v)\} \\
&= \sup_{Q \in \mathcal{P}}\{(-Q) \cdot X - \alpha(Q)\} \\
&= \sup_{Q \in \mathcal{P}}\{\sum_{i=1}^{n} Q_i(-X_i) - \alpha(Q)\} \\
&= \sup_{Q \in \mathcal{P}}\{E_Q[-X] - \alpha(Q)\}
\end{aligned}
$$

where Q_i, X_i denote the i'th components of the vectors Q, X respectively.

Proof, conclusion

Hence $\rho(X) = \rho^{**}(X) = \sup_{Q \in \mathcal{P}}\{E_Q[-X] - \alpha(Q)\}$.

This concludes the proof of the theorem.

Interpretation of the dual representation

Theorem 10 says that any convex risk measure $\rho : \mathbb{R}^n \to \mathbb{R}$ is the expected value of the negative of a contingent claim, $-X$, minus a penalty function, $\alpha(\cdot)$, under the worst case probability.

Note that we consider the expectation of $-X$, not X, since losses are negative in our context.

The penalty function $\alpha(Q)$

K. R. Dahl & A. B. Huseby

Know: The penalty function α in Theorem 10 is of the form $\alpha(Q) = \rho^*(-Q)$.

Luthi and Doege proved that it is possible to derive a more intuitive representation of α.

Theorem

Let $\rho : \mathbb{R}^n \to \mathbb{R}$ be a convex risk measure, and let $\mathcal{A}_\rho$ be its acceptance set (in the sense of Definition 12). Then, Theorem 10 implies that $\rho(X) = \sup_{Q \in \mathcal{P}}\{E_Q[-X] - \alpha(Q)\}$, where $\alpha : \mathcal{P} \to \mathbb{R}$ is a penalty function. Then, α is of the form

$$\alpha(Q) = \sup_{X \in \mathcal{A}_\rho} \{E_Q[-X]\}.$$

Recall: The acceptance set

K. R. Dahl & A. B. Huseby

Before we do the proof, let's recall:

Associated with every convex risk measure ρ, there is a natural set of all acceptable portfolios: The acceptance set, $\mathcal{A}_\rho$, of ρ.

Definition (The acceptance set of a convex risk measure, $\mathcal{A}_\rho$)

A convex risk measure ρ induces a set

$$\mathcal{A}_\rho = \{X \in \mathbb{X} : \rho(X) \leq 0\}$$

The set $\mathcal{A}_\rho$ is called the acceptance set of ρ.

Proof:

It suffices to prove that for all $Q \in \mathcal{P}$,

$$\rho^*(-Q) = \sup_{X \in \mathbb{X}} \{E_Q[-X] - \rho(X)\} = \sup_{X \in \mathcal{A}_\rho} \{E_Q[-X]\} \qquad (3)$$

since we know that $\alpha(Q) = \rho^*(-Q)$ (from the dual representation of convex risk measures).

For all $X \in \mathcal{A}_\rho$, $\rho(X) \leq 0$ (see the definition of acceptance sets), so

$$E_Q[-X] - \rho(X) \geq E_Q[-X].$$

Hence, since $\mathcal{A}_\rho \subseteq \mathbb{X}$

$$\rho^*(-Q) \geq \sup_{X \in \mathcal{A}_\rho} \{E_Q[X] - \rho(X)\} \geq \sup_{X \in \mathcal{A}_\rho} \{E_Q[-X]\}.$$

Proof, ctd.

K. R. Dahl & A. B. Huseby

To prove the opposite inequality (and hence to prove equation (3)):

Assume for contradiction that there exists $Q \in \mathcal{P}$ such that

$$\rho^*(-Q) > \sup_{X \in \mathcal{A}_\rho} \{E_Q[-X]\}.$$

From the definition of supremum, there exists a $Y \in \mathbb{X}$ such that

$$E_Q[-Y] - \rho(Y) > E_Q[-X] \text{ for all } X \in \mathcal{A}_\rho.$$

Proof, ctd.

Note that $Y + \rho(Y)1 \in \mathcal{A}_\rho$ since

$$\rho(Y + \rho(Y)1) = \rho(Y) - \rho(Y) = 0.$$

Therefore,

$$\begin{aligned} E_Q[-Y] - \rho(Y) &> E_Q[-(Y + \rho(Y)1)] \\ &= E_Q(-Y) + \rho(Y)E_Q[-1] \\ &= E_Q(-Y) - \rho(Y) \end{aligned}$$

This is a contradiction. Hence, the result is proved.

Recap: Dual representation of convex risk measures for finite scenario space

Together, Theorem 10 and Theorem 11 give thourough understanding of convex risk measures when Ω is finite:

Any convex risk measure $\rho : \mathbb{R}^n \to \mathbb{R}$ can be written in the form

$$\rho(X) = \sup_{Q \in \mathcal{P}} \{E_Q[-X] - \alpha(Q)\},$$

where $\alpha(Q) = \sup_{X \in \mathcal{A}_\rho} \{E_Q[-X]\}$ and $\mathcal{A}_\rho$ is the acceptance set of ρ.

The case where Ω is infinite

K. R. Dahl & A. B. Huseby

It is also possible to derive a similar characterization in the case where Ω is infinite.

This is left as self-study for those interested.

This is omitted here, and is left to self-study for those interested.

Two commonly used measures of financial risk

K. R. Dahl & A. B. Huseby

Aim: Present two measures of monetary risk which are frequently used in practice.

One of these measures, called Value-at-Risk, is not coherent, or even convex in general.

Why use measures not satisfying the economically reasonable conditions of convex and coherent risk measures?

Why is Value-at-Risk used in practice?

- Old habits die hard: VaR was used before the concepts coherent and convex risk measures were introduced. Hence, people are so used to the old measures that they are hesitant to implement others.
- Simplicity: As we will see, Value-at-Risk is a very intuitive concept.
- Good enough: In many practical situations, the results attained are sufficient, though the measures in question are economically unreasonable in theory.

Value-at-Risk

Value at risk, VaR, is the most commonly used risk measure in practice (insurance, banks, investment funds etc.).

Interpretation: For a given portfolio, time horizon and probability λ, VaR is the maximum potential loss over the time period after excluding the λ percent worst cases.

Let X be a random variable representing a financial position.

X may represent one stock, a portfolio of stocks or the financial holdings of an entire firm.

Negative values of $X(\omega)$ correspond to losses, and positive values to profit.

Mathematical definition of VaR

VaR is defined as follows: Fix some level $\lambda \in (0, 1)$ (typically close to 0), and define $Y := -X$.

Note that for the random variable Y, losses are positive numbers.

Then, VaR_λ is defined as the $(1 - \lambda)$-quantile of Y

$$\mathrm{VaR}_\lambda(X) := F_Y^{-1}(1 - \lambda) \tag{4}$$

where F_Y is the cumulative distribution function of Y.

VaR is the maximum potential loss after excluding the worst cases

One can show (this is left as an exercise) that

$$\mathrm{VaR}_\lambda(X) = -\inf\{x \in \mathbb{R} : F_X(x) > \lambda\}. \tag{5}$$

That is, VaR_λ is the maximum potential loss over the time period after excluding the λ worst cases.

VaR is decreasing

K. R. Dahl & A. B. Huseby

Lemma

$VaR_\lambda(X)$ *is decreasing in* λ.

Proof: Left as an exercise to the reader.

Alternative interpretation of VaR

K. R. Dahl & A. B. Huseby

One can alternatively define VaR as

$$\mathrm{VaR}_\lambda(X) := \inf\{m \in \mathbb{R} | P(X + m < 0) \leq \lambda\}, \qquad (6)$$

i.e., $\mathrm{VaR}_\lambda(X)$ is the smallest amount of money that needs to be added to X in order for the probability of a loss to be less than λ.

Applications of VaR

K. R. Dahl & A. B. Huseby

Financial firms and banks use VaR to quantify the risk of their investments: Can monitor their current risk at any time to measure their potential losses.

In practice, firms will specify their VaR depending on the confidence level λ, but also depending on some time horizon.

The definition of VaR (4) is typically not used directly for computing the value at risk, since this formula requires that we know the exact distribution of X.

Most banks, insurance firms etc. use historical data as an approximation to the exact distribution, and compute the quantile in (4) based on this.

Monte Carlo to calculate VaR

As an alternative, some use Monte Carlo methods based on a stochastic model of the financial markets.

Monte Carlo methods are based on random sampling from the stochastic model.

The Monte Carlo approach is more time-consuming, and usually involves additional work by an analyst in order to fit the parameters of the model to the relevant problem based on historical data.

Drawback with the historical data method: This implicitly assumes that the future distribution will be the same as the past one, no further randomness or adaptation to the general economic situation is included.

Time and VaR

Note that in our mathematical definition of VaR above, we didn't mention time at all.

For practical interpretations of VaR, one should think of our random variable X as the profit/loss random variable for the financial position.

So, if a bank wants to compute their one-month VaR, X is the (uncertain) difference between the current value of the banks financial holdings and the value a month from now.

Practical VaR

How is VaR used and interpreted in practice?

Say a portfolio of stocks has a one-day 2% VaR of NOK 10 million.

Then, there is a 0.02 probability that the value of the portfolio will decrease by more than NOK 10 million during this day, assuming no trading.

On average, the bank will expect to lose more than this 1 out of 50 days.

Note that it is very important for the VaR calculation that there is no trading happening in the portfolio. If there is trading, the distribution of the portfolio will change, and hence also the VaR.

Drawbacks of VaR

Despite its use in practice, value at risk has major drawbacks:

- In general, VaR_λ is not convex, see Föllmer and Knispel: Diversification may increase the risk w.r.t. VaR. This is economically unreasonable.
- From equation (4): VaR_λ ignores extreme losses which occur with small probability. This tail insensitivity makes it an unsuitable measure of risk in situations where the consequences of large losses are very bad (e.g., insurance companies not being able to pay their customers).

These drawbacks of VaR lead to the development of the theory of convex and coherent risk measures.

Still widely used in practice, despite its deficiencies.

Average value at risk

Average value at risk (AVaR), also called expected shortfall (ES) or conditional value at risk (CVaR), was introduced to mend the deficiencies of value at risk.

For $\lambda \in (0,1]$, the average value at risk is defined as

$$\mathrm{AVaR}_\lambda(X) := \frac{1}{\lambda}\int_0^\lambda \mathrm{VaR}_\alpha(X)d\alpha. \tag{7}$$

Average value at risk can be interpreted as the expected loss in a presupposed percentage of worst cases.

AVaR $\geq$ VaR

Note that

$$\mathrm{AVaR}_{\lambda}(X) \geq \mathrm{VaR}_{\lambda}(X),$$

(the proof is left as an exercise).

So, when considering the same level λ, the average value at risk is always greater than or equal the value at risk.

AVaR is a coherent risk measure

Föllmer and Schied prove that AVaR_λ is a coherent risk measure, with a dual representation

$$\mathrm{AVaR}_\lambda(X) = \max_{Q \in \mathcal{Q}_\lambda} E_Q[-X]$$

where $\mathcal{Q}_\lambda := \{Q << P | \frac{dQ}{dP} \leq \lambda\}$. That is, $\mathcal{Q}_\lambda$, is the set of all measures Q that are absolutely continuous w.r.t. the measure P given that the Radon-Nikodym derivative of Q w.r.t. P is less than or equal λ (see Shilling for more on these measure theoretical concepts).

Note also that for $\lambda = 1$, average value at risk reduces to $E_P[-X]$, i.e., the expected loss.

Other examples of convex risk measures

Other examples of convex risk measures are:

- Shortfall risk.
- Divergence risk measures

We refer those interested to Föllmer and Knispel.

Next week

Recap!

We were supposed to have a physical lecture, but this is cancelled due to the current corona situation.

There will be a lecture recording instead.