Risk and performance measurement
– lecture notes –

dr hab. Marcin Pitera
Jagiellonian University

May 19, 2026

Contents

Introduction **2**

1 Preliminaries **2**
1.1 Random variables and their characteristics . . . 2
1.2 Mean-variance criterion . . . 4
1.3 Expected utility based criterions . . . 4

2 Normative properties of risk and performance measures **7**
2.1 Definition of normative properties . . . 7
2.2 Selected interactions between normative properties . . . 10
2.3 Normative properties of certainty equivalents and entropic utility measure . . . 11

3 Important families of risk measures **12**
3.1 Monetary risk measures . . . 12
3.2 Regulatory risk measures: value-at-risk and expected shortfall . . . 13
3.3 Coherent risk measures . . . 18
3.4 Convex risk measures . . . 21

4 From risk to performance measurement **24**
4.1 Risk-to-reward ratios . . . 25
4.2 Coherent acceptability indices . . . 26
4.3 Quasi-concave acceptability indices . . . 28

5 Estimating risk and evaluating the performance **30**
5.1 Plug-in estimation procedures . . . 31
5.2 Risk estimators and their statistical properties . . . 33
5.3 Statistical sample construction and its impact on risk estimation . . . 35
5.4 Value-at-risk backtesting and exception rate performance measure . . . 40
5.5 Evaluation of point forecasts and mean score functions . . . 43
5.6 Evaluation of density forecasts and PIT-backtests . . . 45

6 Selected topics in risk and performance measurement **48**
6.1 Time-consistency and dynamic measures . . . 48
6.2 Risk contribution and incremental analysis . . . 51
6.3 Risk allocation and the Euler principle . . . 52

7 Selected applications of risk and performance measures **54**
7.1 Risk and performance measures in various disciplines . . . 54
7.2 Objectives based on risk and performance measures . . . 59
7.3 Example: Averaged long-run risk sensitive stochastic control . . . 61
7.4 Example: Acceptability good-deal bounds and market efficiency evaluation . . . 63

Introduction

In this course, we assume familiarity with fundamental concepts, techniques, and theorems from Mathematical Analysis, Measure Theory, Probability Theory, and Algebra. Although this lecture is designed to be independent and self-contained, parts of it align with standard references in risk theory and risk management; see, for instance, [FS02a] and [MFE10] for further reading. Some parts of the lecture notes are also based on [BR11].

Warning: The main goal of this lecture is to present the theoretical part of the fundamentals of risk and performance measures. However, bearing in mind that this is an introductory course, we have devoted some portion of it to the presentation of examples aimed at developing probabilistic intuition. These examples should be supplemented by exercises completed during workshop class. We also encourage you to look into selected books (problem collections), where many interesting problems can be found. In addition to the aforementioned sources, problems can also be found, for example, in [Ale09].

1 Preliminaries

W1

The terms *risk* and *performance* have become part of everyday language. Although their intuitive meaning is clear, no single unifying definition has been established. They are perhaps best regarded as *primitive notions*, closely related to ideas such as *preference*, *hazard*, *fortune*, *safety*, *opportunity*, and *uncertainty*, each allowing diverse interpretations.

From a mathematical perspective, our interest lies in *quantifying* risk and performance, that is, developing measurement methods. Over time, many elegant concepts have emerged, with major contributions from Daniel Bernoulli, John von Neumann, and Oskar Morgenstern, among others. For comprehensive historical accounts and overviews, see [Fis88; Kre88; FS02a].

In this work, we focus on the quantification of *utility*—a concept encompassing both risk and performance—within the modern normative framework introduced in [Art+97]. This framework includes the notion of *coherent risk measures*, which has inspired extensive theoretical and practical research, leading to developments such as *convex risk measures* [FS02b] and *acceptability indices* [CM09], both of which will be introduced and discussed during the lecture.

In this lecture, we concentrate primarily on the *static setup*, where random variables represent future (discounted) returns or profit-and-loss (P&L) values of investments. We define various functionals that assign a real number to each position, thereby quantifying the *utility* of the corresponding random variable.[1] The main emphasis is on the normative approach of [Art+97], which imposes axioms such as *cash-additivity*, *convexity*, and *monotonicity*, each with clear financial interpretation. For simplicity, we focus our attention on performance measurement of financial and actuarial positions; however, the theory extends naturally to other settings and can serve as a foundation for constructing efficient objective functions in stochastic optimization problems, including applications in machine learning.

1.1 Random variables and their characteristics

Unless stated otherwise, we work on a fixed probability space $(\Omega, \mathcal{F}, \mathbb{P})$. All random variables introduced in this course are assumed to be real and all inequalities are to be understood in the

[1] We consider a standard setup in which the measurement procedure could be applied and do not consider problems in which e.g. *Knightian Uncertainty* emerges, see [MFE10, Section 1.1] foe details.

almost sure sense. Intuitively, as noted in the introduction, one should link random variable output to a future (discounted) returns or profit-and-loss (P&L) of a financial investment.

We assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is rich enough and does not have any atoms. More formally, this can be achieved by assuming that $(\Omega, \mathcal{F}, \mathbb{P})$ is an atomless *standard* probability space[2], that is, it is isomorphic to $([0,1], \mathcal{B}([0,1], \lambda)$, where $\mathcal{B}(\cdot)$ denotes the standard Borel σ-algebra on $[0,1]$ and λ is a Lebesgue measures on $[0,1]$. Note that the standard probability space is rich enough to define a sequence of i.i.d. random variables on it. While almost all of the results could be easily reformulated for the general case, we make this assumption to avoid many technical difficulties.

For $p \in \mathbb{N}_+$ and σ-algebra $\mathcal{G}$, such that $\mathcal{G} \subseteq \mathcal{F}$, we denote by $L^p(\Omega, \mathcal{G}, \mathbb{P})$ the space of all (a.s. identified) $\mathcal{G}$-measurable random variables with finite pth moment, that is, satisfying inequality $\int_\Omega |X|^p \, d\mathbb{P} < \infty$. Also, we consider the limit cases and use $L^0(\Omega, \mathcal{G}, \mathbb{P})$ to denote the space of all random variables, and $L^\infty(\Omega, \mathcal{G}, \mathbb{P})$ to denote the space of (a.s.) bounded random variables. For brevity, we also set $L^p := L^p(\Omega, \mathcal{F}, \mathbb{P})$, for $p \in [0, \infty]$.

To avoid technical expositions, most statements in this lecture notes are given for L^∞ and we consider the norm $\|X\|_\infty := \operatorname{ess\,sup} |X|$, for $X \in L^\infty$; most object could be easily redefined on a more general spaces linked e.g. to L^p, Orlicz hearts, or L^0-modules. If deemed necessary, we provide a top-level comments about the generalisations and necessary result statement modification.

For any random variable X we use F_X to denote it cumulative distribution function (CDF), that is, we set $F_X(t) := \mathbb{P}[X \leq t]$, for $t \in \mathbb{R}$, and sometimes use $\mathbb{P}_X$ to denote its probability distribution function (PDF). We are interested in the systemic analysis of functionals that map random variables to real numbers, that is mappings of the form

$$\rho\colon L^\infty \to \mathbb{R} \cup \{+\infty\}, \tag{1.1}$$

which extract specific characteristic of the random variables linked to their performance or risk. For completeness, we allow ρ to be infinite but typically, especially in risk measurement context, we assume finiteness. The most common distribution characteristic are the expected value and variance, that, for any $X \in L^\infty$, are given by

$$\mathbb{E}[X] := \int_\Omega X \, d\mathbb{P}, \quad \text{and} \quad \operatorname{Var}[X] := \int_\Omega (X - \mathbb{E}[X])^2 \, d\mathbb{P}$$

Essentially, one could look at both $\mathbb{E}[\cdot]$ and $\operatorname{Var}[\cdot]$ as mappings of the form (1.1) which assign numerical values to any element of L^∞. Note that mean and variance depend only on the PDF (CDF) of the underlying random variable since $\mathbb{E}[X] = \int_{-\infty}^{\infty} x \mathbb{F}_X(x)$; this is linked to so called *law-invariance* property of mapping (1.1) that will be introduced later.

Of course, one cannot fully represent a random variable (or its CDF) by a single number. Still, we aim to establish a partial order in the set of random variables using their characteristics such as mean that will help us to establish decision rules. If one is only interested in potential rewards, then one could use the expected value to establish an order $\succeq$ by simply saying that

$$X \succeq Y \text{ if any only if } \mathbb{E}[X] \geq \mathbb{E}[Y]. \tag{1.2}$$

In the financial context, we can say that we prefer investment X over Y if on average we earn more from X. For the mean, we get basic ordering properties from integral representation of the expected value. Namely, for any $X, Y \in L^\infty$ we get the following properties: (1) If $X \geq Y$, then $\mathbb{E}[X] \geq \mathbb{E}[Y]$; (2) for any $a, b \in \mathbb{R}$, we have $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$, which allow us to better understand how the order imposed by $\mathbb{E}$ would look like. One of the major flaws of the expected value is that it does not quantify the risk of X, e.g. represented by the heaviness of its loss-tail.

[2] also called *Lebesgue-Rokhlin* probability space, see [Rok49].

1.2 Mean-variance criterion

If one is interested in the criterion which incorporates both reward and risk, a popular choice is the mean-variance criterion.

Definition 1.1 (Mean-variance criterion)**.** For any parameter $\gamma \in \mathbb{R}_-$ we define the mapping

$$\mathrm{MV}_\gamma(X) := \mathbb{E}[X] + \tfrac{\gamma}{2}\,\mathrm{Var}[X], \quad X \in L^\infty \tag{1.3}$$

and call it **mean-variance criterion** with risk aversion coefficient γ.[a]

[a] The mean-variance criterion is often parameterized with a positive coefficient $\tilde{\gamma} > 0$, related to our convention by $\tilde{\gamma} := -\gamma$. In that case, the risk aversion parameter satisfies $\tilde{\gamma} \in \mathbb{R}_+$, and the sign in (1.3) is reversed.

The mapping defined in (1.3) could be seen as a generalization of the mean criterion in which we introduce a penalty function linked to the dispersion of the underlying random variable: the bigger it is, the more we subtract from the reward. The $\gamma \in \mathbb{R}_-$ coefficient sets out our preferences with respect to the risk-to-reward tolerance. Of course, this criterion is also not perfect, as it misses a lot of information linked, e.g., to skewness, could over-penalize gain-side asymmetry, and is in fact not monotone, that is, for any $\gamma \in \mathbb{R}_-$ we could get $\mathrm{MV}_\gamma(X) > \mathrm{MV}_\gamma(X)$ even if $X < Y$. Despite this major drawback, the mean-variance criterion is among the most frequently used criterion in optimization problems and is critical, for example, for modern portfolio theory; see [MFE10].

Of course, one can define multiple other characteristics of probability distribution (among the classical choices, we can find central moments, skewness, kurtosis, or quantile-based measures) and try to build criterions with them using, for example, risk-to-reward additive functions such as the one presented in (1.3) or even built other types of functions such as risk-to-reward ratios, e.g. Sharpe ratio, which we will define at a later stage. The analysis of mean and mean-variance criterions show us that a more systematic study of the maps which assign numerical values to random variables X is required.

Remark 1.2 (Monotonicity of mean-variance criterion for Gaussian random variables)**.** Even though the mean-variance criterion is not monotone in general, one can show that it is monotone e.g. if we fix a multivariate normal random variable $X = (X_1, \ldots, X_n)$ and restrict the space of random variables to the linear combinations of its margins. In fact, assuming that X is non-degenerate, one can show that for any $\gamma \in \mathbb{R}_-$ and weighted allocation (with weights summing to one), there exists a unique allocation which maximizes criterion (1.3). This observations, following directly from the linear-quadratic optimal control theory, is important in the context of Markowitz's portfolio theory in which X is used to express future random portfolio returns of equity stocks on some market, see [MFE10]. ♢

1.3 Expected utility based criterions

The *expected utility* based objective functions are one of the cornerstones of the modern decision making (under uncertainty) theory. They are used in many fields such as mathematical finance, economy, or physics. In this section, we provide a brief excerpt from the expected utility theory that is focused on the objective function construction. For more details and mathematical exposition, we refer to [FS02a, Section 2]. In the expected utility theory, the objective functions are taken as expected values for a given utility function.

Definition 1.3 (Utility function and expected utility criterion)**.** We call a function $U\colon \mathbb{R}\to\mathbb{R}$ a *utility function* if it is strictly concave and strictly increasing. Given utility function U, we define mapping $\varphi_U : L^\infty \to \mathbb{R}$, where

$$\varphi_U(X) := \mathbb{E}[U(X)], \quad X \in L^\infty, \tag{1.4}$$

and call it **the expected utility criterion** based on U.

It should be noted that we have pre-imposed *strictly increasing* and *strictly concave* property of function U to ensure that the order imposed by U is rational. Namely, the strictly increasing property of U is required to ensure that the standard order (in the space of random variables) is preserved after applying the criterion, that is, for any $a, b \in \mathbb{R}$ such that $a > b$ we get $\varphi_U(a) > \varphi_U(b)$. On the other hand, the *strictly concave* property is required to get the risk-aversion property which ensures that we prefer deterministic outcome when compared with a random outcome with the same expected value, that is, for any non-deterministic $X \in L^\infty$ we get $\varphi_U(\mathbb{E}(X)) > \varphi_U(X)$; see [FS02a, Proposition 2.35] for proofs. Also, note that utility function U is continuous as it is strictly increasing and concave.

Notably, one can show that any order that satisfies certain axioms related to completeness, transitivity, continuity, and independence have a numerical representation (called con Neumann-Morgenstern utility representation) of the form as in (1.4). Nevertheless, the detailed analysis of this remarkable result, often called *von Neumann–Morgenstern utility theorem*, is out of scope of this lecture, see e.g. [FS02a, Theorem 2.28].

From the risk and performance measurement point of view, we can apply any strictly increasing function on φ_U without changing the ordering. A specific choice of such function, which allow us to maintain deterministic outcome values, is U^{-1} which leads to the definition of certainty equivalent.

Definition 1.4 (Certainty equivalent)**.** Given utility function $U\colon \mathbb{R}\to\mathbb{R}$, we define function $\rho_U : L^\infty \to \mathbb{R}$, where

$$\rho_U(X) := U^{-1}\mathbb{E}[U(X)], \quad X \in L^\infty, \tag{1.5}$$

and call it **the certainty equivalent criterion** based on U.

Note that $\rho_U(X)$ is simply determining a deterministic outcome under which we are indifferent between X and this outcome, that is, we have $\rho_U(X) = \rho_U(\rho_U(X))$.

Another object important in the utility theory that will be later used in the lecture is *Arrow–Pratt coefficient of absolute risk aversion.*

Definition 1.5 (Arrow–Pratt measure of absolute risk aversion)**.** Given utility function $U\colon \mathbb{R}\to\mathbb{R}$ that is twice continuously differentiable, we define function $a_U : \mathbb{R}\to\mathbb{R}$, where

$$a_U(x) := -\frac{U''(x)}{U'(x)}, \quad x \in \mathbb{R}, \tag{1.6}$$

and call it **Arrow–Pratt measure of absolute risk aversion**. Furthermore, for $x \in \mathbb{R}$, we call $a_U(x)$ Arrow–Pratt coefficient of absolute risk aversion at x.

For smooth U, we can use Taylor expansion around $\mathbb{E}(X)$ to get

$$\rho_U(X) \approx \mathbb{E}[X] - \tfrac{1}{2}a_U(\mathbb{E}(X))\,\mathrm{Var}[X],$$

which allows us to establish an approximate link between mean-variance criterion and expected utility based criterion; coefficient $a_U(\mathbb{E}(X))$ can be used to check mean-variance tradeoff value factor. In particular, we can establish a special class of functions for which $a_U(\cdot)$ is either constant or proportional to the argument.

Example 1.6 (Exponential utility and CARA property)**.** The class of exponential utility functions

$$U_\gamma(x) = \begin{cases} \frac{e^{\gamma x}-1}{\gamma} & \gamma \neq 0, \\ x & \gamma = 0, \end{cases} \tag{1.7}$$

for parameter $\gamma \in \mathbb{R}$ constitutes a unique class of utility functions (defined up to affine transform) which have constant absolute risk aversion (CARA) equal to $-\gamma$; note that U_γ is a utility (in the sense of Definition 1.3) only for $\gamma \leq 0$ as otherwise it is not concave but we can also consider CARA property for convex functions. Indeed, let us assume that utility function $U : \mathbb{R} \to \mathbb{R}$ is twice differentiable and such that $a_U \equiv -\gamma$ for some $\gamma \in \mathbb{R} \setminus \{0\}$. This leads to differential equation $(U')'(x) = \gamma U'(x)$ from which we get $U'(x) = Ce^{\gamma x}$ and consequently $U(x) = Ae^{\gamma x} + B$, for some $A, B \in \mathbb{R}$. Thus, using base affine transform, we can convert U to U_γ and conclude the proof for $\gamma \neq 0$. The case of $\gamma = 0$ is left as a simple exercise. ♦

Example 1.7 (CRRA utility and power functions)**.** The class of utilities

$$U_\gamma(x) = \begin{cases} \frac{1}{\gamma} x^\gamma & \gamma \neq 0, \\ \log x & \gamma = 0, \end{cases}$$

for $\gamma \in \mathbb{R}$ leads to specific class of utilities for which $x a_U(x)$ is constant, that is, they exhibit constant relative risk aversion (CRRA). This family of utility functions, together with CARA family, is, arguably, the most common choice is economic studies. ♦

Many empirical studies show that in reality the investors might not act according to axioms embedded into the expected utility theory. This is, arguably, best illustrated by the Allais Paradox.

Example 1.8 (Allais Paradox)**.** Let us consider two lottery pay-off functions given by

$$X_1 := 2400 \qquad \text{and} \qquad X_2 := \begin{cases} 0 & \text{with prob. } 0.01 \\ 2400 & \text{with prob. } 0.66 \\ 2500 & \text{with prob. } 0.33 \end{cases}$$

Numerous empirical experiments show that most people ($\sim 80\%$) prefer X_1 over X_2 even though we have $\mathbb{E}[X_1] = 2400 < 2409 = \mathbb{E}[X_2]$. On the other hand, if we consider

$$Y_1 := \begin{cases} 0 & \text{with prob. } 0.67 \\ 2500 & \text{with prob. } 0.33 \end{cases} \qquad \text{and} \qquad Y_2 := \begin{cases} 0 & \text{with prob. } 0.66 \\ 2400 & \text{with prob. } 0.34 \end{cases}.$$

then people ($\sim 80\%$) tend to prefer the riskier investment, that is, Y_1, which is in agreement with expectation, as $\mathbb{E}[Y_1] = 825 < 816 = \mathbb{E}[Y_2]$. As pointed out by M. Allais in 1953, this choice empirically violates the independence axiom, which is one of the cornerstones of the von Neumann-Morgenstern utility theory. In a nutshell, given order relation $\succ$, the axiom states that if $X_1 \succ X_2$ and $Y_1 \succ Y_2$ and lotteries (random variables) are independent, then for any $\alpha \in (0,1)$ we should have

$$Z_1^\alpha \succ Z_2^\alpha, \quad \text{where} \quad Z_1^\alpha \sim (\alpha F_{X_1} + (1-\alpha)F_{Y_1}) \text{ and } Z_2^\alpha \sim (\alpha F_{X_2} + (1-\alpha)F_{Y_2}).$$

the axiom essentially states that lottery combination should not affect our preferences. By picking $\alpha = 0.5$ we arrive at $\frac{1}{2}X_1 + \frac{1}{2}Y_1 \succ \frac{1}{2}X_2 + \frac{1}{2}Y_2$ which contradicts the fact that Z_1^α and Z_1^α are both distributed as

$$Z = \begin{cases} 0 & \text{with prob. } 0.335 \\ 2400 & \text{with prob. } 0.5 \\ 2500 & \text{with prob. } 0.165. \end{cases}$$

This indicates that at least 65% (80%[2]) of people do not behave consistently with the independence axiom, which could be seen as empirical evidence against von Neumann-Morgenstern (descriptive) theory; the empirical analysis also show that the choice of utility $U(x) = x$ and expectation value as objective function might be not empirically justifiable. We refer to [FS02a, Example 2.32] and references therein for more details on this paradox. ♦

Example 1.8 shows that a single utility function might be not sufficient to properly quantify the utility of all investors, and one often need to consider various (individual) measures of risk and performance before making an investment choice.[3] W1 ⊣ W2

2 Normative properties of risk and performance measures

In this section we discuss normative properties that a generic mapping

$$\rho\colon L^\infty \to \mathbb{R} \cup \{+\infty\} \tag{2.1}$$

should satisfy to be considered a risk measure or a performance measure. We start with definitions of normative (axiomatic) properties and then show selected interactions between them.

2.1 Definition of normative properties

We start with monotonicity properties that allows us to establish some order, cf. (1.2).

Definition 2.1 (Order properties)**.** We say that $\rho : L^\infty \to \mathbb{R} \cup \{+\infty\}$ is

1) *Monotone increasing (MI)* if $X \geq Y$ imply $\rho(X) \geq \rho(Y)$, for $X, Y \in L^\infty$;

2) *Monotone decreasing (MD)* if $X \geq Y$ imply $\rho(X) \leq \rho(Y)$, for $X, Y \in L^\infty$;

Furthermore, we say that ρ is *monotone* (M) if it is either (MI) or (MD).

In general, a map ρ is used for risk measurement if it is (MD), that is, the more preferable the position, the smaller the risk. On the other hand, for performance measurement, we use (MI) property, as it reflects the standard preference order. Note that, assuming $\rho(\cdot) < \infty$, one can easily shift between (MD) and (MI) by taking the negative values, that is, considering mappings $\tilde{\rho}(X) := -\rho(X)$, $X \in L^\infty$.[4]

While being critical, monotonicity property alone does not establish a meaningful class of risk or performance measures. The other constitutive properties are telling us how the measure behaves under affine transforms.

[3] I personally also recommend M. Lewis's popular science book *The Undoing Project* where this and other behavioral paradoxes are presented in the psychological context, and where the work of D. Kahneman and A. Tversky is outlined.

[4] In the risk measurement literature both (MI) and (MD) are used to denote *risk measures* and sometimes the loss-side convention is also used, that is, we consider maps of the form $\bar{\rho}(X) := \rho(-X)$, for $X \in L^\infty$. For consistency, we define all risk measurement maps using (MD) and not taking negative values.

Definition 2.2 (Affine transform properties)**.** We say that $\rho : L^\infty \to \mathbb{R} \cup \{+\infty\}$ is

1) *Cash additive* (CA), if $\rho(X + m) = \rho(X) - m$, for $X \in L^\infty$ and $m \in \mathbb{R}$;[a]

2) *Translation equivariance* (CA'), if $\rho(X + m) = \rho(X) + m$, for $X \in L^\infty$ and $m \in \mathbb{R}$;[b]

3) *Positively homogeneous* (PH), if $\rho(\lambda X) = \lambda\rho(X)$ for $X \in L^\infty$ and $\lambda > 0$;

4) *Scale invariant* (SI), if $\rho(\lambda X) = \rho(X)$, for $X \in L^\infty$ and $\lambda > 0$;

5) *Sub-scale invariant* (SSI), if $\rho(\lambda X) \leq \rho(X)$, for $X \in L^\infty$ and $\lambda > 1$;[c]

[a] Sometime a name *translation property* is used instead of *cash additivity*, e.g. when ρ does not have a monetary interpretation. Note that (CA) is fit for (MD) maps, to ensure proper ordering

[b] We decided to use (CA') due to the link between this property and (CA) property, e.g. when considering mapping $\tilde{\rho}(X) := -\rho(X)$, $X \in L^\infty$

[c] A specific form of sub-scale invariance is when $\rho(\lambda X) = \frac{1}{\lambda}\rho(X)$, for $\lambda > 0$ and $X \in L^\infty$; this is linked to *inverse positive homogeneity* (IPH) property.

As we later show, the affine properties listed in Definition 2.2, together with the monotonicity properties in Definition 2.1, already characterize several general classes of risk and performance measures. However, to enable the use of tools from convex and functional analysis that guarantee the existence of minimizer in various optimization problems, we must impose certain forms of convexity (or concavity). This is essentially used to control the risk of convex combination of two random variables. For brevity, for $a, b \in \mathbb{R}$, we use standard notation

$$a \vee b := \max\{a, b\} \quad \text{and} \quad a \wedge b := \min\{a, b\},$$

that is handy when dealing with quasi-convex and quasi-concave property.

Definition 2.3 (Linear combination properties)**.** We say that $\rho : L^\infty \to \mathbb{R} \cup \{+\infty\}$ is

1) *Convex* (CV) if $\rho(\lambda X + (1-\lambda)Y) \leq \lambda\rho(X) + (1-\lambda)\rho(Y)$, for $X, Y \in L^\infty$ and $\lambda \in (0, 1)$;

2) *Concave* (CC) if $\rho(\lambda X + (1-\lambda)Y) \geq \lambda\rho(X) + (1-\lambda)\rho(Y)$, for $X, Y \in L^\infty$ and $\lambda \in (0, 1)$;

3) *Quasi-convex* (QCV) if $\rho(\lambda X + (1-\lambda)Y) \leq \rho(X) \vee \rho(Y)$, for $X, Y \in L^\infty$ and $\lambda \in (0, 1)$;

4) *Quasi-concave* (QCC) if $\rho(\lambda X + (1-\lambda)Y) \geq \rho(X) \wedge \rho(Y)$, for $X, Y \in L^\infty$ and $\lambda \in (0, 1)$.

Furthermore, we say that ρ is[a]

5) *Additive* (A) if $\rho(X + Y) = \rho(X) + \rho(Y)$, for any $X, Y \in L^\infty$;

6) *Sub-additive* (SBA) if $\rho(X + Y) \leq \rho(X) + \rho(Y)$,for any $X, Y \in L^\infty$;

7) *Super-additive* (SPA) $\rho(X + Y) \geq \rho(X) + \rho(Y)$, for any $X, Y \in L^\infty$;

8) *Comonotone* (CM) if $\rho(X + Y) = \rho(X) + \rho(Y)$ for any comonotone $X, Y \in L^\infty$.[b]

[a] In a nutshell, additivity based properties could be seen as a combination of convexity/concavity based properties combined with positive homogeneity or sub-scale invariance.

[b] A pair $X, Y \in L^\infty$ is comonotone if $(X(\omega_1) - X(\omega_2))(Y(\omega_1) - Y(\omega_2)) \geq 0$ for $\omega_1, \omega_2 \in \Omega$ (a.s. on $\mathbb{P}\otimes\mathbb{P}$). One can show that X and Y are comonotone if and only if there exists $Z \in L^\infty$ and increasing functions $f, g : \mathbb{R} \to \mathbb{R}$ such that $X = f(Z)$ and $Y = g(Z)$; see [FS02a, Lemma 4.83] for details.

Of course, assuming $\rho(\cdot) < \infty$ and considering map $\tilde{\rho}(X) := -\rho(X)$, $X \in L^\infty$, we can quickly switch from convexity-based properties (CV, QCV) to concavity-based properties (CC,QCC), and vice versa; risk measures are typically convex, while performance measures are typically concave. By combining properties introduced so far, we can already define meaningful classes of risk and performance measure and study their dual representation by analyzing convex (bi)conjugates, and using tools from convex optimization, such as Fenchel–Moreau theorem. Nevertheless, before we focus on that, we need to define supplementary properties and study the interactions between all properties introduced so far.

Supplementary normative properties are used to exclude degenerate cases, allow normalization, ensure certain level of continuity, simplify the narrative, etc. Most of these properties are typically required on spaces larger than L^∞ or are induced by a combination of properties stated in Definition 2.1, Definition 2.2 and Definition 2.3, but we state them here for completeness. We start with basic continuity properties.

Definition 2.4 (Continuity properties)**.** We say that $\rho : L^\infty \to \mathbb{R} \cup \{+\infty\}$ admits

1) *Fatou property* (FP), if $\rho(X) \geq \limsup_{n\to\infty} \rho(X_n)$, for any L^∞-dominated sequence $(X_n)_{n=1}^\infty$, where $X_n \in L^\infty$ and $X_n \to X$ (a.s.), as $n \to \infty$.[a]

2) *Lebesgue property* (LP), if $\rho(X) = \lim_{n\to\infty} \rho(X_n)$, for any L^∞-dominated sequence $(X_n)_{n=1}^\infty$, where $X_n \in L^\infty$ and $X_n \to X$ (a.s.), as $n \to \infty$.

[a] A sequence of random variables $(X_n)_{n=1}^\infty$ is L^∞-dominated, if $\sup_{n\in\mathbb{N}} \|X_n\|_\infty < \infty$.

Of course, (LP) is stronger than (FP). One can also show that, under concavity, (FP) is essentially equivalent to upper semi-continuity on $\sigma(L^\infty, L^1)$, see [Del02] for details. Let us now move to consistency-related properties.

Definition 2.5 (Consistency-related properties)**.** We say that $\rho : L^\infty \to \mathbb{R} \cup \{+\infty\}$ is

1) *Normalized*, if $\rho(0) = 0$;

2) *Proper*, if there exists $X \in L^\infty$ such that $\rho(X) < \infty$;

3) *Finite*, if $\rho(X) < \infty$ for any $X \in L^\infty$.

4) *Arbitrage consistent*, if $\rho(X) = \infty$ holds if and only if $X \geq 0$, for $X \in L^\infty$.

5) *Risk-averse*, if $\rho(X) \leq \rho(\mathbb{E}[X])$, for $X \in L^\infty$;[a]

6) *Risk-seeking*, if $\rho(X) \geq \rho(\mathbb{E}[X])$, for $X \in L^\infty$;

7) *Second-order consistent*, if $\rho(X) \leq \rho(Y)$, for $X, Y \in L^\infty$ such that $X \preceq_2 Y$; [b]

8) *Law invariant* (LI) if $\rho(X) = \rho(Y)$ for $X, Y \in L^\infty$, such that $F_X(t) = F_Y(t)$, $t \in \mathbb{R}$.

[a] Sometimes this property is also called *expectation consistency*, see e.g. [CM09].
[b] $X \preceq_2 Y$ means that $\mathbb{E}[u(X)] \leq \mathbb{E}[u(Y)]$ for any non-decreasing and concave (utility) function $u \colon \mathbb{R} \to \mathbb{R}$.

While normalization is typically assumed for risk measures, arbitrage consistency is used for specific performance measures, which aim is to check arbitrage opportunities on the market. Also, we note that the law invariance is, arguably, the most important property from the practical (statistical) perspective, as it states that the risk or performance of a given position $X \in L^\infty$ depends only on its law, i.e. F_X. For law-invariant map ρ one can show that there exists distribution-based version of ρ, say R, such that $\rho(X) = R(F_X)$, for $X \in L^\infty$. This allow us to treat $\rho(X)$ as a specific distribution characteristic such as mean or variance. Also, it allows us to use statistical tools to estimate the value of $\rho(X)$ given a sample from X, e.g. by using so called plug-in approach, in which we first estimate the distribution of X and then apply the mapping to the estimated distribution of X rather than directly to X.

Finally, it is worthwhile to note that for law-invariance (risk measurement) maps, the continuity properties are typically automatically satisfied on L^∞, see [JST06].

2.2 Selected interactions between normative properties

We already indicated that one does not need to state all of the normative properties listed in Section 2.1 as one can choose certain subset of normative properties that is sufficient to ensure specific form of continuity, etc. Let us now present selected interactions. We start with the basic ones that were already indicated in the previous section.

Proposition 2.6 (Basic interactions)**.** Let $\rho : L^\infty \to \mathbb{R} \cup \{+\infty\}$ be finite. Then
1) ρ is (MI) $\Leftrightarrow$ $-\rho$ is (MD);
2) ρ is (MI) $\Leftrightarrow$ $\bar{\rho}$ is (MD), where $\bar{\rho}(X) = \rho(-X)$, $X \in L^\infty$;
3) ρ is (CV) $\Rightarrow$ ρ is (QCV);
4) ρ is (CC) $\Rightarrow$ ρ is (QCC);
5) ρ is (QCC) $\Leftrightarrow$ $-\rho$ is (QCV);
6) ρ is (CC) $\Leftrightarrow$ $-\rho$ is (CV).

The proof of Proposition 2.6 is left as a simple exercise. Let us now study the typical interactions between specific sets of properties.

Proposition 2.7 (Selected normative implications)**.** Let $\rho : L^\infty \to \mathbb{R} \cup \{+\infty\}$ be finite. Then
1) ρ is (PH) $\Rightarrow$ ρ is (N);
2) ρ is (PH) and (SBA) $\Rightarrow$ ρ is (CV);
3) ρ is (PH) and (CV) $\Rightarrow$ ρ is (SBA);
4) ρ is (QCV) and (CA) $\Rightarrow$ ρ is (CV);
5) ρ is (QCC) and (CA') $\Rightarrow$ ρ is (CC);

Proof. Let $\rho : L^\infty \to \mathbb{R} \cup \{+\infty\}$ be finite. Let $X, Y \in L^\infty$ and $\lambda > 0$.

1) $\rho(0) = 2\rho(0) - \rho(0) \overset{(PH)}{=} \rho(2 \cdot 0) - \rho(0) = 0$

2) $\rho(\lambda X + (1-\lambda)Y) \overset{(SBA)}{\leq} \rho(\lambda X) + \rho((1-\lambda)Y) \overset{(PH)}{=} \lambda\rho(X) + (1-\lambda)\rho(Y)$.

3) $\frac{1}{2}\rho(X+Y) \overset{(PH)}{=} \rho(\frac{1}{2}X + \frac{1}{2}Y) \overset{(CV)}{\leq} \frac{1}{2}\rho(X) + \frac{1}{2}\rho(Y)$, for $\lambda = \frac{1}{2}$, so (SBA) follows.

4)
$$\begin{aligned}
\rho(\lambda X + (1-\lambda)Y) &= \rho(\lambda X + \lambda\rho(X) - \lambda\rho(X) + (1-\lambda)Y + (1-\lambda)\rho(Y) - (1-\lambda)\rho(Y)) \\
&\overset{(CA)}{=} \lambda\rho(X) + (1-\lambda)\rho(Y) + \rho\big(\lambda(X+\rho(X)) + (1-\lambda)(Y+\rho(Y))\big) \\
&\overset{(QCV)}{\leq} \lambda\rho(X) + (1-\lambda)\rho(Y) + \max\{\rho(X+\rho(X)), \rho(Y+\rho(Y))\} \\
&\overset{(CA)}{=} \lambda\rho(X) + (1-\lambda)\rho(Y) + 0
\end{aligned}$$

5) The proof is as in 4); we consider $\tilde{\rho}(\cdot) := -\rho(\cdot)$ and follow the same logic. □

2.3 Normative properties of certainty equivalents and entropic utility measure

In Definition 1.4 we have establishes an important class of performance measures (that can be transferred to risk measures by taking minus sign) called certainty equivalent criterions. Let us now quickly check what normative properties are satisfied by maps $\rho_U \colon L^\infty \to \mathbb{R}$ of the form

$$\rho_U(X) := U^{-1}\mathbb{E}[U(X)],$$

where $U \colon \mathbb{R} \to \mathbb{R}$ is an utility function. Note that those maps are the cornerstone of the expected utility theory, so one expects they satisfy the core properties linked to monotonicity or concavity which is indeed the case.

Proposition 2.8 (Base normative properties of certainty equivalents)**.** Let $U : \mathbb{R} \to \mathbb{R}$ be a (concave and increasing) utility function and let $\rho_U \colon L^\infty \to \mathbb{R}$ be the certainty equivalent criterion based on U. Then
1) ρ_U in monotone increasing (MI);
2) ρ_U in law-invariant (LI);
3) ρ_U in normalized (N);
4) ρ_U is quasi-concave (QCC).
5) ρ_U is risk-averse.

We leave the proof of Proposition 2.8 as a simple exercise. A natural question is if we can get some additional properties by considering specific utility functions. One can say that, arguably, exponential utility is the one which is used most commonly in risk-sensitive stochastic optimization.

Definition 2.9 (Entropic utility measure)**.** For any risk-aversion parameter $\gamma \in \mathbb{R}$, we define the mapping $\mathrm{Ent}_\gamma \colon L^\infty \to \mathbb{R}$ given by

$$\mathrm{Ent}_\gamma(X) := \begin{cases} \frac{1}{\gamma} \ln \mathbb{E}[e^{\gamma X}] & \gamma \neq 0 \\ \mathbb{E}[X] & \gamma = 0 \end{cases},$$

and call it *entropic utility measure* for risk-aversion γ.

Let us now state selected properties for entropic utility measure.

Proposition 2.10 (Properties of entropic utility)**.** Let $\mathrm{Ent}_\gamma \colon L^\infty \to \mathbb{R}$ be the entropic utility for $\gamma \in \mathbb{R}$. Then
1) Ent_γ is the certainty equivalent, i.e. $\mathrm{Ent}_\gamma(X) = U_\gamma^{-1}\mathbb{E}[U_\gamma(X)]$, for U_γ defined in (1.7).
2) For any $\gamma \in \mathbb{R}$, Ent_γ is (MI), (LI), (N), and (CA').
3) For any $\gamma \leq 0$, Ent_γ is (CC) and risk-averse.
4) For any $\gamma \geq 0$, Ent_γ is (CV) and risk-seeking.
5) For any $X \in L^\infty$ and $\gamma_1, \gamma_2 \in \mathbb{R}$ such that $\gamma_1 > \gamma_2$, we get $\mathrm{Ent}_{\gamma_1}(X) \geq \mathrm{Ent}_{\gamma_2}(X)$.
6) For any $X \in L^\infty$ the mapping $\gamma \to \mathrm{Ent}_\gamma(X)$ is continuous with limits

$$\lim_{\gamma\to-\infty} \mathrm{Ent}_\gamma(X) = \operatorname{ess\,inf}(X) \quad \text{and} \quad \lim_{\gamma\to\infty} \mathrm{Ent}_\gamma(X) = \operatorname{ess\,sup}(X).$$

Proof. Fix $\gamma < 0$; the proof for $\gamma \geq 0$ is analogous.

1) Since $U_\gamma(x) = \frac{e^{\gamma x}-1}{\gamma}$, we get $U_\gamma^{-1}(x) = \frac{\ln(1+\gamma x)}{\gamma}$, and consequently

$$U_\gamma^{-1}\mathbb{E}[U_\gamma(X)] = \frac{1}{\gamma} \ln\left(1 + \gamma \mathbb{E}\left[\frac{e^{\gamma X}-1}{\gamma}\right]\right) = \mathrm{Ent}_\gamma(X), \quad X \in L^\infty.$$

2) (MI), (LI) and (N) is due to Proposition 2.8. To show (CA'), fix $X \in L^\infty$ and $m \in \mathbb{R}$. Directly from the definition and logarithm function properties we get

$$\mathrm{Ent}_\gamma(X+m) = \tfrac{1}{\gamma}\ln \mathbb{E}[e^{\gamma(X+m)}] = \tfrac{1}{\gamma}\ln e^{\gamma m}\mathbb{E}[e^{\gamma X}] = \tfrac{1}{\gamma}\ln \mathbb{E}[e^{\gamma X}] + m = \mathrm{Ent}_\gamma(X) + m.$$

3) (CC) follows from (QCC), induced by 1), and (CA'), see 4) in Proposition 2.8 and 5) in Proposition 2.7.

4) For $\gamma \geq 0$ the proof is analogous to 3).
5) Fix $X \in L^\infty$. For simplicity let us consider only the risk-averse case; other cases are similar. Fix $0 > \gamma_1 > \gamma_2$. By convexity of the exponential function and Jensen's inequality, we have

$$\mathbb{E}[e^{\gamma_2 X}] = \mathbb{E}[(e^{\gamma_1 X})^{\gamma_2/\gamma_1}] \geq \left(\mathbb{E}[e^{\gamma_1 X}]\right)^{\gamma_2/\gamma_1}.$$

by taking logarithm on both sides and dividing by $\gamma_2 < 0$ we conclude the proof.

6) For $\gamma \neq 0$ the proof of continuity is straightforward and is due to the continuity of the utility function. For $\gamma = 0$, the proof follows from L'Hôpital's rule. Indeed, we get

$$\lim_{\gamma\to 0}\mathrm{Ent}_\gamma(X) = \lim_{\gamma\to 0}\frac{\log \mathbb{E}[e^{\gamma X}]}{\gamma} = \lim_{\gamma\to 0}\frac{\mathbb{E}[Xe^{\gamma X}]/\mathbb{E}[e^{\gamma X}]}{1} = \frac{\mathbb{E}[X\cdot 1]/1}{1} = \mathbb{E}[X],$$

as $\frac{d}{d\gamma}\log \mathbb{E}[e^{\gamma X}] = \frac{\frac{d}{d\gamma}\mathbb{E}[e^{\gamma X}]}{\mathbb{E}[e^{\gamma X}]} = \frac{\mathbb{E}[\frac{d}{d\gamma}e^{\gamma X}]}{\mathbb{E}[e^{\gamma X}]} = \frac{\mathbb{E}[Xe^{\gamma X}]}{\mathbb{E}[e^{\gamma X}]}$.
Let us now show the proof for the limit case $\gamma \to -\infty$; the proof for $\gamma \to \infty$ is similar. Fix $X \in L^\infty$. W2 ⊣ W3
For $Y := X - \operatorname{ess\,inf}(X)$ we get $Y \geq 0$ and $\operatorname{ess\,inf}(Y) = 0$. If $Y \equiv 0$, the proof is straightforward. Otherwise, for any $\epsilon > 0$ there exists set $A \in \mathcal{F}$ such that $\mathbb{P}[A] > 0$ and $Y \leq \epsilon 1_A + M 1_{A'}$, where $M := \operatorname{ess\,sup} Y$. From (MI), for any $\gamma < 0$, we get

$$\mathrm{Ent}_\gamma(Y) \leq \mathrm{Ent}_\gamma(\epsilon 1_A + M 1_{A'}) = \tfrac{1}{\gamma}\ln\left(\mathbb{P}[A]e^{\gamma\epsilon} + \mathbb{P}[A']e^{\gamma M}\right) \leq \tfrac{1}{\gamma}\ln\left(\mathbb{P}[A]e^{\gamma\epsilon}\right) = \tfrac{1}{\gamma}\ln\left(\mathbb{P}[A]\right) + \epsilon.$$

Consequently, by letting $\gamma \to -\infty$ we get $\mathrm{Ent}_\gamma(Y) \leq \epsilon$, for any $\epsilon > 0$. Combining this with the fact that $Y \geq 0$, which implies $\mathrm{Ent}_\gamma(Y) \geq 0$, we conclude that $\lim_{\gamma\to-\infty}\mathrm{Ent}_\gamma(Y) = 0$ which concludes the proof due to (CA) property. □

The entropic utility satisfies many more interesting and unique properties among certainty equivalents and even risk measures (where we take negative of it), which makes it so popular is stochastic optimization. We will discuss those properties at a later stage.

3 Important families of risk measures

In this section, we focus on specific type of maps $\rho\colon L^\infty \to \mathbb{R}\cup\{+\infty\}$ which purpose is to measure the risk of a generic financial P&L position X. We outline selected concepts from the risk measure theory and show dual representation of coherent and convex risk measures; we refer to [FS02a, Section 4] for a more comprehensive introduction to risk measures.

3.1 Monetary risk measures

To ensure a generic mapping is appropriate for risk measurement, it is necessary to apply two normative properties: monotonicity (MD) and cash additivity (CA). Additionally, normalization

(N) is applied to enhance the interpretability of the risk.[5] Recall that (MD) monotonicity convention is measuring the size of loss, that is, the more risky the position X, the bigger the value of $\rho(X)$. We start with a generic definition of a *monetary risk measure*.

Definition 3.1 (Monetary risk measure)**.** We say that $\rho : L^\infty \to \mathbb{R}$ is a *monetary risk measure* if it is monotone decreasing (MD), cash additive (CA), and normalized (N).[a]

[a]Note that we restricted the mapping range to finite maps, but finiteness is in fact induced by (MD) and (CA) on L^∞, as we get $-\|X\|_\infty \leq \rho(X) \leq \|X\|_\infty$.

The word *monetary* in Definition 3.1 reflects the fact that the units in which $\rho(X)$ is measured should correspond to the units in which X is measured due to cash-additivity properties. Of course, the generic definition of *monetary risk measure* is rather general and it is hard to obtain characterization property for it or deduce some other (non-trivial) normative properties. That saying, note that any monetary risk measures also satisfy Lipschitz continuity with respect to the supremum norm, that is, for $X, Y \in L^\infty$, we have

$$|\rho(X) - \rho(Y)| \leq \|X - Y\|,$$ [6]

so (MD) together with (CA) is already leading to a family with a certain degree of smoothness. In most practical examples, we will be dealing with maps that are additionally law-invariant (LI) but for now we decided to not include it directly in the definition. Let us now present a notable example of monetary risk measure, and then impose additional convexity-based normative properties to obtain more specific families of risk measures.

3.2 Regulatory risk measures: value-at-risk and expected shortfall

Arguably, the best recognized and most commonly used monetary risk measure is *value-at-risk* (VaR). While being relatively simple from a mathematical viewpoint, as it corresponds to the negative of a left-continuous quantile function, its found its way in multiple economic areas. It can be used to measure risks of P&L losses, exposures, financial investments, or discounted payoffs and is a key metric that can be used in market risk, counter-party credit risk, insurance risk, or for setting up margin requirements. For instance, the VaR is the main reference metric that is used in the banking sector to report Pillar 1 Market Risk capital within Basel II *internal model approach* (IMA) framework. Noting that VaR metric is currently being substituted by *expected shortfall* as the main IMA capital reporting metric as part of the Basel III (FRTB) reforms, we emphasize that VaR still remains an integral part in the financial regulatory framework, e.g. due to its maintained use in regulatory backtesting exercise.[7] For a comprehensive economical, statistical, financial, and mathematical exposure of value-at-risk see [Jor07], [Ale09], and references therein.

[5]In the risk measure literature one can find objects referred as *risk measures* in which those properties are relaxed (e.g. cash-additivity is substituted with cash-sub-additivity) but in this lecture, we only consider the base case.

[6]We have $X \leq Y + \|X - Y\|$ which implies $\rho(X) \leq \rho(Y + \|X - Y\|) = \rho(Y) - \|X - Y\|$, and we can switch X with Y.

[7]For Basel II recommendations as well as exemplary local regulatory implementations, see e.g. *Market Risk – The Internal Models Approach* section in *BCBS International Convergence of Capital Measurement and Capital Standards* or *European Union Regulation No 575/2013 – Capital Requirements Regulation 2*. More details on Basel III (FRTB) framework could be found e.g. in MAR32 within *Minimum capital requirements for market risk* consolidated Basel Framework document issued by BCBS.

Definition 3.2 (Value-at-risk)**.** For any $\alpha \in (0,1)$, we define the mapping $\mathrm{VaR}_\alpha \colon L^\infty \to \mathbb{R}$ given by

$$\mathrm{VaR}_\alpha(X) := -\inf\{x \in \mathbb{R} \mid \mathbb{P}[X \leq x] > \alpha\}, \quad X \in L^\infty, \tag{3.1}$$

and call it *value-at-risk* at (confidence/threshold) level α.[a]

[a]There are two primary conventions for expressing confidence levels in VaR: the left-tail convention, which is used in this lecture note and is prevalent in the risk measurement literature, and the right-tail convention, which presents thresholds like $1-\alpha$ (such as 99%, 97.5%, 95%). The latter is more commonly used in risk management and regulatory contexts, where the right tail denotes losses.

The main reason behind value-at-risk popularity is linked to its economic and financial interpretability: it can be viewed as the currency amount an institution can lose over a certain period of time with a given probability threshold $\alpha \in (0,1)$. Alternatively, VaR_α can be interpreted as the amount of extra capital that a company needs in order to reduce the probability of going bankrupt to α. To see this, it is sufficient to use property $\mathrm{VaR}_\alpha(X + \mathrm{VaR}_\alpha(X)) = 0$ and show that VaR_α admits alternative representation

$$\mathrm{VaR}_\alpha(X) = \inf\{x \in \mathbb{R} \mid \mathbb{P}[X + x < 0] \leq \alpha\};$$

the proof is left as a simple exercise. We refer to Figure 1 for an illustrative of VaR financial interpretability.

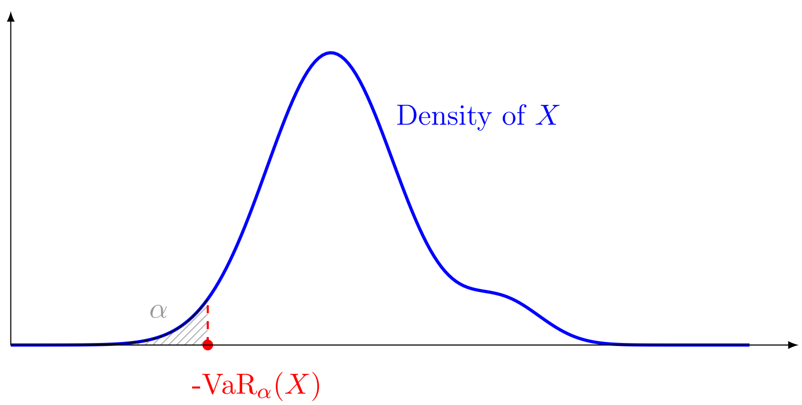

Figure 1: *The plot presents illustration of* $\mathrm{VaR}_\alpha(X)$ *for small value of* $\alpha \in (0,1)$ *and exemplary (absolutely continuous) random variable* X*. The shaded area under the graph is equal to* α*, so that* X *has value bigger than* -$\mathrm{VaR}_\alpha(X)$ *with probability* $1-\alpha$*. In other words, after adding the capital reserve* $\mathrm{VaR}_\alpha(X)$*, the secured position* $X + \mathrm{VaR}_\alpha(X)$ *is positive with probability* $1-\alpha$*.*

In most practical applications, the level $\alpha \in (0,1)$ is set to 1%, 2.5%, or 5%. For example, in the Basel II Pillar 1 Market Risk IMA reporting, the main reference metric is $\mathrm{VaR}_{1\%}$ calculated for a 10-day holding period, which means that the underlying random variable X in $\mathrm{VaR}_{1\%}(X)$ encodes the risk of a future 10-day P&L move linked to the financial institution position. The values 2.5% or 5% could be used e.g. for internal management or as part of IMA model validation tools.[8] Notably, for backtesting purposes, the 1-day holding period is typically used, and many institutions calculate the 10-day $\mathrm{VaR}_{1\%}$ by scaling a 1-day $\mathrm{VaR}_{1\%}$.[9] Apart from 1%, 2.5%, and 5% levels and 1-day

[8]See e.g. *Market Risk Standards* (C164/2018) in *CBUAE Rulebook*.
[9]See e.g. *EBA 2024 Report on the Market Risk Benchmarking Exercise*.

and 10-day holding periods, one can find also other setups, e.g. when Economic Capital within Internal Capital Adequacy Assessment Process (ICAAP) is reported, or when Incremental Risk Charge (IRC) is being computed.[10]

Remark 3.3 (VaR and quantile inverse functions)**.** By investigating (3.1) we can see that VaR_α defined here is a negative of the left-continuous quantile function of X, that is, $\mathrm{VaR}_\alpha(X) = -F_X^{\rightarrow}(\alpha)$, where $F_X^{\rightarrow}(t) := \inf\{x \in \mathbb{R} \mid F_X(x) > t\}$; this definition is consistent with [FS02a, Definition 4.40]. A common alternative is the right-continuous quantile $F_X^{-1}(t) := \inf\{x \in \mathbb{R} \mid F_X(x) \geq t\}$, choice which leads to a more conservative outcome (that might be desirable from a regulatory perspective). $\lozenge$

Now, let us summarize normative properties that are satisfied by VaR.

Proposition 3.4 (Properties of VaR)**.** Let $\mathrm{VaR}_\alpha \colon L^\infty \to \mathbb{R}$ denote value-at-risk at confidence level $\alpha \in (0,1)$. Then
1) VaR_α is a monetary risk measure, that is, it is (MD), (CA), and (N).
2) VaR_α is (PH) and (LI),
3) For any $X \in L^\infty$, the mapping $\alpha \to \mathrm{VaR}_\alpha(X)$ is decreasing

We leave the proof of Proposition 3.4 as a simple exercise. Notably, VaR does not structurally admit risk-aversion property (see Definition 2.4) and does not admit convexity/sub-additivity. Lack of sub-additivity is often used as a main critique point of VaR and is often brought up as a main reason behind VaR to *expected shortfall* migration in finance. To better understand this, let us now illustrate lack of VaR sub-additivity on simple examples in which this is discussed in the context of diversification benefits.

Example 3.5 (VaR and diversification benefit – first example)**.** Let us consider 10% interest loan of \$1000 given to a client that can default with probability $p = 0.7\%$. In this case the future P&L can be described by

$$X := \begin{cases} -\$1000 & \text{with pr. } \ \ 0.7\%; \\ \ \ \$100 & \text{with pr. } 99.3\%; \end{cases} \quad \text{so that} \quad \mathrm{VaR}_{1\%}(X) = -\$100,$$

which essentially means that X has negative risk, so that the position is acceptable without any (risk) capital add-on. Now, let us assume we issue two 10% interest loans, each equal to \$500, to two independent clients, each of them having probability of default of $p = 0.7\%$; default events are assumed to be independent. Then, the future P&L can be described by

$$Y = \begin{cases} -\$1000 & \text{with pr. } \ \ 0.0049\%; \\ -\$450 & \text{with pr. } \ \ 1.3902\%; \\ \ \ \$100 & \text{with pr. } 98.6049\%; \end{cases} \quad \text{so that} \quad \mathrm{VaR}_{1\%}(Y) = \$450,$$

which means that need to secure our position with \$450 to make it acceptable. While the rational decision would be to diversify our portfolio and issue multiple independent loans, from the risk perspective the position X is safer than Y which is questionable from the economic perspective. $\blacklozenge$

Example 3.6 (VaR and diversification benefit – second example)**.** Consider two insurance companies. The first company issued \$1,000,000 insurance policies against earthquakes to 100 homes

[10]Economic Capital aims to cover economic effects of risk-taking activities in a winder sense that done within Pillar 1, see e.g. *BCBS Range of practices and issues in economic capital frameworks* or *ECB Guide to ICAAP*, while Incremental Risk Charge quantifies the risk of potential losses from credit default and migration in the trading book.

within the same city, each carrying a premium of \$50,000. The second company did the same, but issued if for 100 homes spread across different cities worldwide. The likelihood of an earthquake occurring is assumed to be 4% in every city, independently of other cities. Let us use X to denote the P&L vector of the first company, and Y to denote the P&L for the second company. One can show that under these conditions, we get $\mathrm{VaR}_{5\%}(X) = -\$5,000,000$ and $\mathrm{VaR}_{5\%}(Y) \approx \$2,000,000$. This shows that $\mathrm{VaR}_{5\%}$ marks the investment strategy of the first company as safer than the investment strategy of the second company, which is not consistent with a typical (rational) insurance company strategy, in which diversification should have a positive effect on risk. ♦

Using more mathematical setup, in Example 3.6 and Example 3.6 we have proved that VaR is not sub-additivity (SBA) by considering an i.i.d. sequences $(X_i)_{i=1}^k$, for some $k \in \mathbb{N}$, confronting risk of $X := kX_i$ with $Y := \sum_{i=1}^k X_i$, and the showing that $\mathrm{VaR}_\alpha(X) < \mathrm{VaR}_\alpha(Y)$ which leads to a direct contradiction with (SBA) property due to positive homogeneity of VaR. On the other hand, one can show that in many situations the (SBA) property is satisfied by VaR, especially when it is used reasonably.

Example 3.7 (VaR under normality)**.** VaR can be easily extended from L^∞ to L^0 space and it is finite for a generic random variable. For example, assume that $X \sim N(\mu, \sigma)$. Then, using (PH) and (CA), we have

$$\mathrm{VaR}_\alpha(X) = -\mu + \sigma\, \mathrm{VaR}_\alpha(Z) = -\left(\mu + \sigma \Phi^{-1}(\alpha)\right),$$

where $Z \sim N(0, 1)$ and Φ denoted standard normal CDF. Notably, for a jointly normal vector (X,Y) with margins $X \sim N(\mu_1, \sigma_1)$ and $Y \sim N(\mu_2, \sigma_2)$ we get the sub-additivity of VaR since

$$\begin{aligned}
\mathrm{VaR}_\alpha(X+Y) &= -(\mu_1+\mu_2) - \sqrt{\sigma_1^2+\sigma_2^2+2\mathrm{Cov}(X,Y)}\Phi^{-1}(\alpha)\\
&\le -(\mu_1+\mu_2) - \sqrt{\sigma_1^2+\sigma_2^2+2\sigma_1\sigma_2}\Phi^{-1}(\alpha)\\
&= -(\mu_1+\mu_2) - (\sigma_1+\sigma_2)\Phi^{-1}(\alpha)\\
&= \mathrm{VaR}_\alpha(X) + \mathrm{VaR}_\alpha(Y),
\end{aligned}$$

which actually proves (SBA) when combinations of jointly normal vectors are considered, see [MFE10, Section 6.1.2] for further discussion on this topic. ♦

That saying, to structurally avoid non-rational situations as those presented in Example 3.6 and Example 3.6 we can force the risk measure to satisfy the (SBA) normative property which leads to a family of coherent risk measures. Before we do that, let us define another important risk measure which admits (SBA) and is tightly linked to VaR, that is, *expected shortfall* (ES).

Definition 3.8 (Expected shortfall)**.** For any $\alpha \in (0, 1)$, we define the mapping $\mathrm{ES}_\alpha \colon L^\infty \to \mathbb{R}$ given by

$$\mathrm{ES}_\alpha(X) := \frac{1}{\alpha}\int_0^\alpha \mathrm{VaR}_t(X)\,\mathrm{d}t, \quad X \in L^\infty, \tag{3.2}$$

and call it *expected shortfall* at (confidence/threshold) level α.[a]

[a] In the literature, we can find many other names for this risk measure. The most common alternative names are *average balue-at-risk* and *conditional value-at-risk*. Also, note that we follow the same confidence level labeling convention as for VaR, that is, we use 1% instead of 99%, etc.

ES is often interpreted as an average of VaRs beyond a certain threshold or the negative of the expected loss exceeding the α-quantile. The second interpretation is based on the observation that W3 ⊣ W4 for continuous random variables and $\alpha \in (0,1)$, ES_α can be expressed as

$$\mathrm{ES}_\alpha(X) = \mathbb{E}[-X \mid X \leq -\mathrm{VaR}_\alpha(X)], \quad X \in L^\infty, \tag{3.3}$$

see [MFE10, Lemma 2.16] for the proof. In this case, it also has a clear graphical illustration being an integral over the area presented in Figure 1. As we already said, as part of Basel III (FRTB) reforms, VaR at level 1% has been replaced by ES at level 2.5% which is now scheduled to become the main reference metric for IMA Pillar 1 Market Risk capital reporting. Let us now discuss selected properties of ES.

Proposition 3.9 (Properties of ES)**.** Let $\mathrm{ES}_\alpha \colon L^\infty \to \mathbb{R}$ denote expected shortfall at confidence level $\alpha \in (0,1)$. Then
1) ES_α is a monetary risk measure, that is, it is (MD), (CA), and (N).
2) ES_α is (PH), (LI), and risk-averse.
3) For any $X \in L^\infty$, the mapping $\alpha \to \mathrm{ES}_\alpha(X)$ is decreasing.
4) ES_α is (SBA) and (CV),
5) ES_α admits *Rockafellar–Uryasev identity*, that is, for any $X \in L^\infty$, we have

$$\mathrm{ES}_\alpha(X) = \inf_{m\in\mathbb{R}} \left\{ m + \tfrac{1}{\alpha}\,\mathbb{E}\big[(X+m)^-\big] \right\}. \tag{3.4}$$

Proof. For brevity, we skip detailed proof and provide a high-level (sketch) comment with references. The proof of 1), 2), and 3) is a direct implication of Proposition 3.9 combined with integral operator properties; the risk-averse property can be seen by using 3) and noting that $\lim_{\alpha\to 1} \mathrm{ES}_\alpha(X) = -\mathbb{E}[X]$, see [FS02a, p. 179] for details. In the proof of 4) it is sufficient to prove (SBA), as (CV) is implied by (SBA) and (CA) due to Proposition 2.7. The proof of (SBA) and 5) can be found in [EW15], in which seven different proofs of (SBA) for ES are presented. In fact, given 5), the proof of (SBA) is straightforward since, for $X, Y \in L^\infty$, using basic property $(x+y)^- \leq x^- + y^-$ for $x, y, \in \mathbb{R}$, we get

$$\begin{aligned}
\mathrm{ES}_\alpha(X+Y) &= \inf_{m\in\mathbb{R}} \left\{ m + \tfrac{1}{\alpha}\,\mathbb{E}\big[(X+Y+m)^-\big] \right\}\\
&= \inf_{m_1,m_2\in\mathbb{R}} \left\{ m_1 + m_2 + \tfrac{1}{\alpha}\,\mathbb{E}\big[(X+m_1+Y+m_2)^-\big] \right\}\\
&\leq \inf_{m_1,m_2\in\mathbb{R}} \left\{ m_1 + m_2 + \tfrac{1}{\alpha}\,\mathbb{E}\big[(X+m_1)^- + (Y+m_2)^-\big] \right\}\\
&\leq \inf_{m_1\in\mathbb{R}} \left\{ m_1 + \tfrac{1}{\alpha}\,\mathbb{E}\big[(X+m_1)^-\big] \right\} + \inf_{m_2\in\mathbb{R}} \left\{ m_2 + \tfrac{1}{\alpha}\,\mathbb{E}\big[(Y+m_2)^-\big] \right\}\\
&= \mathrm{ES}_\alpha(X) + \mathrm{ES}_\alpha(Y);
\end{aligned}$$

we refer to [EW15, Section 3.2] for details and proof of 5). □

One can also obtain a simple explicit formula for ES under normality.

Example 3.10 (Expected shortfall under normality)**.** Expected shortfall can be easily extended from L^∞ to L^1 space as it is finite for a generic integrable random variable. For example, assuming that $X \sim N(\mu, \sigma)$, we get

$$\mathrm{ES}_\alpha(X) = -\mu + \sigma \frac{\phi(\Phi^{-1}(\alpha))}{\alpha}, \tag{3.5}$$

and we see that, as in VaR case, ES is the affine-transformed version of ES for standard normal variable; see [MFE10, Example 2.18] for details and the proof. Also, it is useful to note that

$$\frac{\mathrm{VaR}_{1\%}(X) - ES_{2.5\%}(X)}{\sigma} \approx 1.1\%,$$

which shows that $\mathrm{VaR}_{1\%}(X)$ is very close to $\mathrm{ES}_{2.5\%}(X)$ under normality, which might explain why the regulator chose to replace VaR 1% confidence level with ES 2.5% confidence level as part of Basel III (FRTB) reforms. ♦

One of the key arguments used in the discussion why ES is better than VaR is linked to the fact that ES is convex (CV) and it is in fact a *coherent risk measure*, this is, ES satisfies (SBA) property. In fact, one can show that ES is the smallest law-invariant risk measure which dominates VaR and is convex; this means that ES is best conservative approximation of VaR if we additionally require convexity.

Proposition 3.11. Let $\alpha \in (0,1)$. Then, ES_α is the smallest (LI) and (CV) monetary risk measure that dominates VaR_α, that is, it satisfies property $ES_\alpha(X) \geq \mathrm{VaR}_\alpha(X)$, for $X \in L^\infty$.

For the proof of Proposition 3.11 we refer to [FS02a, Theorem 4.61] and Theorem 3.16 that is formulated later. Let us now focus on the generic family of coherent risk measures and explain the placement of VaR and ES in this context.

3.3 Coherent risk measures

The seminal paper [Art+99] was the first work in which the authors proposed a modern normative (axiomatic) approach to risk measurement in the financial context; see also [Wan96]. They formulated financial properties which led to the definition of a *coherent risk measure*, see also [CGM01]. For completeness, let us recall all normative properties that contribute to the definition of this class of maps.

Definition 3.12 (Coherent risk measure)**.** We say that $\rho : L^\infty \to \mathbb{R}$ is a *coherent risk measure* (CRM) if it is

1) *Monotone decreasing* (MD) if $X \geq Y$ imply $\rho(X) \leq \rho(Y)$, for $X, Y \in L^\infty$;

2) *Cash additive* (CA), if $\rho(X + m) = \rho(X) - m$, for $X \in L^\infty$ and $m \in \mathbb{R}$;

3) *Positively homogeneous* (PH), if $\rho(\lambda X) = \lambda\rho(X)$ for $X \in L^\infty$ and $\lambda > 0$;

4) *Sub-additive* (SBA) if $\rho(X + Y) \leq \rho(X) + \rho(Y)$,for any $X, Y \in L^\infty$.

In other words, ρ is CRM if it is a monetary risk measure that is additionally (PH) and (SBA).

Note that positive homogeneity (PH) induces normalization (N), and positive homogeneity (PH) together with sub-additivity (SBA) induces convexity (CV); see Section 2.2. The financial interpretation of the axioms is straightforward. For completeness, let us briefly recall it: (MD) - better performing portfolio bears less risk; (CA) - adding a cash of $\$m$ to the position decreases the risk by the same amount; (PH) - the risk of a re-scaled portfolio is re-scaled; (SBA) - diversification reduces the risk.

As have already shown, expected shortfall is a prominent example of a coherent risk measure. For completeness, let us now show two additional example.

Example 3.13 (Worst-case risk measure is CRM)**.** Let us consider the *worst-case risk measure* given by

$$\rho_{\min}(X) := -\operatorname{ess\,inf} X, \quad X \in L^\infty. \tag{3.6}$$

One can easily check that (MD), (CA), and (PH) properties hold for this measure. Furthermore, the (SBA) property is also true due to base properties of the essential infimum. Consequently, $\rho_{\min}$ given in (3.6) is a coherent risk measure. ♦

Example 3.14 (Expected value based risk measure is CRM)**.** Let us consider the *expected value* induced risk measure given by

$$\rho_{\text{mean}}(X) := -\mathbb{E}[X], \quad X \in L^\infty. \tag{3.7}$$

One can easily check that (MD), (CA), and (PH) properties hold for this measure. Furthermore, the expectation is additive, so it also does satisfy (SBA). Consequently, ρ_{mean} given in (3.7) is a coherent risk measure. ♦

Now, let us provide a dual representation result, which is one of the most important results in the theory of coherent risk measures.

Theorem 3.15 (Robust representation of CRM)**.** Assume that $\rho : L^\infty \to \mathbb{R}$ satisfies (FP).[a] Then, ρ is a coherent risk measure if and only if there exists a (closed convex) set of probability measures $\mathcal{Q}$, all absolutely continuous with respect to $\mathbb{P}$, such that

$$\rho(X) = \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_{\mathbb{Q}}[-X], \quad X \in L^\infty. \tag{3.8}$$

[a] For CRMs, the Fatou property (FP) can be restated as: if $0 \leq X \leq 1$ and $X_n \downarrow 0$, then $\rho(X_n) \downarrow 0$. Recall this is a technical assumption that impose specific form of 'continuity' on ρ.

Proof. (⇐) Assume that ρ is defined by (3.8) for some set of probability measures $\mathcal{Q}$. We show that ρ satisfies (MD), (CA), (PH), and (SBA), as stated in Definition 3.12.

(MD): Fix $X, Y \in L^\infty$ such that $X \geq Y$. By the monotonicity of expected value, we have $\mathbb{E}_{\mathbb{Q}}[X] \geq \mathbb{E}_{\mathbb{Q}}[Y]$, and consequently $\mathbb{E}_{\mathbb{Q}}[-X] \leq \mathbb{E}_{\mathbb{Q}}[-Y]$, for any $\mathbb{Q} \in \mathcal{Q}$. By taking supremum of both parts in the last inequality we obtain

$$\sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_{\mathbb{Q}}[-X] \geq \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_{\mathbb{Q}}[-Y],$$

which concludes the proof of (MD).

(CA): Fix $X \in L^\infty$ and $m \in \mathbb{R}$. Since m is a constant, we have

$$\rho(X+m) = \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_{\mathbb{Q}}[-(X+m)] = \sup_{\mathbb{Q}\in\mathcal{Q}} (\mathbb{E}_{\mathbb{Q}}[-X] - m) = \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_{\mathbb{Q}}[-X] - m = \rho(X) - m.$$

(PH): Fix $X \in L^\infty$ and $\lambda > 0xs$. Since λ is a strictly positive constant, we have

$$\rho(\lambda X) = \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_{\mathbb{Q}}[-\lambda X] = \sup_{\mathbb{Q}\in\mathcal{Q}} \lambda \mathbb{E}_{\mathbb{Q}}[-X] = \lambda \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_{\mathbb{Q}}[-X] = \lambda\rho(X).$$

(SBA): Fix $X, Y \in L^\infty$. Recalling that supremum operator is sub-additive, we get

$$\rho(X+Y) = \sup_{\mathbb{Q}\in\mathcal{Q}} (\mathbb{E}_{\mathbb{Q}}[-X] + \mathbb{E}_{\mathbb{Q}}[-Y]) \leq \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_{\mathbb{Q}}[-X] + \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_{\mathbb{Q}}[-Y] = \rho(X) + \rho(Y).$$

(⇒) For brevity, we omit the detailed proof of the converse implication. The proof essentially relies on a version of Hahn-Banach separation theorem (in which convex sets can be separated by a hyperplane). See [FS02a, Proposition 2.85 and Corollary 4.35] or [MFE10, Proposition 6.11] for the proofs; see also [Art+99] for the original proof. □

The robust (dual) representation presented in Theorem 3.15 allow an alternative interpretation of a coherent risk measure output. We can look at the set $\mathcal{Q}$ in (3.8) as the set of potential probability (risk) scenarios for tomorrow. Then, according to Theorem 3.15, the CRM risk measure represents the greatest expected loss $\mathbb{E}_Q[-X]$ we might face across all conceivable 'scenarios' from $\mathcal{Q}$. Also, Theorem 3.15 allow us to easily construct CRMs by considering different sets $\mathcal{Q}$. Surprisingly, the (FP) property imposed in Theorem 3.15 is not very restrictive, especially from a practical point of view, where we deal with law invariant (LI) risk measures on a sufficiently reach probability space, e.g. atomless standard probability space, as in this lecture notes.

Theorem 3.16 (Law invariant risk measures satisfy Fatou property)**.** Let $\rho : L^\infty \to \mathbb{R}$ be a monetary risk measure that satisfies (CV) and (LI). Then, ρ satisfies (FP).

For the proof of Theorem 3.16 see [JST06, Theorem 2.1], where the proof for standard probability spaces is provided. Note that from Theorem 3.16 we get that, for sufficiently reach probability space, we can replace (FP) by (LI) in the dual representation statement in Theorem 3.15; recall that CRMs satisfy (CV) due to Proposition 2.7.

Remark 3.17 (Robust representations of risk measures)**.** In the literature, one can find multiple variants of robust representations, similar to the one presented in Theorem 3.15. Those results have been extended to various other space, e.g. L^p spaces or Orlicz hearts, and for various other types of maps, for which some form of convexity or concavity holds. In particular, it is worthwhile to note that one can remove the need of (FP) property by replacing the set $\mathcal{Q}$ in (3.8) by an appropriate subset of finitely-additive and normalized set functions, see e.g. [FS02a, Section 4.2]. ♢

Let us now check what are the robust representations for CRMs we considered so far.

Example 3.18 (Robust representation of ES)**.** Fix $\alpha \in (0,1)$ and consider expected shortfall risk measure ES_α from Definition 3.8. Then, ES_α admits robust representation (3.8) with the set

$$\mathcal{Q}_\alpha := \left\{ \mathbb{Q} \ll \mathbb{P} \mid \tfrac{d\mathbb{Q}}{d\mathbb{P}} \leq \tfrac{1}{\alpha} \right\}. \tag{3.9}$$

Furthermore, for any $X \in L^\infty$ the supremum in (3.8) is attained for a risk measure $\mathbb{Q}_X \in \mathcal{Q}_\alpha$ with density given by

$$\frac{d\mathbb{Q}_X}{d\mathbb{P}} = \tfrac{1}{\alpha}\left(\mathbb{1}_{\{X<q_\alpha\}} + \frac{\alpha - \mathbb{P}[X<q_\alpha]}{\mathbb{P}[X=q_\alpha]} \mathbb{1}_{\{X=q_\alpha\}} \right), \tag{3.10}$$

where $q_\alpha := -\mathrm{VaR}_\alpha(X)$ is α-quantile of X and $\frac{0}{0} = 0$ convention is used in (3.10). In particular, for continuous random variables, the second summand in (3.10) disappears, and we simply get

$$\mathrm{ES}_\alpha(X) = \mathbb{E}[-X|X<q_\alpha] = -\tfrac{1}{\alpha}\mathbb{E}[\mathbb{1}_{\{X<q_\alpha\}}X] = -\mathbb{E}[\tfrac{1}{\alpha}\mathbb{1}_{\{X<q_\alpha\}}X] = -\mathbb{E}[\tfrac{d\mathbb{Q}_X}{d\mathbb{P}}X] = \mathbb{E}_{\mathbb{Q}_X}[-X].$$

For details and proofs we refer to [FS02a, Theorem 4.47]. ♦

Example 3.19 (Robust representation of the worst case risk measure and the expected value based risk measures)**.** Let us consider the worst-case risk measure and expected value based risk measures defined in (3.6) and (3.7). It can be easily shown that they admit robust representations with sets $\mathcal{Q}^{\min} := \{\mathbb{Q} \mid \mathbb{Q} \ll \mathbb{P}\}$ and $\mathcal{Q}^{\mathrm{mean}} := \{\mathbb{P}\}$, respectively. ♦

As discussed, from practical point of view, we are mostly interested in CRMs which admit (LI). There is a natural question if such risk measures could admit any alternative representations based on standard distribution characteristics such as quantiles (value at risk measures) or other ones (e.g. expected shortfall measures). This is indeed the case, as the following two results will show.

Theorem 3.20 (Kusuoka's representation of CRMs)**.** Let $\rho : L^\infty \to \mathbb{R}$ be (LI). Then, ρ is CRM if and only if it there is a (compact convex) set of probability measures on $[0,1]$ denoted by $\mathcal{M}$ such that

$$\rho(X) = \sup_{\nu \in \mathcal{M}} \int_0^1 \mathrm{ES}_\alpha(X)\nu(d\alpha). \tag{3.11}$$

The proof of Theorem 3.20 can be found in [Kus01]. From Theorem 3.20 we see that the family of expected shortfall risk measure could be seen as a building blocks from which any coherent risk measure could be constructed. Notably, one can show that the supremum in (3.11) can be dropped if any only if ρ admits (CM).

Proposition 3.21 (Kusuoka's representation of comonotone CRMs)**.** Let $\rho : L^\infty \to \mathbb{R}$ be (LI). Then, ρ is a CRM that admits (CM) if and only if there is a probability measures on $[0,1]$ denoted by ν such that

$$\rho(X) = \int_0^1 \mathrm{ES}_\alpha(X)\nu(d\alpha). \tag{3.12}$$

For the proof and further discussion, we refer to [Kus01, Theorem 7]. The family of comonotone CRM plays an important role in financial mathematics. Recalling that ES is constructed as integral over VaR risk measures, see Definition 3.8, we can use Proposition 3.21 to easily recover the class of spectral risk measures that are build on VaR instead of ES. W4 ⊣ W5

Example 3.22 (Spectral risk measures)**.** Let $\phi : [0,1] \to \mathbb{R}_+$ be a *risk spectrum*, that is, a non-increasing and right-continuous function such that $\int_0^1 \phi(t)dt = 1$. Then, the mapping $\rho_\phi : L^\infty \to \mathbb{R}$ given by

$$\rho_\phi(X) := \int_0^1 \phi(t)\,\mathrm{VaR}_t(X)dt, \quad X \in L^\infty, \tag{3.13}$$

is called a *a spectral risk measure* with risk spectrum ϕ. The family of spectral risk measures was defined in [Ace02] where its properties, e.g. in terms of risk aversion, have been studied. From Proposition 3.21 and ES definition, we get that the spectral risk measure ρ_ϕ is CRM that is (LI) and (CM). In fact, one can also show that the properties imposed of ϕ are necessary for ρ_ϕ to be a CRM, see [Ace02, Theorem 4.1]. ♦

Remark 3.23 (Interaction between comonotonicity and convexity)**.** It can be shown that if ρ is a monetary risk measures, then comonotone property (CM) implies positive positive homogeneity (PH). Consequently, Proposition 3.21 also characterize the class of monetary risk measures that are (LI), (CV), and (CM). To see this, it is sufficient to note that for any $X \in L^\infty$, the pair (X,X) is a comonotone pair, so that we get $\rho(2X) = 2\rho(X)$ from which, by recursive argumentation, $\rho(\lambda X) = \lambda\rho(X)$ follows for any rational $\lambda > 0$; Lipshitz-continuity allow us to expand this to any real $\lambda > 0$, see [FS02a, Lemma 4.77] for details and more discussion. ♢

Remark 3.24 (Other characterizations of CRMs)**.** The class of CRM maps that are (LI) and (CM), defined in (3.12) can be also alternatively represented using so called *concave distortions* and *Choquet integrals*. We refer to [FS02a, Section 4.6] for more details. ♢

3.4 Convex risk measures

Now, let us investigate what happens if we replace (SBA) and (PH) with (CV) property.

Definition 3.25 (Convex risk measure)**.** We say that $\rho : L^\infty \to \mathbb{R}$ is a *convex risk measure* if ρ is is a monetary risk measure that satisfies (CV).

The class of convex risk measures is broader than the class of CRMs; the interpretations of normative properties for convex risk measures remain similar to those for CRMs. Note that the key difference between non-coherent convex risk measure and CRM is positive homogeneity (PH); recall that any convex risk measure that satisfies (PH) is CRM. That being said, note that from (CV) we immediately get

$$\rho(\lambda X) \geq \lambda \rho(X), \quad \text{for } \lambda > 1,\ X \in L^\infty, \tag{3.14}$$

and reverse inequality for $\lambda \in (0,1)$; to see this, use λ^{-1} in (3.14) and use (CV) property for X and 0. This means that position up-scaling leads to (potentially nonlinear) risk increase; this often reflects real-world situations in which concentrated positions could lead to additional risks. Similar remark apply to the effect of diversification – coherent risk measures might overstate the diversification benefits in some situations and let to non-rational portfolio selection. As expected, the family of convex risk measures also admits dual representation.

Theorem 3.26 (Robust representation of convex risk measure)**.** Assume that $\rho : L^\infty \to \mathbb{R}$ satisfies (FP). Then, ρ is a convex risk measure if and only if there exists a set of probability measures $\mathcal{Q}$, all absolutely continuous with respect to $\mathbb{P}$, and a function $\alpha : \mathcal{Q} \to [-\infty, \infty]$[a] such that

$$\rho(X) = \sup_{\mathbb{Q} \in \mathcal{Q}} \left(\mathbb{E}_{\mathbb{Q}}[-X] - \alpha(\mathbb{Q}) \right), \quad X \in L^\infty. \tag{3.15}$$

[a] we call α a *penalty function* as it provided a penalty linked to specific probability scenario choice

For the proof of Theorem 3.26 we refer to [FS02a, Theorem 4.31]. As already noted in Theorem 3.16, law invariant (LI) and convex risk measures automatically satisfy the Fatou property (FP), so this assumption is not very restrictive in most pragmatic cases. Also, note that the convex risk measure is CRMs if its penalty function is equal to zero. The penalty functions can be redefined on the set of all probability measures with additional condition that $\alpha(\mathbb{Q}) = +\infty$ whenever $\mathbb{Q} \notin \mathcal{Q}$, where $\mathcal{Q}$ is the set from (3.15). More explicitly, we know that (3.15) holds if and only if

$$\rho(X) = \sup_{\mathbb{Q} \in \mathcal{P}(\Omega, \mathcal{F})} \left(\mathbb{E}_{\mathbb{Q}}[-X] - \tilde{\alpha}(\mathbb{Q}) \right), \quad X \in L^\infty, \tag{3.16}$$

where $\mathcal{P}(\Omega, \mathcal{F})$ denotes the set of all probability measures on (Ω, Σ) and $\tilde{\alpha} \colon \mathcal{P}(\Omega, \mathcal{F}) \to [-\infty, \infty]$ is given by $\tilde{\alpha}(\mathbb{Q}) = \mathbb{1}_{\{\mathbb{Q} \in \mathcal{Q}\}} \alpha(\mathbb{Q}) + \mathbb{1}_{\{\mathbb{Q} \notin \mathcal{Q}\}}(+\infty)$. Note that this representation is also true for coherent risk measure with $\tilde{\alpha}(\mathbb{Q}) = \mathbb{1}_{\{\mathbb{Q} \notin \mathcal{Q}\}}(+\infty)$ and it allows to define risk measure using only the penalty function. For more information about the construction of penalty functions, and the definition of so called *minimal penalty function* (note that the penalty function not always uniquely defines a risk measure), we refer to [FS02a, p. 165]. For brevity, we often use representation (3.16) instead of (3.15).

Remark 3.27 (Link to functional analysis)**.** Theorem 3.26 can be viewed as a special case of the duality in convex analysis. Namely, it is a special case of *Fenchel–Moreau theorem* (bi-conjugation theorem) with *Legendre–Fenchel transformation* of the convex functions on L^∞. As such, this result can be extended to $L^p(\Omega, \Sigma, \mathbb{P})$ spaces (which are Banach spaces). For more information about (infinite-dimensional vector) spaces of random variables, the transfer of classical separation results from functional analysis, comments on dual spaces, and convex functions representations, we refer to [FS02a, Appendix A.7]. $\diamondsuit$

We have already discussed a couple of example of convex risk measures, as every coherent risk measure is also a convex risk measure. Arguably, the most important example of non-coherent convex risk measures is the entropic utility measure which is a negative of entropic utility measure from Definition 2.9.

Example 3.28 (Entropic risk measure and its dual representation)**.** For any risk-aversion parameter $\gamma < 0$, the mapping $\rho_\gamma \colon L^\infty \to \mathbb{R}$ given by

$$\rho_\gamma(X) := -\operatorname{Ent}_\gamma(X) = \begin{cases} -\frac{1}{\gamma}\ln \mathbb{E}[e^{\gamma X}] & \gamma \neq 0 \\ -\mathbb{E}[X] & \gamma = 0 \end{cases},$$

is called an *entropic risk measure* for risk-aversion γ and it is a convex risk measure.[11] The proof that ρ_γ is a convex risk measure, for $\gamma < 0$, follows directly from Proposition 2.10 and we omit it. The name of this family of risk measures comes from their dual representation given by

$$\rho_\gamma(X) = \sup_{\mathbb{Q}\in\mathcal{P}(\Omega,\mathcal{F})} \left\{ \mathbb{E}_{\mathbb{Q}}[-X] - \tfrac{1}{-\gamma} H[\mathbb{Q}\,\|\,\mathbb{P}] \right\}, \tag{3.17}$$

where $H[\mathbb{Q}\,\|\,\mathbb{P}]$ is the relative entropy of $\mathbb{Q}$ with respect to $\mathbb{P}$ given by

$$H[\mathbb{Q}\,\|\,\mathbb{P}] := \begin{cases} \mathbb{E}_{\mathbb{Q}}\left[\ln \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right] & \text{if } \mathbb{Q} \ll \mathbb{P}, \\ +\infty & \text{otherwise.} \end{cases} \tag{3.18}$$

This representation can be easily deduced by direct calculations, using Jensen's inequality and the fact that the exponential function is convex. That is, for any $\mathbb{Q} \ll \mathbb{P}$ and $\gamma < 0$, we have

$$\begin{aligned} \rho_\gamma(X) &= -\frac{1}{\gamma}\ln \mathbb{E}_{\mathbb{Q}}\left[\tfrac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}}\exp(\gamma X)\right] = -\frac{1}{\gamma}\ln \mathbb{E}_{\mathbb{Q}}\left[\exp\left(\gamma X - \ln \tfrac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right)\right] \\ &\geq -\frac{1}{\gamma}\ln\exp\left(\mathbb{E}_{\mathbb{Q}}\left[\gamma X - \ln\frac{d\mathbb{Q}}{d\mathbb{P}}\right]\right) = \mathbb{E}_{\mathbb{Q}}[-X] - \tfrac{1}{-\gamma}H[\mathbb{Q}\,\|\mathbb{P}]. \end{aligned} \tag{3.19}$$

On the other hand, by considering so called *Esscher transform of X*, which is a probability measure $\mathbb{Q}_X \ll \mathbb{P}$ defined via Radon-Nikodym derivative $\frac{d\mathbb{Q}_X}{d\mathbb{P}} = \frac{e^{\gamma X}}{\mathbb{E}[e^{\gamma X}]}$, we recover

$$\tfrac{1}{\gamma}H[\mathbb{Q}_X\,\|\,\mathbb{P}] = \tfrac{1}{\gamma}\mathbb{E}\left[\tfrac{\mathrm{d}\mathbb{Q}_X}{\mathrm{d}\mathbb{P}}\ln\tfrac{\mathrm{d}\mathbb{Q}_X}{\mathrm{d}\mathbb{P}}\right] = \frac{1}{\gamma}\mathbb{E}\left[\frac{e^{\gamma X}}{\mathbb{E}[e^{\gamma X}]}(\gamma X - \ln\mathbb{E}[e^{\gamma X}])\right] = \tfrac{\mathbb{E}[Xe^{\gamma X}]}{\mathbb{E}[e^{\gamma X}]} + \rho_\gamma(X).$$

Combining this with the fact that $\mathbb{E}_{\mathbb{Q}_X}[-X] = -\frac{\mathbb{E}[Xe^{\gamma X}]}{\mathbb{E}[e^{\gamma X}]}$, we get

$$\rho_\gamma(X) = \mathbb{E}_{\mathbb{Q}_X}[-X] - \tfrac{1}{-\gamma}H[\mathbb{Q}_X\,\|\,\mathbb{P}]. \tag{3.20}$$

Combining (3.19) and (3.20), we get (3.17). ♦

Remark 3.29 (A note on the usage of convex risk measures)**.** While convex risk measures are popular in the mathematical finance literature, they are not used that often in practical P&L risk management, and both financial and insurance sector regulations are essentially build on measures that satisfy (PH) property. That saying, (negatives of) convex risk measures are becoming more and more popular in financial stochastic optimization where they are often applied to random variables

[11] As in Definition 1.1, many authors use different parametrization of the entropic risk measure, that is, they consider parameter $\tilde{\gamma} > 0$ defined as $\tilde{\gamma} := -\gamma$, see e.g. [KS09].

representing logarithmic return-rates rather than P&L vectors. Note that, in such context, the interpretation of normative axioms might significantly change. To give an example, the logarithmic return rate of a linearly scaled financial position is the same as the logarithmic return rate for unscaled position, that is, for any $\lambda > 0$ and strictly positive position price process $S_t \in L^\infty$ (defined for $t \in \{0, 1\}$), we get $\ln \frac{\lambda S_1}{\lambda S_0} = \ln \frac{S_1}{S_0}$. This means that positive homogeneity (on a position scaling level) is inherently satisfied when risk measure is applied to log-return vectors; of course, the re-construction of a future price λS_t when our initial capital is raised from S_0 to λS_0 might be impossible due to other reasons linked e.g. to the liquidity, price impact effect, etc. For more information on this topic and examples of financial application with entropic risk we refer to the *risk-sensitive stochastic control* literature, see e.g. [BP99; BP03; Ste99; PS23]. $\lozenge$

W5 ⊣ W6

4 From risk to performance measurement

In the previous section, we focused on the risk measures. Apart from being used to establish prudent capital reserves, risk measures are often combined with specific rewards criteria (e.g. mean) to constitute a new class of measures that are used to measure the performance; the mean-variance criterion could be seen as a toy example of such measure. In this section, we focus on such measures, set a specific set of axioms which a mapping should satisfy to become a proper performance measures, and then provide a comment on the duality between risk measures and acceptability indices.

While the measures used for performance measurement could be, in principle, applied to financial position P&Ls directly, they are often applied to other quantities. This includes:

- Application to simple or logarithmic return rates (e.g. in the portfolio optimization context) on fixed-point time intervals, using infinite discounted schemes, etc.
- Relative-type measurements using excess P&Ls or excess returns that allows incorporation of risk-free rates, discounting schemes, benchmark position returns, or capital reserve securitization.

The exact application is often driven by the underlying problem formulation and over-arching optimization objective, see [CH09] for more details and other applications. For example, if our goal is the financial investment performance measurement, we should compare ourselves with popular benchmark portfolios (e.g. stock index returns) or with secure asset returns (e.g. risk-free rate) in order to avoid unnecessary risk exposures. In contras, if we want to measure the performance (security/safeness) of a position X that was secured by a specific amount of capital, say $\rho(X)$ we might want to measure the performance of a secured position $Y = X + \rho(X)$ rather than X.

Remark 4.1 (Kelly criterion)**.** As noted in Remark 3.29 one can embed return rate transform into the measurement map and study its properties. For example, assuming that our initial capital is $S_0 \in \mathbb{R}_+$ and considering future (strictly positive) investment strategies leading to capital $S_0 \in L^\infty$, we can define map

$$(S_0, S_1) \mapsto \mathbb{E}\left[\ln \tfrac{S_1}{S_0}\right],$$

which corresponds to so called *Kelly criterion* performance measures. Note that this map (understood as the function of S_1) is no longer additive and does not satisfy the usual mean properties. Despite its simplicity, this map is a very popular financial investment criterion, also used by top-financial investors, see [MTZ11] for details. $\lozenge$

For simplicity and to streamline the narrative, if not stated otherwise, in this lecture notes we still assume that random variables represent financial position future P&L (for some fixed time-horizon) or simple returns.

4.1 Risk-to-reward ratios

We start this section with, arguably, the most popular measure that are used to measure financial position performance, that is, the *Sharpe ratio*. Essentially, it is the ratio of (excess) expected return to standard deviation; see [Ale09] for economic insight and [Sha66] for historical background.

Definition 4.2 (Sharpe ratio)**.** We define the mapping $\mathrm{SR}\colon L^\infty \to [0,+\infty]$ given by

$$\mathrm{SR}(X) := \begin{cases} \frac{\mathbb{E}[X]}{\mathrm{Std}(X)}, & \mathbb{E}[X] > 0,\\ 0, & \text{otherwise,} \end{cases} \tag{4.1}$$

with convention $\frac{1}{0} = \infty$, and call it *Sharpe Ratio* (SR).[a]

[a] In the literature you can find other names, e.g. *Sharpe index* or *Sharpe measure*.

In the financial literature, SR is typically applied to $X = R_a - R_f$ denoting the excess return rate in which the risk-free return R_f (or benchmark portfolio return) is subtracted from the underlying asset return R_a, before the measurement is taken. Then, SR measures expected excess return of a portfolio in units of portfolio's standard deviation. Sharpe Ratio is often used as a classical tool to rank portfolios according to their *reward-to-risk* and to evaluate the attractiveness of financial investments. That saying, using standard deviation to quantify risk is considered to be the major drawback of Sharpe Ratio. The reason is that positive returns also contribute to this measure of risk and SR is in general not monotone (example construction is left as an exercise).

To eliminate the aforementioned and other drawbacks of SR, multiple alternatives have been proposed. Among the most popular alternatives, that are often use in economic literature, one can find *Sortino Ratio* and *Gain-Loss Ratio*. For completeness, let us define those mapping and briefly discuss their properties, we refer to [SP94] and [BL00] for economic details.

Definition 4.3 (Gain Loss Ratio and Sortino Ratio)**.** We define the mappings $\mathrm{GLR}\colon L^\infty \to [0,+\infty]$ and $\mathrm{SOR}\colon L^\infty \to [0,+\infty]$ given by

$$\mathrm{GLR}(X) := \begin{cases} \frac{\mathbb{E}(X)}{\mathbb{E}(X^-)}, & \mathbb{E}[X] > 0,\\ 0, & \text{otherwise,} \end{cases} \quad \text{and} \quad \mathrm{SOR}(X) := \begin{cases} \frac{\mathbb{E}[X]}{\sqrt{\mathbb{E}[(X^-)^2]}}, & \mathbb{E}[X] > 0,\\ 0, & \text{otherwise,} \end{cases} \tag{4.2}$$

with convention $\frac{1}{0} = \infty$, and call them *Gain Loss Ratio* (GLR) and *Sortino Ratio* (SOR).

Both GLR and SOR focus on downside risk and share some common desirable features (that we will prove below): they are unit-less (eventual units cancel each other), they are increasing functions of reward and decreasing functions of risk, they are scale invariant, and, as will be later shown, the diversification will lead to performance increase. Instead of taking the downside risk, one can incorporate the generic risk measurement into the ratio.

Definition 4.4 (Risk Adjusted Return on Capital)**.** We say that $\alpha\colon L^\infty \to [0,+\infty]$ is a *Risk Adjusted Return on Capital (RAROC)* if it can be represented as

$$\alpha(X) := \begin{cases} \frac{\mathbb{E}[X]}{\rho(X)}, & \mathbb{E}[X] > 0, \\ 0, & \text{otherwise}, \end{cases} \tag{4.3}$$

for some coherent risk measure $\rho\colon L^\infty \to \mathbb{R}$; we use convention $\alpha(X) = +\infty$ if $\rho(X) \leq 0$.

For more comprehensive study on risk-to-reward ratios and numerous other examples, we refer to [CK12]. As for GLR and SOR, RAROCs also admit specific normative properties linked e.g. to scale invariance. This observation prompts a natural desire to study performance measure in a unified mathematical framework.

Example 4.5 (RAROC for ES under normality)**.** As risk measures, the risk-reward ratios could be easily extended to other spaces of random variables, e.g. L^1. Fix $\alpha \in (0,1)$ and consider expected shortfall risk measure ES_α from Definition 3.8 as an input to RAROC criterion defined in (4.3) (defined for $X \in L^1$). Assume that $X \sim N(\mu,\sigma)$ for $\mu \in \mathbb{R}$ and $\sigma > 0$. Then, using (3.5), we get

$$\tilde{\alpha}(X) = \frac{\mathbb{E}[X]}{\mathrm{ES}_\alpha(X)} = \frac{\mu}{-\mu + \sigma\frac{\phi(\Phi^{-1}(\alpha))}{\alpha}} = \frac{\mu}{\sigma}\cdot\left(\frac{\phi(\Phi^{-1}(\alpha))}{\alpha} - \frac{\mu}{\sigma}\right)^{-1},$$

so that RAROC could be seen as a function of SR. In fact, this is the case for all introduced risk-to-reward ratios. ♦

4.2 Coherent acceptability indices

Let us now formalize the concept of a performance measure by imposing a set of specific normative properties. If we consider (MD) and (CA) as constituting properties of monetary risk measures (see Definition 3.1), then (MI) and (SI) are constituting properties of a performance measure. The systematic study of such maps, in a coherent framework, has been initiated in [CM09], where the term *acceptability index* has been coined as a mathematical terminology for performance measures. Now, we provide a definition of a *coherent acceptability index*.

Definition 4.6 (Coherent acceptability index)**.** We say that $\alpha : L^\infty \to [0,+\infty]$ is a *coherent acceptability index* (CAI) if it is

1) *Monotone increasing* (MI) if $X \geq Y$ imply $\alpha(X) \geq \alpha(Y)$, for $X, Y \in L^\infty$;
2) *Scale invariant* (SI), if $\alpha(\lambda X) = \alpha(X)$, for $X \in L^\infty$ and $\lambda > 0$;
3) *Quasi-concave* (QCC) if $\alpha(\lambda X + (1-\lambda)Y) \geq \alpha(X) \wedge \alpha(Y)$, for $X, Y \in L^\infty$ and $\lambda \in (0,1)$.
4) *Fatou continuous*, i.e. it satisfies Fatou property (FP) – for any $x \in \mathbb{R}$, if $\alpha(X_n) \geq x$ then $\alpha(X) \geq x$, assuming that $|X_n| \leq 1$ are such that $X_n \to X$, as $n \to \infty$, in probability.

As before, all properties have rather direct financial interpretation; note that scale invariance is the one which stands out, as it is specific for acceptability indices. Note that Quasi-concavity (QCC) implies that a diversified portfolio performs at higher level than its components. Notably, any CAI has a duality representation through CRMs, that is, any CAIs could be represented through a family of descending probability scenarios which could be later translated to the family of increasing family of CRMs, which justifies the term *coherent* in the name.

Theorem 4.7 (Robust representation of CAIs)**.** Assume that $\alpha : L^\infty \to [0,+\infty]$ is unbounded from above. Then, α is a coherent acceptability index if and only if there exists an ascending family of probability measures absolutely continues with respect to $\mathbb{P}$, say $(\mathcal{Q}_x)_{x\in\mathbb{R}_+}$ [a], such that

$$\alpha(X) = \sup\left\{x \in \mathbb{R}_+ \colon \inf_{\mathbb{Q}\in\mathcal{Q}_x} \mathbb{E}_{\mathbb{Q}}[X] \geq 0\right\}, \quad X \in L^\infty, \tag{4.4}$$

where the convention $\inf\emptyset = \infty$ and $\sup\emptyset = 0$ is used.

[a] The ascending property reflect the fact that $\mathcal{Q}_x \subseteq \mathcal{Q}_y$ for $x, y \in \mathbb{R}_+$ such that $x < y$.

Proof. For brevity, we only show that the map defined in (4.4) is CAI. The proof of the reverse implication can be found in [CM09, Theorem 1]; the core idea is to define a family of so called *acceptance sets* linked to α, that is, sets $\mathcal{A}_x := \{X \in L^\infty \colon \alpha(X) \geq x\}$, for $x \in \mathbb{R}_+$, then recover the risk measures $\rho_x(X) := \sup\{m \in \mathbb{R} \colon X - m \in \mathcal{A}_x\}$ from α and show they are coherent, and next use robust representation to get the final representation.

(MI) Assume $X \geq Y$. For any $x \in \mathbb{R}_+$ and $\mathbb{Q} \in \mathcal{Q}_x$ we get $\mathbb{E}_{\mathbb{Q}}[X] \geq \mathbb{E}_{\mathbb{Q}}[Y]$, and consequently if $\mathbb{E}_{\mathbb{Q}}[Y] \geq 0$ then $\mathbb{E}_{\mathbb{Q}}[X] \geq 0$. The same inequality holds after applying inf operator, which implies $\alpha(X) \geq \alpha(Y)$.

(SI) Fix $\lambda > 0$ and $X \in L^\infty$. By direct calculations, using (PH) property of expectation, we get

$$\alpha(\lambda X) = \sup\left\{x \in \mathbb{R}_+ \colon \lambda \inf_{\mathbb{Q}\in\mathcal{Q}_x} \mathbb{E}_{\mathbb{Q}}[X] \geq 0\right\} = \sup\left\{x \in \mathbb{R}_+ \colon \inf_{\mathbb{Q}\in\mathcal{Q}_x} \mathbb{E}_{\mathbb{Q}}[X] \geq 0\right\} = \alpha(X).$$

(QCC) Fix $X, Y \in L^\infty$ and $\lambda \in (0,1)$. For any $x \in \mathbb{R}_+$ and $\mathbb{Q} \in \mathcal{Q}_x$ we know that if $\mathbb{E}_{\mathbb{Q}}[X] \geq 0$ and $\mathbb{E}_{\mathbb{Q}}[Y] \geq 0$, then $\mathbb{E}_{\mathbb{Q}}[\lambda X + (1-\lambda)Y] \geq 0$ for any $\lambda \in (0,1)$. Same holds after applying inf operator, which implies $\alpha(\lambda X + (1-\lambda)Y) \geq \alpha(X) \wedge \alpha(Y)$ and concludes the proof.

(FP) Take any sequence $(X_n)_{n\in\mathbb{N}}$ such that $|X_n| < 1$, $X_n \to X$ in probability, as $n \to \infty$, and $\alpha(X_n) \geq x$ for some $x \in \mathbb{R}_+$. Using the (FP) property of the expectation for any $y \in \mathbb{R}_+$ such that $y < x$ and $\mathbb{Q} \in \mathcal{Q}_y$ we know that $\mathbb{E}_{\mathbb{Q}}[X_n] \geq 0$, for $n \in \mathbb{N}$, and consequently $\mathbb{E}_{\mathbb{Q}}[X] \geq 0$. As the choice of y has been arbitrary, we get $\alpha(X) \geq x$, and conclude the proof. □

One can look at CAI as a performance measures which check how much we can increase the set of a given probabilistic scenarios (following a prescribed addition scheme) to keep the position acceptable. By combining Theorem 4.7 and Theorem 3.15 and performing some basic algebraic operations we see that CAI can be in fact represented via an increasing family of coherent risk measures, say $\rho_x \colon L^\infty \to \mathbb{R}$, for $x \in \mathbb{R}_+$, satisfying (FP) and such that[12]

$$\alpha(X) = \sup\{x \in \mathbb{R}_+ \colon \rho_x(X) \leq 0\}, \quad X \in L^\infty. \tag{4.5}$$

As noted above, one can show that while SR satisfies (SI), (QCC), and (FP), it is not CAI due to the lack of (MI). Let us now show that GLR, SOR, and RAROC are in fact CAIs.

Proposition 4.8 (GLR, SOR and RAROC are CAIs)**.** The mappings GLR, SOR, and RAROC for CRM satisfying (FP), defined in (4.2) and (4.3), are coherent acceptability indices.

[12] by increasing family we mean that for any $X \in L^\infty$ and $x, y \in \mathbb{R}_+$ satisfying $x > y$, we get $\rho_x(X) \geq \rho_y(X)$

Proof. For brevity, we only show the sketch of the proof presented jointly for all mappings. We fix an arbitrary coherent risk measure ρ satisfying (FP) that is used in the definition of RAROC.

(MI) Recalling that expectation is (MI), it is enough to note that the denominator of GLR, SOR, and RAROC is (MD). We first note that ρ in the definition of RAROC is (MD) as it is CRM. Next, for any $X, Y \in L^\infty$, if $X \le Y$, then $X^- \ge Y^-$ and $(X^-)^2 \ge (Y^-)^2$, and thus $\mathbb{E}[X^-] \ge \mathbb{E}[Y^-]$ and $\mathbb{E}(X^-)^2 \ge \mathbb{E}(Y^-)^2$, which yields (MD) property for both denominators and (MI) of GLR and SOR.

(SI) This property follow directly from the fact that both numerators and denominators of all considered mappings are (PH).

(QCC) The proof for RAROC is immediate, as expectation is a linear operator and ρ is (QCV). Now, let us show the proof for GLR. Suppose that $\mathrm{GLR}(X) \ge x$ and $\mathrm{GLR}(Y) \ge x$, for some $x \in \mathbb{R}_+$. Then, we have that $\mathbb{E}[X] \ge x\mathbb{E}[X^-]$ and $\mathbb{E}[Y] \ge x\mathbb{E}[Y^-]$. Thus, by the convexity of the function x^-, we get

$$x\mathbb{E}[(\lambda X + (1-\lambda)Y)^-] \le x(\lambda\mathbb{E}[X^-] + (1-\lambda)\mathbb{E}[Y^-]) \le \mathbb{E}[\lambda X + (1-\lambda)Y],$$

which implies $\mathrm{GLR}(\lambda X + (1-\lambda)Y) \ge x$, and concludes the proof of (QCC) for GLR. Next, we show the proof for SOR. By the Cauchy-Schwarz inequality, we have

$$\mathbb{E}[X^-Y^-] \le \sqrt{\mathbb{E}[(X^-)^2]\mathbb{E}[(Y^-)^2]}\,.$$

Thus, by direct calculations, we get

$$\begin{aligned}\Big(\lambda\sqrt{\mathbb{E}[(X^-)^2]} + (1-\lambda)\sqrt{\mathbb{E}[(Y^-)^2]}\Big)^2 &\ge \lambda^2\mathbb{E}[(X^-)^2] + (1-\lambda)^2\mathbb{E}[(Y^-)^2] + 2\lambda(1-\lambda)\mathbb{E}[X^-Y^-]\\ &= \mathbb{E}\Big[\Big(\lambda X^- + (1-\lambda)Y^-\Big)^2\Big]\\ &\ge \mathbb{E}\Big[\Big((\lambda X + (1-\lambda)Y)^-\Big)^2\Big],\end{aligned}$$

from which (QCC) for SOR follows.

(FP) This follow directly from the (FP) property of all maps used in the definition of GLR, SOR, and RAROC; note that the proof for $x = 0$ is immediate and convergence in probability imply convergence in L^p norm for uniformly bounded (by a constant) random variables. ☐

For the statement of additional economic properties of CAIs, and discussion about other families of CAIs such as AIT, AIW, AIMIN, AIMAX, AIMINMAX, or AIMAXMIN we refer to [CM09].

4.3 Quasi-concave acceptability indices

As in the risk measure case, we can generalize the family of coherent acceptability indices from a coherent to a convex setup. This is done by removing the (SI) property in a similar way that (PH) was removed in the risk measurement setup.

Definition 4.9 (Quasi-concave acceptability index)**.** We say that $\alpha : L^\infty \to [0, +\infty]$ is a *quasi-concave acceptability index* (QAI) if it is

1) *Monotone increasing* (MI) if $X \ge Y$ imply $\alpha(X) \ge \alpha(Y)$, for $X, Y \in L^\infty$;

2) *Quasi-concave* (QCC) if $\alpha(\lambda X + (1-\lambda)Y) \ge \alpha(X) \wedge \alpha(Y)$, for $X, Y \in L^\infty$ and $\lambda \in (0, 1)$.

3) *Fatou continuous*, i.e. it satisfies Fatou property (FP) – for any $x \in \mathbb{R}$, if $\alpha(X_n) \ge x$ then $\alpha(X) \ge x$, assuming that $|X_n| \le 1$ are such that $X_n \to X$, as $n \to \infty$, in probability.

As shown in [RS13], one can link the family of QAIs to convex risk measures via robust representation. For brevity, we follow the convention in which the penalty function defines the mapping, see (3.16).

Theorem 4.10. Assume that the mapping $\alpha\colon L^\infty \to [0,+\infty]$ satisfies $\alpha(c) = +\infty$ for $c \geq 0$ and $\alpha(c) = 0$ for $c < 0$.[a] Then, α is a quasi-concave acceptability index if and only if there exists a decreasing family of penalty functions $\nu_x\colon \mathcal{P}(\Omega,\mathcal{F}) \to [-\infty,\infty]$, $x \in \mathbb{R}_+$, such that

$$\alpha(X) = \sup\left\{x \in \mathbb{R}_+\colon \inf_{\mathbb{Q}\in\mathcal{P}(\Omega,\mathcal{F})} (\mathbb{E}_{\mathbb{Q}}[X] + \nu_x(\mathbb{Q})) \geq 0\right\}, \quad X \in L^\infty, \tag{4.6}$$

where the convention $\inf\emptyset = \infty$ and $\sup\emptyset = 0$ is used.

[a] Note that this assumption is similar to unboundedness in Theorem 4.7.

For the proof, see [RS13, Proposition 3]. Again, using basic calculations one can rewrite (4.4) using the family of increasing convex risk measures, cf. (4.5). Notably, from the (QCC) property, we know that QAIs are in fact sub-scale invariant, that is, we have

$$\begin{aligned} \alpha(\lambda X) &\geq \alpha(X) \text{ for } \lambda \in (0,1),\\ \alpha(\lambda X) &\leq \alpha(X) \text{ for } \lambda \in (1,+\infty), \end{aligned}$$

which might be linked to market liquidity property. Again, we refer to [RS13]; see also [Rig24] for some further generalizations of acceptability indices.

Example 4.11 (QAIs based on CEs)**.** One can also define specific families of QAIs using scaled utilities and the corresponding families of certainty equivalents. Indeed, we can consider a concave, strictly increasing, and bounded from above *utility function* $U \in C^2(\mathbb{R})$, the scaled family $(U_x)_{x\in\mathbb{R}_+}$, where $U_x(z) := U(x\cdot z)$ for $\gamma > 0$ and $z \in \mathbb{R}$, and the corresponding family of certainty equivalents given by

$$\mu_x(X) := -U_x^{-1}(\mathbb{E}[U_x(X)]) = -\tfrac{1}{x}U^{-1}(\mathbb{E}[U(xX)]). \tag{4.7}$$

While $(\mu_x)_{x\in\mathbb{R}_+}$ typically fails to be a risk measure due to lack of (CA), one can consider utilities for which the Arrow-Pratt risk aversion function is non-decreasing (so that the mapping $x \to \mu_x(\cdot)$ is increasing) and define the mapping

$$\alpha(X) = \sup\{x \in \mathbb{R}_+ : \mu_x(X) \leq 0\}, \quad X \in L^\infty.$$

The mapping α is indeed QAI and could be re-represented using a family of convex risk measures). Notably, this mapping satisfies inverse positive homogeneity, that is, we have $\alpha(\lambda X) = \frac{1}{\lambda}\alpha(X)$, for $\lambda > 0$. We refer to [PR24] for proofs and more details. ♦

In the end of this section, we note that robust representations in (4.7) and (4.10) effectively allow us to represent acceptability indices via the family of increasing risk measures with the convention that the index should attain any value in $\mathbb{R}_+$. Of course, those results could be parametrized so that the values of the index represent risk aversion coefficient or confidence threshold. This immediately leads to an interesting question: *Can we meaningfully define performance measure based on the* $(\mathrm{VaR}_\alpha)_{\alpha\in(0,1)}$ *family via a formula similar to* (4.5)*?* The answer is positive, see next example.

Example 4.12 (Performance measure based on VaR)**.** Let us consider the performance measure given by

$$\tilde{\alpha}_{\mathrm{VaR}}(X) := \sup\{\alpha \in (0,1)\colon\ \mathrm{VaR}_\alpha(X) \leq 0\}, \quad X \in L^\infty$$

that is based on the family of VaR risk measures. This measure attains values in $[0,1]$ and is not an acceptability index in the classical sense (e.g. as VaR family of risk measures is not convex). That saying, since VaR family of risk measure plays crucial role in mathematical finance, it is natural to ask if the performance measure based on the VaR family is still used. To answer this question, we first note that for continuous random variables, we get that

$$\tilde{\alpha}_{\mathrm{VaR}}(X) = \mathbb{P}[X \leq 0], \tag{4.8}$$

so that α is effectively checking what is the minimal confidence that makes our position safe, that is, not exceeding a loss beyond the respective confidence interval. In particular, given any (estimated) capital reserve Z, that was used to secure position X at level $\alpha \in (0,1)$, we can check if the performance of the secured position $X+Z$ is as expected, that is, if we get $\tilde{\alpha}_{\mathrm{VaR}}(X+Z) = \alpha$. Of course, if we knew the actual capital needed for securitization and set $Z = \mathrm{VaR}_\alpha(X)$, then we would immediately get $\mathbb{P}[X+Z \leq 0] = \mathbb{P}[X \leq -\mathrm{VaR}_\alpha(X)] = \alpha$, from which $\tilde{\alpha}_{\mathrm{VaR}}(X+Z) = \alpha$ follows. Unfortunately, this is often not the case in practice, where we need to estimate the risk of the financial position as well as the corresponding capital reserve. In fact, as we later show in Section 5.4, the empirical version of (4.8) constitutes the exception rate statistic, which is a key metric in VaR regulatory backtesting being an integral part of Market Risk IMA capital framework. ⧫

Example 4.13 (QAI based on entropic risk measure)**.** Assume that QAI is defined via the family of entropic risk measures with reversed risk-aversion specification, that is, let

$$\alpha_{\mathrm{Ent}}(X) := \sup\{\gamma \in \mathbb{R}_+ \colon -\mathrm{Ent}_{-\gamma}(X) \leq 0\}, \quad X \in L^\infty.$$

This definition could be extended to random variables with all finite moments. For example, assume that $X \sim N(\mu,\sigma)$. Noting that $-\mathrm{Ent}_{-\gamma}(X) = -\mu + \frac{\gamma}{2}\sigma^2$, for $\gamma \in \mathbb{R}_+$, we get

$$\alpha_{\mathrm{Ent}}(X) = \sup\{\gamma \in \mathbb{R}_+ \colon \tfrac{\gamma}{2}\sigma^2 \leq \mu\} = 2\frac{\mu}{\sigma^2}, \tag{4.9}$$

which recovers the value of the risk-aversion coefficient in the mean-variance stochastic optimization problems (when we fix μ and σ), cf. Definition 1.1 and [Whi90]. Note that inverse positive homogeneity property is due to the usage of variance instead of standard deviation in the denominator in (4.9).

⧫

W6 ⊣ W7

5 Estimating risk and evaluating the performance

In the previous sections, we have defined mappings which assign numerical values to random variables. In this section we restrict ourselves to mapping which satisfy the *law invariance* (LI) property, that is, mappings $\rho\colon L^\infty \to \mathbb{R} \cup \{+\infty\}$ satisfying $\rho(X) = \rho(Y)$ for $X, Y \in L^\infty$, such that $F_X \equiv F_Y$. In most practical application, the distribution of the underlying random variable $X \in L^\infty$ (e.g. future P&L vector) is unknown, and we need to estimate it using the given or reconstructed statistical sample from X. In this section, we follow a standard simplified setup and assume we are given an i.i.d. sample $(X_i)_{i\in\mathbb{N}}$ from X, where $F_{X_i} \equiv F_X$, for $i \in \mathbb{N}$.

To simplify notation, we fix the finite sample size $n \in \mathbb{N}$, and use $\boldsymbol{X} := (X_1,\ldots,X_n)$ to denote the finite subsample from X; we also use $x = (x_1,\ldots,x_n) \in \mathbb{R}^n$ to denote a vector of realizations from $\boldsymbol{X}$. As usual, we also use $x_{(k)}$ to denote the kth order statistics from x, which could be seen as the kth element of the sorted version of vector x. Given the underlying risk or performance measurement mapping $\rho\colon L^\infty \to \mathbb{R}\cup\{+\infty\}$ and sample size $n\in\mathbb{N}$, we use $\hat{\rho}_n\colon \mathbb{R}^n \to \mathbb{R}$ to denote its estimator; we always assume that $\hat{\rho}_n$ is a measurable function. In particular, the value $\hat{\rho}_n(x)$ is estimating the value $\rho(X)$, and $\hat{\rho}_n(\boldsymbol{X})$ is a random variable that is measurable with respect to (random) sample data. Often, with slight abuse of notation, we use $\hat{\rho}$ instead of $\hat{\rho}_n$, or even use $\hat{\rho}$ do refer to the whole family of functions $(\hat{\rho}_n)_{n\in\mathbb{N}}$.

Of course, not any $\hat{\rho}$ (or value $\hat{\rho}(x)$) is properly estimating ρ (or $\rho(X)$) and one needs to impose additional property on it. Those properties could be linked to bias, consistency, etc. In this section we outline the key properties, and also introduce the most common estimation methods.

5.1 Plug-in estimation procedures

It is easy to note that for law invariant (LI) maps, there exists a distribution-based version of ρ.

Definition 5.1. Let $\rho\colon L^\infty \to \mathbb{R}\cup\{+\infty\}$ be (LI). Then, we call a mapping $R:\mathcal{D}\to\mathbb{R}\cup\{+\infty\}$, where $\mathcal{D}$ is the set of all probability distributions, a *distribution-based version of* ρ if

$$\rho(X) = R(F_X), \quad \text{for } X\in L^\infty.$$

From now on, we use R to denote a distribution-based version of the generic mapping ρ.

Most risk and performance measures introduced in this lecture notes satisfy (LI) property and could be re-expressed using R instead of ρ

Example 5.2 (Distribution-based versions of VaR and ES)**.** Fix $\alpha\in(0,1)$. Then, directly from (3.1) and (3.2), we get that the mappings given by

$$R_{\mathrm{VaR}_\alpha}(F) := -\inf\{z\in\mathbb{R} \mid F(z)>\alpha\} \quad\text{and}\quad R_{\mathrm{ES}_\alpha}(F) := \frac{1}{\alpha}\int_0^\alpha R_{\mathrm{VaR}_t}(F)\,\mathrm{d}t, \quad \text{for } F\in\mathcal{D}, \tag{5.1}$$

are the distribution-based versions of VaR_α and ES_α. ♦

The typical procedure that is followed when estimating risk or performance for a given position $X\in L^\infty$, that is $\rho(X)$, goes as follows:

Step 1: Given sample $\boldsymbol{X}$, we estimate the underlying distribution by a classical parametric or non-parametric estimation methodology obtaining estimate $\hat{F}_X$ of F_X.

Step 2: We plug-in the estimated distribution $\hat{F}_X$ into the distribution-based version of ρ obtaining the estimator $\hat{\rho}_{\text{plug-in}}(\boldsymbol{X}) = \rho(\hat{F}_X)$.

This estimation procedure is called a *plug-in estimation* procedure because we plug the estimated distribution into the risk-based functional to get the estimated value. Of course, the distribution-based functional is typically highly non-linear, and it is hard to infer any (non-asymptotic) properties of estimator $\hat{\rho}_{\text{plug-in}}$ from the estimation methodology that was used to get $\hat{F}_X$ from sample $\boldsymbol{X}$. That saying, one can recover most of the risk estimators following the logic described above.

Let us start with a series of examples linked to a non-parametric estimation of VaR and ES; similar reasoning could be applied to any other risk or performance measure.

Example 5.3 (VaR plug-in estimator based on ECDF)**.** Given $x\in\mathbb{R}^n$, a realized sample from X, we can define the standard *empirical cumulative distribution function* (ECDF) of X by setting

$$\hat{F}_X^{\text{emp}}(t) := \tfrac{1}{n}\textstyle\sum_{i=1}^n \mathbb{1}_{\{x_i\le t\}}. \tag{5.2}$$

Then, fixing $\alpha \in (0,1)$ and plugging-in (5.2) into (5.1), we get the estimator

$$\begin{aligned}\hat{\mathrm{VaR}}_{\alpha}^{\mathrm{emp}}(x) := R_{\mathrm{VaR}_{\alpha}}(\hat{F}_X^{\mathrm{emp}}) &= -\inf\left\{z \in \mathbb{R} \mid \tfrac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{x_i \leq z\}} > \alpha\right\} \\ &= -\inf\left\{z \in \mathbb{R} \mid \sum_{i=1}^{n} \mathbb{1}_{\{x_{(i)} \leq z\}} > n\alpha\right\} \\ &= -x_{(\lfloor n\alpha \rfloor + 1)}, \end{aligned} \tag{5.3}$$

with convention $\inf \emptyset = x_{(n)}$, which is one of the common choices for VaR estimator. To ensure a more conservative estimate, the ECDF defined in (5.2) is sometimes modified to provide a more conservative output, leading to estimator $\tilde{\mathrm{VaR}}_{\alpha}^{\mathrm{emp}}(x) = -x_{(\lfloor n\alpha \rfloor)}$, defined for $\alpha > n^{-1}$. ♦

Example 5.4 (Other VaR plug-in estimators based on ECDF)**.** Given $x \in \mathbb{R}^n$, a realized sample from X, we can consider various modifications of the (modified) ECDF that would lead to different risk estimators. For example, let us consider a smoother (linearly-interpolated) version of the ECDF that is given by

$$\hat{F}_X(t) := \begin{cases} 0 & \text{if } t < x_{(1)}, \\ \frac{i}{n} + \frac{t - x_{(i)}}{x_{(i+1)} - x_{(i)}} \cdot \frac{1}{n} & t \in [x_{(i)}, x_{(i+1)}),\ \text{for } i = 1, \ldots, n-1, \\ 1, & \text{if } t > x_{(n)}, \end{cases}$$

Using the plug-in procedure and performing calculations similar to those in Example 5.3, for $\alpha > n^{-1}$, we get the estimator VaR_{α} of the form

$$\hat{\mathrm{VaR}}_{\alpha}(x) = R_{\mathrm{VaR}_{\alpha}}(\hat{F}_X) = \lambda x_{(\lfloor n\alpha \rfloor)} + (1-\lambda) x_{(\lfloor n\alpha \rfloor + 1)},$$

where $\lambda := n\alpha - \lfloor n\alpha \rfloor$. In particular, for $n = 250$ and $\alpha = 1\%$ we get estimator

$$\hat{\mathrm{VaR}}_{1\%}(x) = \tfrac{1}{2} x_{(2)} + \tfrac{1}{2} x_{(3)}$$

that is a standard estimator used in the regulatory setup. There are many other ways of modifying ECDF, which leads to different quantile estimators; see [HF96] for other examples. ♦

Example 5.5 (ES plug-in estimator based on ECDF)**.** The ECDF defined in (5.2) could be also used to define estimators of other risk measures. If we consider ES_{α} and plug this estimator into distribution-based version of ES defined in (5.1), then we get

$$\begin{aligned}\hat{\mathrm{ES}}_{\alpha}^{\mathrm{emp}}(x) := R_{\mathrm{ES}_{\alpha}}(\hat{F}_X^{\mathrm{emp}}) &= -\frac{1}{\alpha}\int_0^{\alpha} x_{(\lfloor nt \rfloor + 1)}\,\mathrm{d}t \\ &= -\frac{1}{\alpha}\left(\frac{1}{n} \cdot \sum_{i=1}^{\lfloor n\alpha \rfloor} x_{(i)} + \left(\alpha - \frac{\lfloor n\alpha \rfloor}{n}\right) x_{(\lfloor n\alpha \rfloor + 1)}\right) \\ &= \frac{1}{n\alpha}\left(\sum_{i=1}^{\lfloor n\alpha \rfloor} x_{i:n} + (n\alpha - \lfloor n\alpha \rfloor) x_{(\lfloor n\alpha \rfloor + 1):n}\right), \end{aligned} \tag{5.4}$$

which is one of the common choices for ES estimator, see e.g. Article 11 in [EU24]. ♦

Similar procedure could be also applied in the parametric setup and allow us to recover well known formulas for parametric risk estimation under different distributional assumptions

Example 5.6 (Normal VaR and ES estimators)**.** Let us assume that a sample $x \in \mathbb{R}^n$ is from the normal distribution. In this case, the distribution of X could be estimated using empirical mean and variance via formula

$$\hat{F}_X^{\text{norm}}(t) := \Phi\left(\frac{t - \hat{\mu}(x)}{\hat{\sigma}(x)}\right),$$

where $\hat{\mu}(x) = \frac{1}{n}\sum_{i=1}^n x_i$ and $\sigma(x) = \sqrt{\frac{1}{n}\sum_{i=1}^n (x - \hat{\mu}(x))^2}$. Plugging this estimator into the distribution-based functional for VaR_α, we recover

$$\begin{aligned}
\hat{\text{VaR}}_\alpha^{\text{norm}}(x) &:= -\inf\left\{z \in \mathbb{R} \mid \Phi\left(\frac{z - \hat{\mu}(x)}{\hat{\sigma}(x)}\right) > \alpha\right\} \\
&= -\inf\left\{z \in \mathbb{R} \mid \frac{z - \hat{\mu}(x)}{\hat{\sigma}(x)} > \Phi^{-1}(\alpha)\right\} \\
&= -\inf\left\{z \in \mathbb{R} \mid z > \hat{\mu}(x) + \hat{\sigma}(x)\Phi^{-1}(\alpha)\right\} \\
&= -\left(\hat{\mu}(x) + \hat{\sigma}(x)\Phi^{-1}(\alpha)\right),
\end{aligned}$$

which is a standard normal VaR estimator, see [MFE10]. We note that this formula could be alternatively obtained by plugging the estimated parameters directly into the normal VaR formula in Example 3.7. Similarly, one can show that the formula for the plug-in gaussian estimator ES_α is given by

$$\hat{\text{ES}}_\alpha^{\text{norm}}(x) = -\hat{\mu}(x) + \hat{\sigma}(x)\frac{\phi(\Phi^{-1}(\alpha))}{\alpha},$$

which is again similar to the risk formula in Example 3.10. ♦

5.2 Risk estimators and their statistical properties

In this section, we recall some basic statistical concepts and link them to risk and performance measure estimators. Given a generic law-invariant risk measure ρ and position X, we are interested in checking statistical properties of the sequence $\rho_n(\boldsymbol{X})$, either for fixed $n \in \mathbb{N}$ or in the limit case when $n \to \infty$.

From the practical perspective, the non-parametric estimators are the most common choice. That saying, even for VaR and ES, one can find multiple propositions based on both plug-in approaches and their extensions. To better understand this, see survey paper [NZC14], where more than 45 estimation methods for ES alone are presented; 11 different non-parametric methodologies are introduced. For completeness, we present selected examples in Table 1.

In particular, all estimators considered in Table 1 are built on order statistics. Since order statistics are the base building block of non-parametric quantile estimator, it is not surprising that they are also building blocks for other estimators. In fact, one can show that if want to mimic properties of a CRM by estimator, understood as a function from $\mathbb{R}^n$ to $\mathbb{R}$, then we indeed need to incorporate linear combinations of order statistics.

Nr	Estimator	Comment
1	$\widehat{\mathrm{ES}}^1_\alpha(x) := \frac{-1}{\lfloor \alpha n \rfloor} \sum_{i=1}^{\lfloor \alpha n \rfloor} x_{(i)}$	Average tail loss ES estimator based on (3.3) and sample conditional mean.
2	$\widehat{\mathrm{ES}}^2_\alpha(x) := \frac{-1}{\alpha n} \left(\sum_{i=1}^{\lfloor \alpha n \rfloor} x_{(i)} + (\alpha n - \lfloor \alpha n \rfloor) x_{(\lfloor \alpha n \rfloor + 1)} \right)$	Non-parametric ES plug-in estimator for ECDF, see Example 5.5.
3	$\widehat{\mathrm{ES}}^3_\alpha(x) := \frac{-1}{\alpha(n+1)} \Big(\frac{3}{2} x_{(1)} + \sum_{i=2}^{M_6 - 1} x_{(i)} + \frac{1 + 2R_6 - R_6^2}{2} x_{(M_6)} + \frac{R_6^2}{2} x_{(M_6+1)} \Big)$	ES plug-in estimator for Type 6 quantile and flat extrapolation.
4	$\widehat{\mathrm{ES}}^4_\alpha(x) := \frac{-1}{\alpha(n+1)} \Big(\left(\frac{1}{2} + \frac{1}{1-\xi} \right) x_{(1)} + \sum_{i=2}^{M_6 - 1} x_{(i)} + \frac{1 + 2R_6 - R_6^2}{2} x_{(M_6)} + \frac{R_6^2}{2} x_{(M_6+1)} \Big)$	ES plug-in estimator for Type 6 quantile and Pareto-type extrapolation.
5	$\widehat{\mathrm{ES}}^5_\alpha(x) := \frac{-1}{M_6} \left(\frac{3}{2} x_{(1)} + \sum_{i=2}^{M_6} x_{(i)} \right)$	Conservative version of ES estimator 3 with restricted integration
6	$\widehat{\mathrm{ES}}^6_\alpha(x) := \frac{-1}{M_6} \left(\left(\frac{1}{2} + \frac{1}{1-\xi} \right) x_{(1)} + \sum_{i=2}^{M_6} x_{(i)} \right)$	Conservative version of ES estimator 4 with restricted integration.

Table 1: *Table presents six different non-parametric ES estimators; see [EBA23, Annex I] for details. For brevity, we use notation $M_6 := \lfloor \alpha(n+1) \rfloor$, $R_6 := \alpha(n+1) - \lfloor \alpha(n+1) \rfloor$. The parameter $\xi \in (0,1)$ corresponds to tail heaviness and needs to be pre-assessed before the estimation takes place.*

Theorem 5.7 (Robust representation of law-invariant coherent risk estimators)**.** A function $\hat{\rho}_n \colon \mathbb{R}^n \to \mathbb{R}$ is a law-invariant coherent risk estimator[a] if and only if there exists a set $M \subset \{a \in \mathbb{R}^n : \sum_{i=1}^n a_i = 1, a_1 \geq a_2 \geq \ldots \geq a_n \geq 0\}$ such that

$$\hat{\rho}_n(x) = \sup_{a \in M} \langle a, -s(x) \rangle, \quad x \in \mathbb{R}^n, \tag{5.5}$$

where $s(x) \in \mathbb{R}^n$ is the sorted version of x.

[a] For exact definition of a coherent risk estimator we refer to [Aic+25]. In a nutshell, the function should satisfy the properties linked to monotonicity, positive homogeneity, sub-additivity, and cash-additivity. In the estimation context, law invariance means that $\hat{(}x) = \hat{(}x')$ for any $x, x' \in \mathbb{R}^n$ such that x' is a permutation of x.

For the proof of Theorem 5.7 we refer to [Aic+25]. This theorem effectively states that one can utilize robust representations also in the estimation context and define estimators based on scenario sets.

Apart from representation theorem, we are also interested in generic statistical properties of the estimators. This includes properties such as consistency or unbiasedness.

Definition 5.8 (Consistent estimator)**.** A sequence of estimators $(\hat{\rho}_n)_{n=1}^{\infty}$ is *consistent* for a measure $\rho\colon L^\infty \to \mathbb{R}\cup\{+\infty\}$ if, for any $X \in L^\infty$ such that $\rho(X) < +\infty$, and i.i.d. sample $\boldsymbol{X} = (X_1, X_2, \ldots)$ from the distribution of X, we have

$$\hat{\rho}_n(\boldsymbol{X}_n) \xrightarrow{a.s.} \rho(X), \quad n \to \infty,$$

where $\boldsymbol{X}_n = (X_1, \ldots, X_n)$, and $\xrightarrow{a.s.}$ stands for $\mathbb{P}$-almost sure convergence[a].

[a] For simplicity, we focus on $\mathbb{P}$-a.s. convergence of the estimators, which corresponds to strong consistency in classical statistics. Nevertheless, most of the results can be extended to weaker forms of convergence, such as convergence in probability (weak consistency).

Definition 5.9 (Unbiased estimator)**.** An estimator $\hat{\rho}_n : \mathbb{R}^n \to \mathbb{R}$ is *unbiased* for a measure $\rho\colon L^\infty \to \mathbb{R}\cup\{+\infty\}$ if, for any $X \in L^\infty$ such that $\rho(X) < +\infty$, and i.i.d. sample $\boldsymbol{X}_n = (X_1, \ldots, X_n)$ from the distribution of X, we have

$$\mathbb{E}[\hat{\rho}_n(\boldsymbol{X}_n)] = \rho(X).$$

Similarly, we say that a sequence of estimators $(\hat{\rho}_n)_{n=1}^{\infty}$ is *asymptotically unbiased* (for ρ) if

$$\mathbb{E}[\hat{\rho}_n(\boldsymbol{X}_n)] \to \rho(X), \quad \text{as } n \to \infty.$$

One can easily show that, under specific continuity assumptions imposed on a risk measure, the non-parametric plug-in estimators based on ECDF are consistent and asymptotically unbiased. The non-asymptotic unbiased property is typically much harder to control in the non-parametric setup, apart from the mean. Furthermore, the financial interpretation, in the risk context, of the unbiasedness property is troublesome – we are effectively requiring that our capital reserve should be **on average** sufficient to cover our losses. Let us now introduce another estimation property which could be seen as a modification of the statistical bias.

Definition 5.10 (Risk unbiased estimator)**.** An estimator $\hat{\rho}_n : \mathbb{R}^n \to \mathbb{R}$ is *risk unbiased* for a measure $\rho\colon L^\infty \to \mathbb{R}\cup\{+\infty\}$ if, for any $X \in L^\infty$ and i.i.d. sample $\boldsymbol{X}_n = (X_1, \ldots, X_n)$ from the distribution of X, we have

$$\rho(\hat{\rho}_n(\boldsymbol{X}_n) + X) = 0.$$

Similarly, we say that a sequence of estimators $(\hat{\rho}_n)_{n=1}^{\infty}$ is *asymptotically risk unbiased* (for ρ) if

$$\rho(\hat{\rho}_n(\boldsymbol{X}_n) + X) \to 0, \quad \text{as } n \to \infty.$$

Note that Definition 5.9 and Definition 5.10 do not coincide, unless $\rho \not\equiv \mathbb{E}$. In Definition 5.9 we require that the position secured by the estimated capital, that is, secured position $S = \boldsymbol{X}_n + X$, does not bear risk in terms of ρ; note that both X and $\boldsymbol{X}_n$ are random. We refer to [PS18] for details and more information on risk unbiasedness. As we show later, the risk unbiasedness property plays an important role in regulatory backtesting, as it could be used to construct a good estimator of VaR performance measure. W7 ⊣ W8

5.3 Statistical sample construction and its impact on risk estimation

So far, we have assumed that we are given a sample from X and focused on how to estimate $\rho(X)$ given an i.i.d. sample $(X_1, \ldots, X_n)$. In practical situation, the proper sample construction is much more challenging task, even when compared with the actual choice of the risk estimator. In this section, we briefly comment on the key challenges that could be encountered in the process.

The construction of the sample $\boldsymbol{X}$ results from modeling choices that can be formalized as a *data-generating transformation*

$$\mathcal{T} : \mathcal{D} \longrightarrow \mathbb{R}^n,$$

mapping available risk factor *data set* $\mathcal{D}$ (historical time series, market simulations, filtered innovations, stress scenarios, *etc.*) into a sample that could be used for risk estimation. The quality of the resulting risk estimate $\hat{\rho}(\boldsymbol{X})$ therefore depends not only on the risk estimator $\hat{\rho}$ but also on $\mathcal{T}$. Please note that the set $\mathcal{D}$ is typically a high-dimensional set (e.g. our portfolio could be sensitive to moves of multiple EQ and FX risk factors) and the procedure typically involves construction of individual position scenario moves which are later aggregated to a portfolio level scenarios. As such, the chosen data-generating transformation should properly account for dependence between all objects to which our portfolio is sensitive to.

Let us now provide a high-level comment on selected aspects of statistical sample construction process and its impact on risk estimation. To ease the exposition, we focus on Market Risk capital reserves and P&L projection models often associating X_i with $\text{P\&L}_i$.

Remark 5.11 (Scenarios in risk models and their impact on sample construction)**.** While in this section we assume that X represents P&Ls linked to financial position moves, it might not always be the case. All methodological choices should be considered in reference to the underlying problem. For example, while the approach based on Historical Simulation might be best for 1-day P&L risk projections (assuming we have adequate data), it might be not plausible for longer holding periods or exposures projections based on simulation models, e.g. when counter-party credit risk exposures are evaluated. In fact, generation of forward-looking exposure scenarios (typically) requires utilization of SDE-based models and risk neutral measure repricing, which in turn require utilization of Monte Carlo methods. ♢

5.3.1 Overarching methodological split

The classical split in the risk literature (see e.g. [MFE10, Section 2.3] or [Ale09, Section IV.1.9]) is to classify the over-arching sample construction methodology into one of the following categorizes (to ease the exposition we are focusing on P&L sample generation models):

- **Variance-Covariance P&L construction methodology**: In this approach, it is assumed that all risk-factors changes encoded in data set $\mathcal{D}$ come from a *multivariate normal distribution*, and one can model their dynamics, including dependency structure, via mean vector and variance-covariance matrix. Furthermore, it is assumed that each trade P&L in the portfolio can be represented as a *linear function of the risk factor changes* (e.g. via trade's deltas). In that case, noting that the combination of margins from a multivariate normal vector is normal, one can estimate VaR of the whole portfolio by estimating it's mean and variance, which can be easily recovered from the over-arching mean and variance-covariance matrix estimates.

- **Historical Simulation P&L construction methodology**: In this approach, we are using historical time series from a given time-period to construct a joint historical sample for all risk factors changes. In this approach, the *shifted scenarios*, that is, a set of scenarios in which the current values of risk factors, say S_{current}, are combined with historical risk factor changes, say $(\Delta S_i)_{i=1}^n$ to constitute a potential future shifts, say $(\tilde{S}_i)_{i=1}^n$. Next, given shifted scenarios, we can compute shifted price for each trade in our portfolio, say $P(\tilde{S}_i)$, and reconstruct P&L sample by considering the difference between $P(\tilde{S}_i)$ and $P(S_{\text{current}})$, for $i = 1, \ldots, n$.

- **Monte Carlo P&L construction methodology**: This approach is an umbrella for all models in which the risk factors changes are assumed to come from a (semi)parametric model from which a statistical sample could be picked. As such, it could be seen as a combination of the previous two approaches in which we allow dependence structure to depend on the historical sample, but assume that we are still able to construct a Monte Carlo engine that will allow us to generate samples (even if we cannot derive explicitly the multivariate distribution imposed on all risk factors).

5.3.2 Historical data scenario construction

Note that in all approaches, we need to be able to construct a statistical sample from the historical data – either to estimate variance-covariance structure, directly simulate historical simulation P&Ls, or calibrate the distribution in the Monte Carlo framework. The exact construction methodology for *shifted scenarios* and/or *return rates* depends on the specific risk factor. For simplicity, let us focus on the Historical Simulation P&L methodology. The two most common approaches to scenario construction are:

- **Absolute (additive) returns**: In this approach, the shift is linked to the Bachelier type dynamics, in which asset price moves could be modeled by a Brownian motion or its extensions. Given a current (baseline) value of a risk factor, say S_{current} and a vector of historical prices $(S_0, \ldots, S_n)$, we define absolute returns $\Delta S_i := S_i - S_{i-1}$, $i = 1, 2 \ldots, n$, and the corresponding shifted historical scenarios given by

$$\tilde{S}_i := S_{\text{current}} + \Delta S_i, \quad i = 1, \ldots, n.$$

- **Relative (multiplicative) returns**: In this approach, the shift is linked to the Black-Scholes type dynamics, in which asset prices moves could be modeled by a geometric Brownian motion or its extensions. Given a current (baseline) value of a risk factor, say S_{current} and a vector of historical prices $(S_0, \ldots, S_n)$, we define relative returns $\Delta S_i := \frac{S_i - S_{i-1}}{S_{i-1}}$, $i = 1, 2 \ldots, n$, and the corresponding shifted historical scenarios given by

$$\tilde{S}_i := S_{\text{current}}(\Delta S_i + 1), \quad i = 1, \ldots, n.$$

  The construction of the corresponding P&L is the same as in (5.6)

Note that in the first approach the size of the shift does not depend on the current value of the risk factor, while in the second one, the move is proportional to the current size of the risk factor. The specific methodological choice depends on the nature of risk factor and/or the class of the underlying asset. In the literature, there are also other *mixed approaches* which combine both methodologies. We note that the relative approach implicitly assumes that the risk factor value is positive and bounded away from zero.

Once we have shifted, scenarios, we need to convert them to P&L. Assuming that a position is sensitive to the risk factors $(S^1, \ldots, S^k)$, we can model its P&L by constructing historical simulation scenarios

$$\text{P\&L}_i = P(\tilde{S}_i^1, \ldots, \tilde{S}_i^k) - P(S^1_{\text{current}}, \ldots, S^k_{\text{current}}), \quad i = 1, 2, \ldots, n, \tag{5.6}$$

where P is the pricing function for the underlying position (if the trade depends also on some non-random inputs that change through time, this could be incorporated in P). In particular, sometimes it is viable to use the delta-approximation instead of a full-revaluation. Namely, instead of re-repricing the trade using function P, we can recover the sensitivity of a position with respect to the moves of risk factors and construct a simplified P&L given by

$$\widetilde{\text{P\&L}}_i = \sum_{j=1}^{k} \delta_j (\tilde{S}_i^j - S_{\text{current}}^j), \quad i = 1, 2 \ldots, n.$$

where δ_j is the sensitivity of the trade with respect to jth risk factor moves (typically, it is delta-risk output from the pricing function P).[13]

Remark 5.12 (Overlapping scenarios)**.** Ideally, scenario returns (whether additive or relative) would form an i.i.d. sample. In practice, this assumption rarely holds. One must therefore account for intra-scenario dependence and for market-regime effects, e.g. the current level of volatility can distort the estimation process. A lack of independence may also arise from the chosen construction methodology. For instance, to improve statistical efficiency when estimating the distribution of 10-day future P&L, it is common to construct overlapping 10-day scenario and P&L samples from daily data. This procedure, however, induces serial correlation in the resulting sample that needs to be controlled; see [Ale09, p. IV.3.2.7] for illustration. ♢

The choice of the risk factor construction methodology depends on the underlying quote characteristics, asset class, market microstructure and quoting convention, etc. For information about the general methodological split for typical market risk factors we refer to [ECB18, Table 28].

5.3.3 Lookback period choice and weighting schemes

From statistical viewpoint, assuming that the sample generation process is stationary, the bigger the sample size, the better. On the other hand, when we are dealing with empirical data this could no longer be the case, and the choice of too long historical windows might lead to non-representable outputs. To better understand this, let us list some common pros and cons for short and long sample size choices:

- **Longer lookback windows** could reduce variance and statistical errors but may introduce structural bias if the return distribution is not stationary.
- **Shorter lookback windows** might increase variance and statistical errors but could better capture the current market conditions and market's regimes.

The exact choice of the lookback windows depends on the underlying problem, e.g., the short the holding period and the less extreme the underlying quantile, the shorted lookback period is acceptable. For standard market risk setup, with 10-day holding period and 1% VaR risk, the minimal lookback windows is annual (as required by the CRR regulations). The typical choice in this setup could range from 250 observations up to 750 observations; see e.g. [EBA25, Figure 10] where choices made by different financial institutions for Market Risk are confronted.

To account for the fact that more recent observations are more adequate, sometime a *weighting scenario scheme* could be introduced. Namely, for a scenario (or P&L) sample $(X_1, \ldots, X_n)$, we can introduce a weight vector $w = (w_1, \ldots, w_n) \in \mathbb{R}^n$ such that $w_i \geq 0$ and $\sum_{i=1}^n w_i = 1$ and weight

[13] Of course, we can also use higher-order approximation, e.g. via delta-gamma approach, as this is a simple Taylor based expansion. Note that the delta-approximation approach is effectively used in the variance-covariance methodology that pre-assumes linear relationship between risk factors and trades.

the scenarios accordingly. For instance, one could construct the weighted empirical cumulative distribution function (weighted ECDF) given by

$$\tilde{F}^{\text{emp}}_{w,X}(t) := \textstyle\sum_{i=1}^{n} w_i \mathbb{1}_{\{x_i \leq t\}},$$

and then use the plug-in risk estimation as explained in Section 5.1.

Example 5.13 (Plug-in VaR estimator for exponentially decaying weights)**.** Let us fix sample size $n \in \mathbb{N}$, decay parameter $\lambda \in (0,1)$, and consider the non-negative weight vector $w = (w_1, \ldots, w_n)$ given by

$$w_i := \frac{(1-\lambda)\lambda^{n-1}}{1-\lambda^n}.$$

Using basic geometric series arguments one can check that $\sum_{i=1}^{n} w_i = 1$. Then, using the same logic as in Example 5.3 for weighted ECDF, we get

$$\begin{aligned}
\hat{\text{VaR}}_\alpha(x) = R_{\text{VaR}_\alpha}(\tilde{F}^{\text{emp}}_{w,X}) &= -\inf\left\{ z \in \mathbb{R} \mid \sum_{i=1}^{n} w_i \mathbb{1}_{\{x_i \leq z\}} > \alpha \right\} \\
&= -\inf\left\{ z \in \mathbb{R} \mid \sum_{i=1}^{n} w_{(i)} \mathbb{1}_{\{x_{(i)} \leq z\}} > \alpha \right\} \\
&= -x_{(K_\alpha)},
\end{aligned} \tag{5.7}$$

where $w_{(i)}$ denotes the weight assigned to the ith worst element of the sample x, and

$$K_\alpha := \min\{k \in \mathbb{N} \colon \textstyle\sum_{i=1}^{k} w_{(i)} > \alpha\};$$

note that K_α depends on the sample realization, that is, on $x \in \mathbb{R}^n$. In other words, we sort the weights according to sample ordering, and then check for which worst case observations the cumulative sample weight exceeds threshold α. ♦

Apart from simple weighting schemes, there are also other approaches to risk estimation based on conditional parameter setup. e.g. when the conditional variance is estimated from the data and is plugged into in the risk formula. Those approaches are rarely used in high-dimensional Market Risk setup, where we deal with thousands of risk factors; for details, we refer to [Ale09]. w8 ⊣ w9

5.3.4 Holding period and confidence level scaling

Another common approach in risk measurement is to estimate risk over a holding period and confidence level for which sufficient data are available, and then project this estimate to the desired target configuration. Time (holding period) and confidence scaling is often achieved by incorporating a specific scalar factors but many alternative approaches (e.g. based on EVT) are available, see [PSS23] for details.

The usage of time scaling methods could be considered as a market standard that is utilized by many financial institutions, see [EBA25, Figure 9]. For time scaling of VaR_α (or ES_α), a commonly used heuristic is the square-root-of-time rule. This approach converts 1-day risk to h-day risk using a simple conversion rule

$$h\text{-day VaR}_\alpha = \sqrt{h} \cdot (\text{1-day VaR}_\alpha). \tag{5.8}$$

Despite its simplicity, such approach allow us to use (non-overlapping) 1-day scenarios instead of (overlapping) h-day scenarios in the estimation process, which often leads to performance increase.

Of course, this approach relies no multiple inherent assumptions and could lead to capital projection under-performance if not properly controlled, see e.g. [DZ06]. Also, note that (5.8) is interlinked with the (PH) property.

Example 5.14 (1-day to 10-day VaR scaling under normality)**.** Assume that X is accounting for a 10-day future P&L of a portfolio, which can be modeled as a sum of independent 1-day P&L moves denoted as $(Z_1, \ldots, Z_{10})$, where $Z_i \sim N(\mu, \sigma)$, for $i = 1, \ldots, 10$. Then, we know that $Z = Z_1 + \ldots + Z_{10} \sim N(10\mu, \sqrt{10}\sigma)$ so that we get

$$\mathrm{VaR}_\alpha(Z) = -10\mu + \sqrt{10} \cdot (\mathrm{VaR}_\alpha(Z_i) + \mu). \tag{5.9}$$

This shows that even in the (theoretical) normal setup the square-root-of-time rule should be used with care. Namely, while the relationship holds for $\mu = 0$ and is approximately accurate if $\mu \ll \sigma$ (that is, volatility dominates the trend), it does not hold in general. In the latter setup, the further we look into the future for the holding period, the more cautious we should be, as the trend might start to dominate the volatility. ♦

The processing of confidence level scaling (e.g. when extreme level quantiles is considered for relatively small sample size) could be also based on scalar factors, but there is no market consensus. For instance, one can use EVT based reasoning to estimate the shape of the tail or calculate the risk for less extreme level and then transfer it to any level using a pre-determined ratio from a reference distribution.

Example 5.15 (Confidence scaling for ES under normality)**.** Assume that $X \sim N(0, \sigma)$ for unknown parameter $\sigma \in \mathbb{R}_+$. Then, given $ES_\alpha(X)$ for $\alpha \in (0, 1)$, we can express $ES_\beta(X)$ for any $\beta \in (0, 1)$, as a function of $ES_\alpha(X)$. Indeed, using (3.5) we get

$$\mathrm{ES}_\beta(X) = \sigma \frac{\phi(\Phi^{-1}(\beta))}{\beta} = d(\alpha, \beta)\, \mathrm{ES}_\alpha(X),$$

where $d(\alpha, \beta) = \frac{\phi(\Phi^{-1}(\alpha))}{\phi(\Phi^{-1}(\beta))} \cdot \frac{\beta}{\alpha}$. Note that, unlike in the time scaling, the confidence scalar value $d(\alpha, \beta)$ depends on the underlying risk measure. For example, if we replace ES with VaR, the scalar would be equal to $\tilde{d}(\alpha, \beta) = \frac{\Phi^{-1}(\alpha)}{\Phi^{-1}(\beta)}$ ♦

For more information on time and confidence scaling, we refer to [PSS23].

5.4 Value-at-risk backtesting and exception rate performance measure

In the literature once can find multiple value-at-risk backtesting frameworks. Among them, arguably, the most important one is related to the number of overshoot count due to its regulatory relevance. In principle, VaR backtests (and also other risk measure backtests) could be broadly split into the following categories:

- **Unconditional coverage VaR backtests**: those tests examine whether the frequency of VaR exceedance (also called overshoots, violations, or breaches) is consistent with the nominal confidence level of the VaR model. For example, if the model estimates VaR at level 1%, then the expected violation probability for a single observation should be close to 1%. The most widely used framework is the *Kupiec Proportion of Failures (POF) test*, that is, a likelihood-ratio test that evaluates whether the observed number of violations follows a Bernoulli process with success probability equal to the VaR level. The statistical test could be based on both

both two-sided confidence interval (if we are interested in the measurement of the appropriateness) and the one-sided confidence interval (if we are interested in the conservativeness of the model). This approach detects models that systematically underestimate or overestimate risk, but does not consider the temporal structure of violations (for instance, linked to clustering). Consequently, unconditional coverage tests detect problems with the model, but they cannot be treated as omnibus tests.

- **Conditional coverage VaR backtests:** Conditional coverage tests extend unconditional coverage test by additionally assessing the independence of violations over time. A well-calibrated VaR model should not only produce the correct overall violation rate, but also generate violations that are serially independent. Clustering of violations indicates that the model does not react appropriately to changing market volatility or structural breakdowns. The canonical approach here is the *Christoffersen Conditional Coverage (CC) test*, which detects risk models exhibiting persistence or clustering of exceedance. The independent-check component test is essentially counting the number of transitions (between *exception* and *no exception*) and confront its with expected outcome using likelihood-ratio approach.

The detailed analysis of time series models and related temporal effects of risk measurement (detected by conditional coverage backtesting) is out of scope of this lecture notes; see [Ale09, p. IV.6.4.3] for details. For completeness, let us now describe in details a variant of the POF test that is used to assess model conservativeness. Let us assume that we are given a sequence of estimated capital reserves, say $\tilde{\mathrm{VaR}}_i$, as well as a sequence of the realized P&Ls, say $\mathrm{P\&L}_i$, for $i = 1, 2, \ldots, n$. We can combine them into a single secured position sample denoted by $y = (y_1, \ldots, y_n)$, where

$$y_i := \mathrm{P\&L}_i + \tilde{\mathrm{VaR}}_i.$$

Then, the *exception rate* backtesting statistic is defined as

$$T_n := \sum_{i=1}^{n} \frac{\mathbb{1}_{\{y_i<0\}}}{n}, \tag{5.10}$$

where $\mathbb{1}_A$ denotes the indicator function of the set A. In other words, we tally the number of exceptions (instances where capital is breached) in the secured sample and divide this count by n to obtain the empirical exception rate. Under the correct model specification, assuming that risk projections are equal to the true risks and observed P&Ls are independent, we should get $T_n \sim B(n, \alpha)$ which allows us to directly set the confidence thresholds for statistical tests.

For a significance level $\alpha = 1\%$ and a sample size of $n = 250$, the supervisory authority evaluates the adequacy of an internal model used for market risk capital determination by partitioning the possible outcomes into three distinct zones, based on the annual number of observed backtesting exceptions (nT_n).[14] By constructing an appropriate confidence interval around binomial distribution, the internal model (IM) is classified into one of the following zones:

- **green zone**: if the number of breaches is fewer than five, then under the correctly specified model this outcome is expected to occur in approximately 90% of all instances and corresponds to the event $T_n \in [0.00, 0.02)$.
- **yellow zone**: if the observed number of breaches falls within the interval from 5 to 9 inclusive, then, under the correctly specified model, this outcome is expected to occur in approximately 10% of all realizations and corresponds to test statistic values satisfying $T_n \in$

[14] Under a correctly specified model employing the 1% Value-at-Risk (VaR) as the reference risk measure, the expected number of exceptions over $n = 250$ trading days is approximately 2 to 3.

[0.02, 0.04). The left interval end point is also a conservative 5% confidence threshold cut-off since $F^{-1}_{B(1\%,250)}(0.95) = 5$.

- **red zone**: If there are 10 or more breaches, then under the correctly specified model this outcome should occur in fewer than 0.01% of all realizations and corresponds to $T_n \in [0.04, 1.00]$. The left interval end point is also a conservative 0.01% confidence threshold cut-off since $F^{-1}_{B(1\%,250)}(0.9999) = 10$.

Note that we employed the nominal count of breaches $(n \cdot T_n)$ for clearer exposition; for further information on Basel regulatory backtesting, see [BCB96].

The idea of counting the number of overshoots is very closely related to the performance measure we considered in Example 4.12. Let us now show that the exception rate statistic is indeed a performance measure dual to the VaR risk measure family, in which VaR is replaced by the empirical VaR estimators (based on ECDF).

Proposition 5.16. For a secured position $y = (y_1, \ldots, y_n)$, consider the test statistic T_n defined in (5.10). Then

$$T_n = \inf\{\alpha \in (0, 1] : \hat{\mathrm{VaR}}^{\mathrm{emp}}_\alpha(y) \le 0\}, \tag{5.11}$$

with the convention $\inf \emptyset = 1$, where $\hat{\mathrm{VaR}}^{\mathrm{emp}}_\alpha$ is empirical VaR plug-in estimator defined in (5.3).

Proof. Fix $n \in \mathbb{N}$ and $y = (y_i)_{i=1}^n \in \mathbb{R}^n$. Assuming $y_{(n)} \ge 0$, we obtain

$$\begin{aligned}
\inf\{\alpha \in (0, 1) : \hat{\mathrm{VaR}}^{\mathrm{emp}}_\alpha(y) \le 0\} &= \inf\{\alpha \in (0, 1) : -y_{(\lfloor n\alpha \rfloor+1)} \le 0\} \\
&= \tfrac{1}{n} \inf\{k \in (0, n) : y_{(\lfloor k \rfloor+1)} \ge 0\} \\
&= \tfrac{1}{n} \inf\{k \in \{0, 1, \ldots, n-1\} : y_{(k+1)} \ge 0\} \\
&= \tfrac{1}{n} \textstyle\sum_{k=1}^n \mathbb{1}_{\{y_{(k)}<0\}} \\
&= \textstyle\sum_{i=1}^n \frac{\mathbb{1}_{\{y_i<0\}}}{n}.
\end{aligned}$$

Conversely, if $y_{(n)} < 0$, then by the convention $\inf \emptyset = 1$ the statement follows. □

From Proposition 5.16, we infer that, for the purpose of evaluating IM performance, the regulator can be viewed as applying an empirical VaR estimator at various confidence levels $\alpha \in (0, 1]$ and then determining for which of these levels the secured position remains acceptable. Since the VaR family is monotone in the confidence level, we can identify the smallest level at which the position is still accepted. This value then serves as a measure of the position's performance. The traffic-light performance thresholds 0.02 and 0.04 are introduced to accommodate possible model misspecification, bias, model risk, and related issues, and are calibrated based on Bernoulli distribution specification. Equivalently, the regulator aims to verify that the model remains conservative when the confidence risk level is (slightly) increased.

Remark 5.17 (ES backtesting)**.** Expected Shortfall is much harder to backtest when confronted with VaR because ES concerns the average severity of losses in the tail, not just whether a loss exceeds a threshold. This means ES backtesting requires many extreme observations, yet tail events are rare, making empirical assessment statistically fragile. VaR, in contrast, produces a binary exception indicator, allowing simple and reliable coverage tests such as Christoffersen's. For this reason, under the FRTB framework regulators use ES as the capital measure – because it is theoretically superior – but rely on VaR-style exception counting for backtesting, since it remains the only method that is both statistic ally robust and operationally straightforward. This hybrid approach preserves the risk sensitivity of ES while retaining the proven practicality of VaR backtesting. For more information on the challenges linked to backtesting of ES we refer to [MP19]. ◊

Remark 5.18 (Good backtesting performance does not indicate good VaR estimator)**.** A simple yet illustrative example that demonstrates the limitations of the exceedance rate as a measure of estimator quality is as follows (see [HE14] for details). Consider $n = 250$, and assume that it is known that the P&L sample $(\text{P\&L}_i)_{i=1}^n$ is supported on the interval $[-0.95, 0.95]$. Then, constructing an estimator that takes the value 1 in 245 cases and the value -1 in the remaining five cases randomly yields perfect backtesting performance when evaluation is based solely on the exceedance rate. Consequently, while exception rate might be good to assess conservativeness of the estimator, it often fails int the comparative performance assessments; see Section 1.2 of [Gne11] for a more elaborate example. ◊

W9 ⊣ W10

5.5 Evaluation of point forecasts and mean score functions

As noted in Remark 5.18, good performance in backtesting might not necessarily imply that a given estimator is good, even in the base VaR setup. The remarkable article [Gne11] critically reviews the evaluation of point forecasts and bring many examples linked to mean, median, or quantile (VaR) forecasting methods. Let us now briefly recall the concepts of scoring functions and elicitability in the context of risk measurement. The concept of traces back to [OR85]. The goal of the approach is to provide a methodology which ensures truthful reporting, by making evaluated observations consistent with the underlying metric; see [Dav16] for an interesting discussion on the topic.

A *scoring function*, or simply a *score*, is a mapping $S : \mathbb{R}^2 \to \mathbb{R}_{\geq 0}$ measuring the deviation from the point forecast r (first argument) to the realization x (second argument). Of course, we need to impose special properties on S that depend on the type of point forecast we want to make, so that the output is meaningful.

Definition 5.19 (Consistent scoring function)**.** The scoring function $S : \mathbb{R}^2 \to \mathbb{R}_{\geq 0}$ is called *consistent* for distribution-based characteristic $R : \mathcal{D} \to \mathbb{R} \cup \{\infty\}$ if

$$\mathbb{E}[S(R(F_X), X)] \leq \mathbb{E}[S(r, X)], \quad X \in L^\infty. \tag{5.12}$$

for all $r \in \mathbb{R} \cup \{+\infty\}$. We say that the scoring function S is *strictly consistent* if it is consistent and equality in (5.12) holds only for $r = R(F_X)$.

We also say that a distribution-based measure R is elicitable if there is a strictly consistent scoring function for it. In summary, the mean value of a scoring function should be the smallest if our point forecast is equal to the true value of the distribution characteristic $R(F_X)$.

Example 5.20 (Squared error scoring function)**.** In predictive statistics, the mean value is arguably the most frequently employed distribution characteristic. In particular, it plays a central role in (standard) linear regression, where the loss or error function is constructed on the basis of the residual sum of squares. This is consistent with the framework of consistent scoring functions. Specifically, one can verify that the scoring function

$$S(r, x) = (r - x)^2 \tag{5.13}$$

is (strictly) consistent for the expectation functional, for any random variable $X \in L^\infty$.[15] being a standard example. In the *stochastic processes* and *linear models (econometrics)* lectures it has been already shown that the mean is the best forecast for L^2-distance so we skip the proof. ♦

Example 5.21 (Absolute error scoring function)**.** One can show that the function given by

$$S(r, x) = |r - x|$$

is consistent for the median. This is aligned with regression models in which *Least Absolute Deviation* metric is used to measure the error between the observed value and its projection. ♦

In our context, we are mainly interested in the consistent scoring functions for risk measures.

Proposition 5.22. The risk measure VaR_α is elicitable with *quantile score function*

$$S(r, x) = (\mathbb{1}_{\{r \geq x\}} - \alpha)(r - x), \tag{5.14}$$

being a strictly consistent scoring function for R_{VaR_α}.[a]

[a] In fact, there exists many more scoring function strictly consistent for R_{VaR_α}, see Theorem 2.6 in [Gne11].

We refer to [Gne11, Section 3.3] for the proof. In the next result, we focus on ES.

Proposition 5.23. The risk measure ES_α is not elicitable, that is, there exists no scoring function that is strictly consistent for R_{ES_α}.

For the proof, we refer to [Gne11, Theorem 3.5]. The results presented in Proposition 5.23 lead to an intensive academic discussion on whether ES is backtestable and whether it could be used in the predictive setup; this also explains why the regulatory update is used VaR or backtesting. Still, it should be pointed out that the current regulatory setup is not based on scoring rules (as exception rate corresponds to so called *identification function* rather than *scoring function*) and it has been shown that ES_α is in fact *jointly elicitable* with VaR_α in a same manner that variance is *jointly elicitable* with mean; see [FZG15] for details. In fact, a lot of competing ES_α backtesting approaches have been proposed in the literature, ranging from sum of VaR forecasts (for different confidence level) evaluation, *VaR-ES identification function* analysis, up to backtest based on the dual relationship between risk and performance measures. The detailed analysis of available ES backtesting measures is out of the scope of this lecture. In this context, it should be also noted that the the class of coherent and elicitable risk measures is fully characterized; it is the class of so called *expectile risk measures.* see [BB15]. Notably, the entropic risk measure, that is convex but not coherent, is elicitable, see [Emb+21].

In practice, the distribution of F_X and the true probability $\mathbb{E}$ present in the formula (5.12) is not known, so we need to estimate both F_X and the mean of the score function. Given a realized sample of risk projections, say $(r_i)_{i=1}^n$, and a realized sample of values, say $(x_i)_{i=1}^n$, we can define the empirical mean score by setting

$$\bar{S} := \tfrac{1}{n}\textstyle\sum_{i=1}^n S(r_i, x_i). \tag{5.15}$$

If we get two competing models producing point forecasts then we can compare their performance by confronting the resulting mean scores; the smaller the value of $\bar{S}$, the better the forecast.

[15] If we expand the space to include non-bounded random variables, the situation is more intricate. Namely, one can show that on L^1 the function (5.13) is consistent but not strictly consistent. It becomes strictly consistent if we restrict our space to L^2.

The backtests that aim to evaluate competing methodologies or estimators using scores similar to the one presented in (5.15) are often referred to as *comparative backtests* , as their goal is to compare different models rather than to assess whether the provided methodology is accurate. Of course, a comparison of the output should be made with care as blunt value check does not provide any answer about the statistical significance of the observed performance increase. That is why the analysis of (5.15) is often accompanied by the Diebold-Mariano testing framework, see Remark 5.25.

Remark 5.24 (Application to VaR backtesting)**.** In the context of VaR backtesting, the forecasts r_i will be taken by the considered risk estimators (i.e. we set $r_i := -\hat{\rho}_i$) and x_i will be the realized cash-flows linked to $\text{P\&L}_i$. Then, the quantile scoring function from (5.14) could be rewritten as

$$S(-\hat{\rho}, x_i) = \alpha(x_i + \hat{\rho}_i)^+ + (1-\alpha)(x_i + \hat{\rho}_i)^-. \tag{5.16}$$

We observe that this procedure results in a weighted penalty scheme: when the secured position is positive, the weight α is used, whereas for negative secured positions, the weight $(1-\alpha)$ is applied. Although motivated by the above considerations, this method does not have a direct connection to the backtesting frameworks currently used in regulation. It should also be emphasized that this procedure penalizes estimators that are overly conservative. ♢

Remark 5.25 (Diebold-Mariano test statistic)**.** The analysis of (5.15) is often accompanied by the Diebold-Mariano testing framework (or other similar tests) in which a version of a standard t-test is applied to the score values, see [DM02] or [PS18] for details. For completeness, let us present a simplified version of this test which does not take into account autocorrelation effects. Let $(r_i)_{i=1}^n$ and $(\hat{r}_i)_{i=1}^n$ be two competing point forecasts that are to be confronted with a sequence of realizations $(x_i)_{i=1}^n$. We define the score differentials sample $d = (d_1, \ldots, d_n)$, where

$$d_i := S(r_i, x_i) - S(\hat{r}_i, x_i), \quad i = 1, \ldots, n. \tag{5.17}$$

The Diebold–Mariano test is a statistical test of the null hypothesis indicating the equality of *mean scores* between the forecasting methods. Assuming that the sample is i.i.d. and imposing suitable moment conditions on d, the Diebold–Mariano statistic is the standard t-statistic given by

$$\text{DM} = \frac{\bar{\mu}(d)}{\bar{\sigma}(d)}, \tag{5.18}$$

and should be asymptotically standard normal, which allows the construction of confidence intervals and consequent statistical tests. When the scoring function S is strictly consistent, the null hypothesis of the equality of mean scores has a clear decision-theoretic interpretation: neither forecast dominates the other in terms of expected score, and hence neither forecast is preferred by a risk-neutral evaluator. We note that in practice the sample independence condition is often replaced by a weak stationarity conditions. In that case, the variance estimator used in (5.18) needs to be replaced with a HAC (heteroskedasticity and autocorrelation consistent) estimator also called Newey–West estimator, see [DM02] for details. Also, note that the choice of score function has a high impact on the values of score differentials in (5.17), and the consequent test power of Diebold-Mariano test. We refer to Section 5 in [Gne11] for a discussion about score function choice and its impact on Diebold-Mariano test power. ♢

5.6 Evaluation of density forecasts and PIT-backtests

In many applications of risk management, a model delivers a full *conditional predictive distribution* rather than a single point forecast, e.g. VaR or ES. Examples include P&L distributions produced by internal market risk models, simulation-based credit portfolio models, or stress-testing frameworks. The holistic statistical evaluation of such models requires tools that assess the adequacy of an entire forecast density rather than evaluation of risk or performance metrics outputs.

Let $(Y_t)_{t=1}^n$ denote a series of future observations, where $n \in \mathbb{N}$, and let us assume we have a series of estimated predictive distributions denoted by $(\hat{F}_t)_{t=1}^n$. In the standard market risk setup, $(Y_t)_{t=1}^n$ relates to 1-day P&Ls from our portfolio, while $\hat{F}_t$ represents ECDF build from (Historical Simulation) sample P&Ls for each day t using the data and scenarios available up to day $t-1$. Note that the true distribution of Y_t, that is, F_{Y_t}, as well as its estimator, that is, $\hat{F}_t$, could change in time. Our goal is to confront the realizations Y_t with the predictive distributions $\{\widehat{F}_t\}$ and to check if the predictive model is consistent with the data using a backtesting framework. This can be done utilizing the concept of the probability integral transform that was initially developed in [Fis34; Ros56] and later adjusted to market risk setup backtesting in [Ber01].

Definition 5.26 (Probability integral transform)**.** Let $Y \in L^\infty$ and let $\widehat{F}_Y$ be a distribution forecast of Y. The *probability integral transform (PIT)* of Y given $\widehat{F}$ is given by $U := \widehat{F}(Y)$.

Given $(Y_t)_{t=1}^n$ and $(\hat{F}_t)_{t=1}^n$, the PIT transform of our observations is given by $(U_t)_{t=1}^n$, where

$$U_t := \widehat{F}_t(Y_t).$$

Under the correct model specification, assuming that all the distributions are continuous, we expect that the sequence $(U_t)_{t=1}^n$ forms an i.i.d. sequence that is uniformly distributed on $[0, 1]$.

Proposition 5.27. Let $(Y_t)_{t=1}^n$ be a sequence of continuous random variables and let $(\hat{F}_t)_{t=1}^n$ be the corresponding sequence of forecast distributions. Assume that the sequence $(Y_t)_{t=1}^n$ is independent and that the forecasts are correct, that is, we have $\hat{F}_t = F_{Y_t}$, for $t = 1, \ldots, n$. Then,

$$U_t \sim \mathsf{U}(0,1) \quad \text{and} \quad (U_t)_{t=1}^n \text{ are independent,}$$

that is, $(U_t)_{t=1}^n$ forms an i.i.d. sequence that is uniformly distributed on $[0, 1]$.

Proof. As $\hat{F}_t = F_{Y_t}$, for $t = 1, \ldots, n$, the distribution forecasts are deterministic and they do not depend on the same historical samples or past observations from $(Y_t)_{t=1}^n$. Consequently, since the sequence $(Y_t)_{t=1}^n$ is independent, we immediately get that the sequence $(U_t)_{t=1}^n$ is also independent. Thus, it is sufficient to show that $U_t \sim \mathsf{U}(0, 1)$, for $t = 1, \ldots, n$. Fix $t \in \{1, \ldots, n\}$ and define generalized inverse $F_{Y_t}^{-1}(y) := \inf\{x \in \mathbb{R} : F_{Y_t}(x) \geq y\}$, for $y \in [0, 1]$. Since F_{Y_t} is continuous and non-decreasing, for any $y \in [0, 1]$, we get

$$\mathbb{P}(U_t \leq y) = \mathbb{P}(F_{Y_t}(Y_t) \leq y) = \mathbb{P}(Y_t \leq F_{Y_t}^{-1}(y)) = F_{Y_t}(F_{Y_t}^{-1}(y)).$$

By continuity of F_X, we have $F_X\big(F_X^{-1}(y)\big) = y$, for $y \in [0, 1]$, and consequently $F_{U_t}(y) = y$, which implies $U_t \sim U(0, 1)$ and concludes the proof. □

Remark 5.28 (Continuity assumption)**.** The continuity assumption imposed on the sequence $(Y_t)_{t=1}^n$ in Proposition 5.27 is necessary. If X has atoms, then $F_{Y_t}(Y_t)$ is no longer uniformly distributed. Still note that $F_{Y_t}(Y_t)$ would be stochastically larger than the uniform distribution. ♢

It is important to note that the result stated in Proposition 5.27 holds irrespective of the parametric form of the true data distribution and remains valid even when the forecast density changes over time. This result could be used to backtest the distributional forecast adequacy – the better our forecasts, the more uniform on $[0, 1]$ the PIT sequence $(U_t)_{t=1}^n$ should be. As in the VaR backtesting, the backtests are typically focused on either conditional or unconditional coverage, i.e., testing of the following properties:

- *Uniformity* of the PITs (correct marginal distribution),
- *Independence* of the PITs (correct dynamics).

This could be achieved by both graphical tools such as PIT histograms, that are widely used in practice, as well as formal statistical frameworks such as classical goodness-of-fit tests for uniformity (e.g. Kolmogorov–Smirnov, Cramér–von Mises, or Kuiper tests). While the goodness-of-fit tests are asymptotically valid, they tend to be low-powered in the sample sizes typically available for risk model backtesting. This issue is particularly acute when the focus lies on tail behavior, as is the case for market and credit risk.

These limitations motivate parametric testing approaches that exploit the structure implied by correct density forecasts or other integral transforms such as the Rosenblatt Gaussian transform, in which we a (potentially) uniform sequence $(U_t)_{t=1}^n$ into the Gaussian sequence. For instance, following [Ber01], given $(U_t)_{t=1}^n$, we can define the transformed sequence

$$Z_t := \Phi^{-1}(U_t)$$

where Φ denotes the standard normal distribution function. Under the correct model specification (null hypothesis of correct density forecasts), directly from Proposition 5.27, we get that

$$Z_t \sim \mathcal{N}(0, 1) \quad \text{and} \quad (Z_t)_{t \geq 1} \text{ is i.i.d.}$$

This transformation maps the problem into a Gaussian setting, allowing the use of goodness-of-fit or likelihood ratio (LR) tests with good finite-sample properties; we can also test for Gaussianity using many tests focused on tail-heaviness specification testing such as Jarque-Bera test, see [JB80].

Remark 5.29 (Regression setup for PIT backtest)**.** Apart from simple goodness-of-fit normality testing, a convenient alternative hypothesis is an auto-regressive Gaussian process,

$$Z_t - \mu = \rho(Z_{t-1} - \mu) + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma^2),$$

with null restrictions H_0 given by $\mu = 0$, $\sigma^2 = 1$, and $\rho = 0$. Note that the likelihood ratio tests on subsets of these restrictions allow separate assessment of unconditional calibration ($\mu = 0$, $\sigma^2 = 1$), dynamic misspecification ($\rho \neq 0$), or joint density adequacy. $\Diamond$

To sum up, density forecast evaluation generalizes classical backtesting by using the full predictive distribution rather than isolated quantiles. PIT-based methods provide a unifying framework in which calibration, independence, and tail behavior can be tested in a mathematically coherent manner. Likelihood-based tests of Gaussian PITs offer particularly strong power in realistic sample sizes and have become a benchmark tool in modern risk model validation. Those tests are now standard in density forecast evaluation and are commonly referred to as *PIT-based backtests*, *p-value backtests* or *density backtests*. Note that the need of holistic distribution assessment is also integrated into many regulatory frameworks, see e.g. Section 4.3.2 in Market Risk Chapters (CRR2 or CRR3) in [ECB18].

6 Selected topics in risk and performance measurement

In this section, we provide a generic information on selected topic linked to risk and performance measurement.

6.1 Time-consistency and dynamic measures

In many optimization problems, one needs to evaluate the risk or performance of a financial position over time and to take future-time controls into account. A canonical example is a multi-period portfolio optimization problem with discrete rebalancing dates. At each time, decisions are taken conditionally on the information available, and their quality must be assessed not only in terms of immediate outcomes, but also in terms of future risk and performance. This naturally leads to the notion of *dynamic risk and performance measures*, that is, families of conditional functionals indexed by time and adapted to the underlying information flow. Let us not very briefly introduce the mathematical setup, and comment on the key ideas that are utilized in the dynamic measurement framework.

We consider a discrete-time filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t=0,\dots,T}, \mathbb{P})$ with a finite time horizon $T \in \mathbb{N}$ and interpret $\mathcal{F}_t$ as the information available at time t. For brevity, we use notation $L^\infty_t := L^\infty(\Omega, \mathcal{F}_t, \mathbb{P})$, for $t = 0, \dots, T$, and assume that $\mathcal{F}_0 := \{\Omega, \emptyset\}$ and $\mathcal{F}_T = \mathcal{F}$, so that $L^\infty_t = L^\infty$ and $L^\infty_0 = \mathbb{R}$. As usual, random variables represent discounted terminal cash flows or P&Ls.

In this setup, we introduce the notion of a *dynamic measure* $\rho := \{\rho_t\}_{t=0}^T$, that is, a family of maps

$$\rho_t \colon L^\infty \to L^\infty_t$$

For each $X \in L^\infty$ and $t \in \{0, \dots, T\}$, the value $\rho_t(X)$ represents the risk or preference of a position X measured at time T; note that $\rho_0(X)$ recovers the usual (static) measure that we considered in the previous section. The axiomatic theory of risk measures could be adjusted to the dynamic setup. Instead of redefining all axioms, we provide an exemplary definition of a dynamic equivalent of a family of risk measures that was defined in previous chapter – coherent risk measures.

Definition 6.1 (Dynamic coherent risk measure)**.** We say that $\rho = (\rho_t)_{t=0}^T$ is a *dynamic coherent risk measure* (dCRM) if, for any $t \in \{0, \dots, T\}$, the mapping $\rho_t \colon L^\infty \to L^\infty_t$ is

1) *Monotone decreasing*, if $X \geq Y$ imply $\rho_t(X) \leq \rho_t(Y)$, for $X, Y \in L^\infty$;
2) *Cash additive*, if $\rho_t(X + m) = \rho_t(X) - m$, for $X \in L^\infty$ and $m \in L^\infty_t$;
3) *Positively homogeneous*, if $\rho_t(\lambda X) = \lambda \rho_t(X)$ for $X \in L^\infty$ and $\lambda \in L^\infty_t$ is such that $\lambda \geq 0$;
4) *Sub-additive*, if $\rho_t(X + Y) \leq \rho_t(X) + \rho_t(Y)$,for any $X, Y \in L^\infty$.

From Definition 6.1, we see that for ρ_t, the standard axioms have been replaced by their $\mathcal{F}_t$-conditional equivalents, as the risk measured in time t (from today's perspective) is a random variable from the space L^∞_t. Of course, one can also formulate the dynamic versions of the representation theorems, and reintroduce most object considered in the previous section in the dynamic setup, see e.g. [FP06].

The key new challenge which emerges in the dynamic setup is to ensure that the risk or performance measurement that are made in different points in time are consistent. Intuitively, a preference or risk assessment made today should not be contradicted by future assessments once additional information becomes available.

For risk measures, the aforementioned principle could be naturally translated into a specific risk measure axiom which is called strong time-consistency.

Definition 6.2 (Time-consistency)**.** A dynamic risk measure $\rho = (\rho_t)_{t=0}^T$ is called *strongly time-consistent* if for all $X, Y \in L^\infty$ and all $t < s$, we have

$$\rho_s(X) \leq \rho_s(Y) \qquad \Longrightarrow \qquad \rho_t(X) \leq \rho_t(Y).$$

The property in Definition 6.2 is essentially linked to Bellman principle of optimality, and could be restated using various dynamic frameworks, see [BCP17]. Indeed, for risk measures, time-consistency is equivalent to a recursive, or Bellman-type, structure and we can restate the definition using recursive condition

$$\rho_t(X) = \rho_t(-\rho_s(X)), \tag{6.1}$$

for all $t < s$; this relation is fundamental for multi-period optimization – see Remark 6.4. Surprisingly, the strong time-consistency axiom is often not satisfied, if we consider standard generalization of risk measures and one needs to be careful, when defining the families of risk measures. W10 ⊣ W11

Example 6.3 (Inconsistency of naive dynamic Expected Shortfall)**.** Let X be a terminal *profit* defined on a two-period binomial tree for $t \in \{0, 1, 2\}$. At time $t = 1$ the process moves to an upper node u or a lower node d, each with probability 1/2. At time $t = 2$ the terminal values of X are given by

Node	Terminal profits		
u	−10	12	14
d	−20	22	22

All terminal outcomes are assumed to be equally likely conditional on the corresponding node at time 1. We consider the filtration generated by the process with $\mathcal{F}_0 = \{\Omega, \emptyset\}$. Consider a dynamic version of Expected Shortfall at level $\alpha = 2/3$ given by

$$\rho_t(X) := \frac{1}{\alpha} \int_0^\alpha \mathrm{VaR}_u(X \mid \mathcal{F}_t)\, du,$$

where $\mathrm{VaR}_u(X \mid \mathcal{F}_t)$ is the conditional quantile; note that $\rho_0(X) = \mathrm{ES}_\alpha(X)$. First, it is easy to calculate the static ES. Indeed, for $\alpha = 2/3$, we get

$$\rho_0(X) = -\frac{(-20) + (-10) + 12 + 14}{4} = 1,$$

so that the position has positive risk. Now, let as calculate $\rho_1(X)$. At node u, the terminal values are −10, 12, 14, and the lower α-tail consists of −10 and 12, giving

$$\rho_1(X)(\omega) = -\frac{(-10) + 12}{2} = -1, \quad \text{for } \omega \in \Omega \text{ linked to node } u.$$

At node d, the terminal values are −20, 22, 22, and the lower α-tail consists of −20 and 22, giving

$$\rho_1(X)(\omega) = -\frac{(-20) + 22}{2} = -1 \quad \text{for } \omega \in \Omega \text{ linked to node } u.$$

Consequently, we end up with a non-rational risk assessment: the position has positive risk at time 0 and negative risk at time 1, and

$$\rho_0(X) = 1 \neq -1 = \rho_0(-\rho_1(X)).$$

For more details, refer to [RS07, Example 8.1]. ♦

While the notion of strong time consistency is well suited for dynamic risk measures, owing in particular to their cash-additivity (CA) property, it is not appropriate for performance or acceptability measures, which are typically assumed to satisfy scale invariance (SI). In fact, it can be shown that there exists no non-degenerate acceptability index that is strongly time consistent. As a consequence, alternative notions of time consistency have been proposed in the literature. A detailed discussion of these concepts is beyond the scope of this lecture; see [BCP17] for an extensive analysis. Also, one can show that, on infinite time horizon and with sufficiently rich filtration structure, the dynamic analog of the entropic risk measure is the only dynamic risk measure which is law invariant, strongly time-consistent, and relevant, see Remark 6.5. Finally, we note that the topic of time-consistency and consistent measurement of preferences is a subject of an on-going research with many recent contributions.

Remark 6.4 (Backward recursion and Bellman principle of optimality)**.** Strong time-consistency admits a constructive interpretation in discrete time that is linked to a one step recursive backward construction and standard Bellman programming principle, see e.g. [RS06]. Consider a finite discrete-time filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t=0}^T, \mathbb{P})$. Let $(\rho_t)_{t=0}^T$ be a dynamic risk measure. From (6.1) we see that the strong time-consistency is in fact equivalent to the existence of a backward recursive representation of the form $\rho_t(X) = \rho_{t,t+1}\big(-\rho_{t+1}(X)\big)$ for a one-step $\mathcal{F}_t$-conditional mappings $\rho_{t,t+1}$. For a finite time horizon, strong time consistency is typically achieved via a recursive construction based on the relation (6.1) with terminal condition $\rho_T(X) = -X$, and the dynamic programming principle. In this approach, the risk assessment at time t is defined in terms of the assessment at time $t+1$, that is, $\rho_t(X)$ is evaluated through $\rho_{t+1}(X)$ rather than directly through X. Consistency across time is then ensured by requiring the recursion (6.1); see [RS06]. This backward treatment is fundamental in applications, as it allows dynamic risk measures to be specified locally in time while guaranteeing global intertemporal coherence, see [CK11; Sha09]. $\diamondsuit$

Remark 6.5 (Dynamic Entropic utility as the only strongly-time consistent risk measure for any generic filtration)**.** It has been shown in [KS09] that the dynamic version of entropic utility is essentially the only law-invariant example of strongly time-consistent risk measure if we consider infinite time-horizon filtration that is suitably rich (e.g. it is generated by a proper i.i.d. sequence). The key reason for this uniqueness lies in the interaction between law-invariance and the recursive structure imposed by strong time-consistency. Law-invariance forces the risk measure to depend only on conditional distributions, while strong time-consistency requires these distributions to aggregate additively over time. Entropy is essentially the only divergence that is both additive under conditioning and compatible with convex duality, which explains why no other law-invariant dynamic risk measure can satisfy strong time-consistency in a non-degenerate way. This could be used to explain the popularity of the (entropy-based) risk-sensitive criterion in the stochastic control literature, see [DS99]. $\diamondsuit$

Remark 6.6 (Achieving time-consistency via precommitment or other methods)**.** When using risk or performance criterions that are not inherently time-consistent, one can enforce good and dynamically consistent behavior in optimization problems by restricting the set of actions or enforcing precommitment. In the latter approach, the decision-maker chooses at time 0 a complete sequence of future actions and evaluates performance only from the initial perspective, without (fully) re-optimizing at later dates. This approach is widely used in dynamic mean–variance portfolio optimization, where the variance term destroys time-consistency. We refer to [Cui+22] for a survey paper on this topic. We also note that under precommitment, the optimization problem remains well-posed and often admits explicit solutions, but the resulting strategy is dynamically inconsistent: if the investor were allowed to reconsider the problem at an intermediate time, the originally optimal strategy would generally no longer be optimal. Standard precommitment thus resolves the inconsistency by assumption, rather than by modifying the risk criterion itself, and should be interpreted as a modeling choice rather than a dynamic optimality principle (made at risk or performance objective constitution level). Note that there are also other approach to embed time-consistency into the optimization problem, e.g. based on subgame-perfect equilibrium definition or non-standard recursive constructions, see e.g. [BM14]. ♢

Remark 6.7 (The role of regularization in the context of time-consistency)**.** Regularization applied to the objective criterion is frequently employed in dynamic optimization and risk evaluation to improve stability, robustness, or tractability of the problem. Its role in achieving time-consistency, however, is more subtle. In a dynamic setting, regularization typically enters through a penalty term added to a nominal objective or risk functional. Examples include penalization of model uncertainty, transaction costs, control variation, or deviation from a reference model. Such regularization often improves numerical behavior and reduces sensitivity to small perturbations of the data. From the perspective of time-consistency, the crucial question is whether the regularization term is compatible with conditioning and backward recursion. This mechanism is well documented in risk-sensitive and robust control, see [HS01]. Note that, as expected, regularization by relative entropy or by divergence measures that satisfy a chain rule allows for dynamically consistent updates. Other popular choices such as variance or Wasserstein penalties, typically improve robustness but do not structurally resolve dynamic inconsistency. Those add-ons might rebuilt time-consistency within a specific subset of dynamic processes but (apart from relative entropy) they do not guarantee structural (strong) time-consistency. ♢

6.2 Risk contribution and incremental analysis

From the risk management perspective, we are often interested how much the risk of a predefined portfolio, say S, will change is we add or remove a specific trade from it, say X. This naturally leads to the concepts of incremental and limit risks, which shows the actual risk impact or marginal impact (sensitivity) of adding or removing position from a portfolio. Given a risk measure $\rho\colon L^\infty \to \mathbb{R}$, the *incremental risk* of X to S is usually computed by considering the increment

$$\rho(S+X)-\rho(S); \tag{6.2}$$

for the regulatory risk measures, this quantity is called Incremental VaR or Incremental ES and is often used for monitoring purposes. In many practical situations, we are often given a portfolio and we want to measure what are the incremental risks for each of the portfolio's constituents. For sensitivity-based analysis, the impact of the marginal limit in (6.2) is often considered, which naturally leads to the definition of risk contribution.

Definition 6.8 (Risk contribution)**.** Let $\rho : L^\infty \to \mathbb{R}$ and let $X, S \in L^\infty$. The *risk contribution* (or *marginal risk*) of X to S is defined by

$$\mathrm{RC}_\rho(X,S) := \frac{\partial}{\partial\varepsilon}\,\rho(S+\varepsilon X)\Big|_{\varepsilon=0}, \tag{6.3}$$

provided that the directional derivative exists.

The value (6.3) is often reported together with (6.2) for portfolio constituents to measure the overall portfolio sensitivity; note that similar values could be computed e.g. in reference to risk factor moves. For regulatory risk measures that satisfy (PH) property, there are formulas that could be used for computations under specific regularity conditions that are often implicitly assumed.

Example 6.9 (Risk contribution for VaR)**.** Let us consider $\rho = \mathrm{VaR}_\alpha$ for some $\alpha \in (0,1)$ and assume that (S, X) is absolutely continuous with full support. Then, we have

$$\mathrm{RC}_{\mathrm{VaR}_\alpha}(X, S) = \frac{\partial}{\partial \varepsilon} \mathrm{VaR}_\alpha(S + \varepsilon X)\Big|_{\varepsilon=0} = \mathbb{E}[-X \mid S = -\mathrm{VaR}_\alpha(S)]\,,$$

where $\mathbb{E}[\cdot \mid S = -\mathrm{VaR}_\alpha(S)]$ is a regular conditional expectation, see [MFE10, Lemma 6.26] for details. Given a Historical Simulation P&L sample, this quantity is often calculated by taking the VaR realizing scenario for S and checking the size of P&L of X under this scenario. ♦

Example 6.10 (Risk contribution for ES)**.** Let us consider $\rho = \mathrm{ES}_\alpha$ for some $\alpha \in (0,1)$ and assume that (S, X) is absolutely continuous with full support. Then, we have

$$\mathrm{RC}_{\mathrm{ES}_\alpha}(X, S) = \frac{\partial}{\partial \varepsilon} \mathrm{ES}_\alpha(S + \varepsilon X)\Big|_{\varepsilon=0} = \mathbb{E}[-X \mid S \leq -\mathrm{VaR}_\alpha(S)]\,.$$

We refer to [MFE10] for details. Given a Historical Simulation P&L sample, this quantity is often calculated by taking the tail scenarios that contribute to the estimation of ES and checking the size of averaged P&L of X under these scenarios. ♦

In the end of this section it is useful to note that the incremental risk computed in (6.2) could be seen as the approximation of the marginal risk from (6.3), for $\varepsilon = 1$. As the portfolios are typically large and we do not want to recompute the aggregated risk, the marginal impact are preferred, as adding small positions would not change the order of scenarios; one needs to be careful if a material position impact is considered as (6.2) could be preferred in that case.

6.3 Risk allocation and the Euler principle

In this section let us provide a brief introduction to the problem of risk allocation and comment on the most popular allocation principle; for simplicity we assume that we are given a pre-defined risk measure ρ – similar problem could be defined also for performance measures. Let us assume that we are given a random variable

$$S := \sum_{i=1}^{d} X_i,$$

that is, the sum of d-random variables $X_1, \ldots, X_d \in L^\infty$; in the financial context, S could be interpreted as the *aggregated portfolio* P&L while the sum constituents could correspond to P&L of single trades, business units, or desks. In this context, we refer to the quantity $\rho(S)$ as *aggregated risk*. We are often interested in the problem of *risk allocation*, which, as the name suggest, concerns the question of how a given aggregate risk should be distributed among sum constituents (several agents, business units, or time periods in a principled way). This question might naturally arise in insurance, finance, and capital regulation, where a firm assesses overall risk but must allocate capital or costs internally (e.g. to set limits or monitor actual risk exposures).

To do this, we need to define a specific *allocation principle* that will provide a set of numbers, say $(k_1, \ldots, k_d) \in \mathbb{R}^d$, corresponding to risk allocation for vector $(X_1, \ldots, X_d)$. The principles considered in the literature are often based on a set of normative properties that are associated with the notions of fairness, incentives, or optimality. For brevity, let us now present a couple of properties that are often assumed for allocation rules:

- **Full allocation**: allocated numbers sum to total risk, that is, $\rho(S) = \sum_{i=1}^{d} k_i$;

- **Symmetric allocation**: shifting does not disturb allocation, that is, for permuted vector $(X_{s_1}, \dots, X_{s_d})$ allocation problem we should get allocations $(k_{s_1}, \dots, k_{s_d})$.

- **Riskless allocation**: for non-random allocations, the value equals its deterministic contribution, that is, if $X_i = c$, then $k_i = -c$.

- **Additive for independent risks allocation**: If $X_1, \dots, X_d$ are independent, then we should get $k_i = \rho(X_i)$.

The detailed study of allocation axioms is out of scope of the lecture notes, see [MFE10, Section 6.3] and references therein. We note there is no universally *correct* allocation principle. Axioms reflect what one wants to preserve: fairness, incentives, marginality, or stability over time. Different choices lead to different internally consistent—allocation rules. Let us now introduce, arguably, the most widely used allocation rule.

Definition 6.11 (Euler allocation principle)**.** Let $\rho : L^\infty \to \mathbb{R}$ be a monetary risk measure that is (PH) and let $S = \sum_{i=1}^d$, where $X_i \in L^\infty$, for $i = 1, \dots, d$. Then, we call the vector $(k_1, \dots, k_d) \in \mathbb{R}^d$ the *Euler allocation principle* (for ρ and S), where

$$k_i := \mathrm{RC}(X_i, S) \qquad i = 1, \dots, d,$$

provided that $\mathrm{RC}(X_i, S)$ exists for $i = 1, \dots, d$.

Essentially, Euler allocation distributes total risk according to marginal risk contributions, and positive homogeneity guarantees that these contributions add up exactly to the total risk, that is, full allocation is satisfied. Let us now consider a couple of example in a rather informal setup.

Example 6.12 (Euler allocation for VaR and ES)**.** From Example 6.9 and Example 6.9 we immediately get the values of risk allocations for VaR and ES. In both cases, the total risk principle check is immediate. Indeed, for VaR_α, we get

$$\begin{aligned}\sum_{i=1}^d RC_{\mathrm{VaR}_\alpha}(X_i, S) &= \sum_{i=1}^d \mathbb{E}[-X_i \,|\, S = -\mathrm{VaR}_\alpha(S)] = \mathbb{E}\left[-\sum_{i=1}^d X_i \,\middle|\, S = -\mathrm{VaR}_\alpha(S)\right] \\ &= \mathbb{E}[-S \,|\, S = -\mathrm{VaR}_\alpha(S)] = \mathrm{VaR}_\alpha(S);\end{aligned}$$

the calculations for ES_α are similar. We also note that for both VaR and ES, the allocation is symmetric, and riskless. ♦

Example 6.13 (Euler allocation for multivariate normal distribution)**.** Assume that the marginal vector $(X_1, \dots, X_d) \sim \mathcal{N}(\mu, \Sigma)$ is a multivariate normal vector and that ρ is a (PH), (LI), and differentiable on the class of normal distributions, e.g. $\rho = \mathrm{VaR}_\alpha$ or $\rho = \mathrm{ES}_\alpha$, for $\alpha \in (0, 1)$. Then, recalling that the risk is proportional to variance, we can recover the Euler allocation assigned to X_i and get

$$k_i = \frac{\mathrm{Cov}(X_i, S)}{\mathrm{Var}(S)} \big(\rho(S) - \mathbb{E}[S]\big) + \mathbb{E}[X_i].$$

Note that k_i is a regression-based risk sharing rule: the total risk premium of the portfolio is distributed among components in proportion to their regression coefficients with respect to the aggregate loss. Each component is charged according to how much it explains fluctuations of the total risk, rather than according to its standalone volatility. ♦

Remark 6.14 (Euler principle does not always allocate risk)**.** The existence of risk contributions does not, by itself, guarantee that total risk can be decomposed into a sum of contributions. This requires additional structural properties of the risk measure. For instance, consider the mean–variance functional

$$\rho(S) = \mathbb{E}[S] + \lambda \ \mathrm{Var}(S), \qquad \lambda > 0.$$

Although ρ is convex and directionally differentiable, it is not positively homogeneous. The marginal risk contributions does not satisfy full allocation even for an extreme case $S = X_1$. Indeed, we get

$$\begin{aligned}
\mathrm{RC}_\rho(S,S) &= \frac{\partial}{\partial \varepsilon}\, \rho(S+\varepsilon S)\Big|_{\varepsilon=0} \\
&= \lim_{\varepsilon\to 0} \frac{\mathbb{E}[S+\varepsilon S] + \lambda\, \mathrm{Var}(S+\varepsilon S) - (\mathbb{E}[S] + \lambda\, \mathrm{Var}(S))}{\varepsilon} \\
&= \mathbb{E}[S] + \lim_{\epsilon\to 0} \frac{((1+\varepsilon)^2-1)\lambda\, \mathrm{Var}(S)}{\epsilon} \\
&= \mathbb{E}[S] + 2\lambda \ \mathrm{Var}(S) \neq \rho(S).
\end{aligned}$$

This example shows that marginal risk is a local sensitivity concept, while risk allocation is a global property. Positive homogeneity of the risk measure is essential for marginal contributions to recover total risk, even in the one-dimensional case. ♢

Remark 6.15 (Risk contribution versus risk allocation)**.** It should be emphasized that risk contribution and risk allocation describe related but distinct concepts. Risk contribution is a local notion that measures the marginal impact of a component on total risk, typically defined via directional derivatives. Risk allocation, by contrast, is a global decomposition of total risk into component-wise quantities that satisfy additional axioms, such as full allocation or symmetry. In particular, marginal risks need not sum to the total risk. The Euler principle provides a bridge between the two notions: for positively homogeneous risk measures, risk contributions coincide with a valid risk allocation and satisfy the full allocation property. Absent such structure, risk contributions should be interpreted as sensitivity measures rather than allocated capital. This is important, e.g. for the entropic utility as the underlying risk measure. ♢

7 Selected applications of risk and performance measures

This summary section illustrates how risk and performance measures enter mathematical models as structural building blocks. The aim is not to introduce new axioms or objects, but to show how maps introduced in this lecture notes could be used in various frameworks.

7.1 Risk and performance measures in various disciplines

Risk and performance measures arise in a variety of disciplines, either as direct measurement tools, decision criteria under uncertainty, or supporting metrics. Note that that are often replacing or refining expected-value-based rules. For completeness, let us present exemplary applications together with top-level comments.

Finance and banking. As already extensively discussed, in banking regulation risk measures such as VaR and ES are used to determine capital buffers. The estimated risks are used for regulatory reporting purposes (e.g. contributing directly to RWAs), capital ratios, and internal management limits. For completeness, we list selected applications below. Note that the Basel III regulatory framework is structured around three complementary pillars (Pillar 1, Pillar 2, Pillar 3) that jointly define how risks are measured, capital is held, and risk information is disclosed in order to ensure the solvency and stability of banking institutions. While a detailed description of the Basel framework is out of scope of this lecture, we provide a few top-level examples.

- **Regulatory capital reporting for traded risk (Pillar 1).** This pillar establishes minimum capital requirements for market risk and counter-party credit risk.

  For **market risk**, if a bank follows the internal models approach, the base market risk capital is determined using VaR- and ES-based metrics in various regulatory contexts. The key IMA risk charge components under Basel III are: (1) **Expected Shortfall Capital Charge (ES)**: Expected Shortfall at the 97.5% confidence level is the base capital metric. It is typically estimated using an annual historical window on which 10-day P&Ls are constructed, with subsequent adjustments to reflect position liquidity. A rolling-window approach is used to identify the stress period that maximizes the ES estimate, leading to the notion of *stressed Expected Shortfall*; (2) **Default Risk Charge (DRC)**: this component complements ES by capturing jump-to-default (JTD) risk. DRC models are typically based on multi-factor copula frameworks, with the reference metric being VaR at the 99.9% confidence level over a one-year horizon; (3) **Non-Modellable Risk Factors Risk Charge (NMRF)**: risk factors lacking sufficient real price observations are classified as non-modellable and excluded from ES calculations; **Residual Risk Add-On (RRAO)**: specific risks not adequately captured elsewhere are capitalized using separate, often regulatory prescribed, methodologies.

  For **counter-party credit risk**, VaR and ES are not used to measure default risk capital, which is instead based on one-year loss models driven by **Probability of Default (PD)**, **Exposure at Default (EAD)**, and **Loss Given Default (LGD)**. However, VaR-type concepts appear indirectly in exposure modeling and explicitly in the CVA capital charge, where counter-party credit risk is treated as a form of market risk. In particular, internal models for EAD often rely on quantiles of exposure distributions: Potential Future Exposure (PFE) is defined as a high quantile of future exposure at a given horizon. This represents a VaR of exposure rather than a VaR of losses and does not enter capital as a risk measure in the sense used for market risk.

- **Supervisory assessment for traded risk (Pillar 2).** This pillar complements Pillar 1 by requiring banks to assess whether regulatory capital adequately captures the full spectrum of traded risk and to hold additional capital where necessary. From a modeling perspective, Pillar 2 relies on internal risk measures applied to the same market and counter-party credit loss distributions as in Pillar 1, but typically under more conservative assumptions. This includes the use of **higher confidence quantiles**, **alternative stress scenarios**, **longer effective holding periods**, **more conservative liquidity assumptions**, and **explicit capital add-ons** for risks insufficiently captured by regulatory ES- and VaR-based metrics, such as concentration risk, model risk, valuation uncertainty, and liquidity risk. While Pillar 1 capital is formula-driven, Pillar 2 capital reflects supervisory judgment informed by internal risk measures and stress testing frameworks; the estimation methodologies are not directly prescribed by regulation. This includes **regulatory stress testing**, such as EBA stress tests, which is a Pillar 2 supervisory tool used to assess capital adequacy under adverse scenarios; the public disclosure of stress-test results falls under Pillar 3.

- **Product control and valuation adjustments.** In trading and product control functions, risk measures are used to quantify valuation uncertainty and model risk rather than capital adequacy. In particular, valuation adjustments such as Prudent Valuation Adjustment (PVA), Additional Valuation Adjustment (AVA), and Funding Valuation Adjustment (FVA) are often constructed using distributional properties of valuation errors, model parameter uncertainty, or funding cost exposures. From a modeling perspective, these adjustments can be interpreted as risk measures applied to distributions of valuation outcomes, with conservative quantiles or tail expectations used to ensure prudent valuation. Unlike market risk capital, these measures are applied at the level of individual trades or portfolios and over valuation-relevant horizons, and their purpose is to ensure balance-sheet prudence and P&L integrity rather than regulatory solvency.

- **Internal risk management and performance measurement.** Beyond regulatory capital calculations, risk and performance measures play a central role in banks' internal risk management frameworks. VaR, ES, and related tail-risk metrics are routinely used to set trading limits, allocate economic capital, and monitor risk consumption at the level of desks, portfolios, and business lines. In this context, risk measures are applied to shorter horizons and higher frequencies (e.g. daily or intraday P&L) than those used for regulatory capital, and are often complemented by stress losses and scenario-based measures. Performance measures, such as risk-adjusted return metrics, are then used to assess profitability relative to the amount of risk consumed, supporting capital allocation, pricing, and management decision-making. Unlike regulatory risk measures, internal metrics are institution-specific and optimized for timeliness and risk sensitivity rather than regulatory comparability.

- **Model monitoring, validation, and risk governance.** Risk measures are also central to model monitoring, independent model validation, and risk governance frameworks. VaR and ES are used as benchmarking and diagnostic tools in backtesting, sensitivity analysis, and model comparison, rather than as decision criteria. In this context, discrepancies between model-implied and realized risk measures are interpreted as indicators of model risk and inform remediation actions, model overlays, or governance escalations rather than capital calculations.

Insurance and actuarial science. In insurance and actuarial applications, risk measures are primarily used to quantify solvency risk arising from uncertain claim payments and to determine capital buffers required to ensure the insurer's ability to meet its obligations to policyholders. Unlike banking applications, risk measures are not applied to short-horizon trading P&Ls, but to *aggregate loss and liability distributions* over medium- to long-term horizons. Performance measures play a secondary role, while acceptability conditions and capital adequacy dominate. For completeness, we list selected applications below. Modern insurance regulation is built around risk-based capital frameworks, such as Solvency II in the European Union, which determine capital requirements using explicit risk measures applied to insurance balance-sheet quantities. The key regulatory quantities are the Solvency Capital Requirement (SCR) and the Minimum Capital Requirement (MCR). Again, note that, in contrast to banking, insurance risk measures are driven primarily by liability risk rather than short-term market fluctuations.

- **Regulatory capital determination (Solvency II).** Under Solvency II, the SCR is defined as a tail risk measure applied to the insurer's one-year change in basic own funds. In practice, this corresponds to the distribution of net losses arising from underwriting risk, market risk, credit risk, and operational risk over a one-year horizon. The standard approach defines the SCR via a high-quantile risk measure (typically VaR at the 99.5% confidence level), while internal models may replace the standard formula subject to supervisory approval.

- **Internal models and economic capital.** Insurers using internal models apply risk measures directly to simulated distributions of aggregate claims, premium income, asset values, and technical provisions. These models explicitly account for claim frequency and severity, claim development over time, and dependence across lines of business. The resulting risk measure defines internal economic capital, which may differ from regulatory capital and is used for strategic decision-making.

- **Reserving risk, technical provisions, and reinsurance design.** Risk measures are applied to the distribution of future claim payments arising from existing policies, in particular for long-tailed lines such as liability or workers' compensation. In this setting, the underlying random variable is the cumulative run-off loss of outstanding claims over their lifetime, projected from one-year claim development models. Tail risk measures are applied to the distribution of reserve deterioration used to quantify reserve uncertainty and to determine risk margins added to best-estimate technical provisions, reflecting the cost of holding capital against non-hedgeable insurance risk.

  Reinsurance contracts are then evaluated as risk-mitigating instruments by comparing risk measures applied to net aggregate losses before and after risk transfer. In practice, VaR or ES at high confidence levels are computed on the distribution of one-year reserve changes or ultimate claim payments under different reinsurance structures (e.g. excess-of-loss or quota share). The reduction in the selected risk measure provides a quantitative assessment of capital relief, while joint analysis of net losses and counter-party exposure allows insurers to balance solvency improvement, concentration risk, and reinsurance cost.

- **Own Risk and Solvency Assessment (ORSA).** Beyond minimum regulatory capital, insurers are required to conduct an Own Risk and Solvency Assessment, in which internal risk measures and stress scenarios are applied to loss and balance-sheet projections over multi-year horizons. ORSA analyses typically incorporate more conservative assumptions and severe but plausible scenarios, and are used to assess solvency under adverse conditions rather than to compute binding capital requirements.

- **Pricing and risk-adjusted performance.** Risk measures enter insurance pricing through risk loadings and capital allocation to lines of business. Risk-adjusted performance measures, such as return on risk capital, are used to evaluate underwriting profitability relative to the capital consumed by different products or portfolios, supporting pricing, portfolio steering, and strategic planning.

- **Model validation and governance.** Risk measures are used in actuarial model validation to assess the stability and sensitivity of loss and reserve models. Changes in VaR or ES estimates under alternative assumptions, parameter perturbations, or data subsets are interpreted as indicators of model risk rather than as capital drivers, and inform governance decisions such as model overlays or methodology changes.

Economics and decision theory. In economic models with ambiguity, model uncertainty, or incomplete preferences, risk and performance measures provide decision criteria that do not rely on an expected value. In this context, risk measures are often interpreted as *certainty equivalents*, mapping uncertain outcomes into deterministic values while accounting for downside risk or ambiguity. Performance measures are used to rank alternatives when preferences are not fully specified. The detailed overview on how and when risk and performance maps are used in economics and decision theory is out of scope of these lecture notes but, for completeness, let us list three exemplary applications:

- **Robust decision-making under ambiguity.** Risk measures are applied to payoff distributions under multiple probability models, with the decision-maker evaluating actions based on worst-case or conservative outcomes. This leads to max–min or distributionally robust formulations in which risk measures replace expected utility as the objective criterion.
- **Dynamic decision problems and time consistency.** In intertemporal choice models, dynamic risk measures are used to evaluate future payoffs in a time-consistent manner, replacing conditional expectations in Bellman equations and allowing for precautionary behavior under uncertainty.
- **Equilibrium and welfare analysis.** Risk measures enter equilibrium models as agents' evaluation functionals, affecting asset prices, risk premia, and welfare comparisons when preferences are non-linear or ambiguity-averse. Performance measures can be used to compare equilibria or policy outcomes without assuming a common utility function.

Engineering, operations research, and portfolio optimization. In engineering and operations research, risk measures are used to model uncertainty in system performance, operating costs, and constraint violations. In contrast to financial applications, risk measures are typically applied to *cost*, *delay*, or *failure* random variables rather than profit-and-loss. Performance measures are used to compare designs or policies under uncertainty. Selected applications include:

- **Risk-constrained optimization.** Risk measures are used to replace probabilistic constraints by deterministic risk constraints of the form $\rho(X) \leq c$, where X represents uncertain cost, demand, or system load. Typical applications include power system operation, supply chain design, and infrastructure planning, where high-impact rare events must be controlled explicitly.
- **Risk-sensitive objective functions.** Instead of minimizing expected cost, decision-makers optimize risk-adjusted objectives in which risk measures are applied directly to total operating cost or system loss. This allows asymmetric penalization of extreme outcomes and leads to solutions that are more robust than expectation-based designs.
- **Stochastic and robust control.** In sequential decision problems, dynamic risk measures are used to evaluate cumulative costs over time, replacing conditional expectations in dynamic programming equations. This framework captures risk-averse or safety-oriented behavior in control systems, such as inventory management, energy systems, and transportation networks.
- **Portfolio optimization beyond mean–variance.** In portfolio optimization, risk measures such as VaR and ES are applied to portfolio returns or losses to define risk constraints or objective functions, generalizing the classical mean–variance framework. This allows for the incorporation of non-Gaussian returns, downside risk, and tail dependence in portfolio construction.
- **Performance evaluation and trade-off analysis.** Performance measures are used to compare alternative designs, control policies, or portfolios when multiple criteria are present. Risk-adjusted performance metrics enable systematic trade-off analysis between efficiency, robustness, and safety.

Other applications and connections to physics, information theory, and learning. Risk and performance measures also appear in a range of disciplines outside economics and engineering, often through their interpretation as certainty equivalents or nonlinear expectation operators. A prominent example is the *entropic risk measure*, which arises naturally from exponential utility and admits a representation in terms of relative entropy. Closely related functionals appear in statistical physics as free-energy or log-partition functionals, linking risk-sensitive evaluation to large-deviation principles. Beyond physics and information theory, risk measures are increasingly used in data-driven decision-making, including reinforcement learning and online optimization. In these settings, risk measures are applied to cumulative reward or cost distributions to penalize rare but severe outcomes and to learn risk-averse policies. In all these contexts, risk measures are not interpreted as capital or safety margins, but as structural tools for aggregating uncertainty, controlling tail behavior, and shaping learning dynamics.

7.2 Objectives based on risk and performance measures

In many optimization problems under uncertainty, the objective functional is no longer given by an expected value or an expected utility, but is instead constructed directly from a risk or performance measure. This section reviews the most common ways in which risk and performance measures are used to define optimization objectives and discusses their mathematical structure.

Throughout this section, let $\mathcal{X} \subset L^p(\Omega, \mathcal{F}, \mathbb{P})$ denote a set of admissible random outcomes (e.g. terminal payoffs, results of admissible portfolio strategies, costs, or losses).

- **Risk minimization and performance maximization problems.** The most direct use of a risk measure ρ is as an objective functional itself. One considers optimization problems of the form

  $$\inf_{X \in \mathcal{X}} \rho(X) \quad \text{or} \quad \sup_{X \in \mathcal{X}} \alpha(X)$$

  for instance when ρ is the standard risk measures or α represents expected return or a risk-adjusted performance index. These formulations replace classical expected-value optimization by criteria that are sensitive to tail behavior, downside risk, or acceptability. If ρ is convex and $\mathcal{X}$ is convex, the minimization problem is a convex optimization problem. If α is concave, the corresponding maximization problem is concave. Often, positive homogeneity of ρ or α implies scale-invariance properties of optimal solutions, while non-smoothness (as in the case of ES or VaR) naturally leads to subgradient-based optimality conditions.

- **Expected value optimization under risk constraints.** An alternative formulation separates performance and risk by retaining the expectation as the objective and imposing risk as a constraint. One considers problems of the form

  $$\sup_{X \in \mathcal{X}} \mathbb{E}[X] \quad \text{subject to} \quad \rho(X) \leq c,$$

  for a prescribed constant $c \in \mathbb{R}$; operator $\mathbb{E}$ could be of course replaced by generic mapping *alpha*. Conversely, one may minimize risk under a performance constraint, for example

  $$\inf_{X \in \mathcal{X}} \rho(X) \quad \text{subject to} \quad \mathbb{E}[X] \geq m,$$

  for some $m \in \mathbb{R}$,. Such formulations are particularly natural in regulatory, engineering, and safety-critical contexts, where admissible risk levels are specified externally. From a mathematical perspective, constrained formulations are closely related to penalized problems via Lagrange multipliers and lead naturally to risk-adjusted objective functions.

- **Risk-adjusted objective functions.** A common way to construct scalar objectives is to combine reward and risk in a single functional. Given a parameter $\lambda > 0$, one considers

$$\sup_{X \in \mathcal{X}} \left[\mathbb{E}[X] - \lambda\, \rho(X) \right],$$

  This formulation generalizes the classical mean–variance criterion by allowing for general, possibly asymmetric and tail-sensitive risk measures. If ρ is convex, then the underlying objective function is concave, and maximization problems admit well-posed optimality conditions. Varying λ yields a family of optimal solutions that can be interpreted as an efficient frontier balancing reward and risk, with λ acting as a shadow price of risk.

- **Acceptability-based formulations.** Risk measures induce acceptance sets of the form $\mathcal{A} := \{X \in L^p : \rho(X) \leq 0\}$. optimization problems can therefore be formulated directly in terms of acceptability, for instance

$$\sup_{X \in \mathcal{X} \cap \mathcal{A}} \mathbb{E}[X],$$

  or, more generally, using families of acceptance sets indexed by performance levels. Such formulations emphasize feasibility and geometric structure over numerical risk values and are particularly useful in constrained optimization, regulatory settings, and robust decision-making.

The structure of the admissible set $\mathcal{X}$ plays a crucial role in determining the nature of the resulting optimization problem. Depending on how $\mathcal{X}$ is defined, the same objective functional may lead to structurally different optimization frameworks, including static portfolio optimization, optimal stopping problems, or optimal policy selection in controlled stochastic systems. Typical optimization areas in which risk and performance measure based objective functions are actively used include:

- **Portfolio optimization.** In portfolio optimization, the set $\mathcal{X}$ consists of terminal portfolio values or portfolio returns generated by admissible trading strategies. Typically, $\mathcal{X}$ is defined via self-financing strategies subject to budget, leverage, or trading constraints, and may incorporate market frictions such as transaction costs or trading constraints. Risk and performance measures are applied to the resulting terminal payoffs or cumulative gains to construct risk-constrained or risk-adjusted portfolio selection problems.

- **Stochastic control and planning.** In stochastic control and planning problems, $\mathcal{X}$ consists of cumulative cost or reward functionals associated with admissible control processes over a finite or infinite horizon. Depending on the formulation, these costs may be discounted, averaged, or aggregated over time. Risk and performance measures are used to evaluate these random costs, resulting in control problems that explicitly penalize tail events or extreme outcomes.

- **Optimal stopping problems.** In optimal stopping problems, the set $\mathcal{X}$ is generated by admissible stopping times. Each stopping time defines a random payoff or cost, such as the value of a process observed at the stopping time or the accumulated cost until stopping. Risk and performance measures are applied to these random outcomes, leading to stopping rules that account for downside risk or variability rather than maximizing expected value alone.

- **Robust and distributionally robust optimization.** In robust and distributionally robust optimization, the set $\mathcal{X}$ reflects both admissible decisions and uncertainty about the underlying probability model. Risk measures are applied to worst-case or ambiguity-adjusted loss distributions, often leading to min–max or saddle-point formulations. Such problems naturally arise when the decision-maker seeks solutions that perform reliably across a family of plausible models.

The formulations presented above provide the foundations for **dynamic and stochastic optimization problems** in which time consistency plays a central role and recursive construction of optimal policies becomes possible; the detailed analysis of those problems is out of scope of these lecture notes. We illustrate this point in the following example in which we introduce the objective function often used in the long-run risk sensitive stochastic control setup. In the second example, we also show how risk and performance measure can structurally impact the whole framework build-up, so that the stated objective is inherently depending on the choice of the risk and performance metric.

Remark 7.1 (Robust representations and optimization)**.** When a risk measure ρ admits a robust (dual) representation, then risk-based optimization problems can be reformulated as min–max problems. For instance, for CRMs, we get,

$$\inf_{X\in\mathcal{X}} \rho(X) \;=\; \inf_{X\in\mathcal{X}} \sup_{Q\in\mathcal{Q}} \mathbb{E}_Q[-X],$$

which admits an interpretation as a zero-sum game between a decision-maker and an adversarial model selector. Under suitable convexity and compactness assumptions, min–max theorems allow the interchange of inf and sup, leading to saddle-point characterizations of optimal solutions. From an optimization perspective, such reformulations are advantageous: they replace a nonlinear objective functional by a family of linear expectations, facilitate the use of convex duality and decomposition techniques, and provide a natural foundation for robust and distributionally robust optimization as well as risk-sensitive stochastic control. ♢

7.3 Example: Averaged long-run risk sensitive stochastic control

Let us consider a discrete-time Markov Decision process (MDP) on a finite state-action space and briefly outline the intuition (logic) behind this model, we refer to [BR11] for details. Note that MDP could be seen as an extension of a Markov Chain defined on some *state space*: in every transition step, we can choose a transition matrix from a given set and the choice is determined by a element from the *actions space*. We use $E = \{x_1, \ldots, x_k\}$ and $U = \{\tilde{u}_1, \ldots, \tilde{u}_l\}$, where $k, l \in \mathbb{N}$, to denote the state space and the action space, respectively. We use $(X_n)_{n\in\mathbb{N}}$ to denote a sequence of random variables with values in the state space E that reflect the consequent states of the controlled process. In every step, to control transition from X_n to X_{n+1}, we can pick any action $a \in U$ that leads to a transition matrix

$$P^a := [p_a(x_i, x_j)]_{i,j=1}^k,$$

and consequent one-step transition probabilities $\mathbb{P}^a(x, A) := \sum_{y\in A} p_a(x, y)$, $A \in 2^E$. We use Π to denote the family of all (measurable) decision rules for all steps, and each $\pi \in \Pi$ is linked to a controlled probability measure $\mathbb{P}^\pi$ defined on $(\Omega, \mathcal{F})$; this typically corresponds to the canonical measurable space $(E^{\mathbb{N}}, (2^E)^{\mathbb{N}})$.

Our goal is to choose $\pi \in \Pi$ that would lead to optimal control and the cost of our actions is quantified via so called *cost function* $c\colon E \times U \to \mathbb{R}$. The top-level transition logic is as follows:

1. at time $i \in \mathbb{N}$ we are in state X_i, that is, E-valued and $\mathcal{F}_n$-measurable random variable;

2. we choose the action a_i that is, U-valued and $\mathcal{F}_n$-measurable random variable;

3. we pay the cost of our choice that is equal to $c(X_i, a_i)$;

4. the process X_i transitions to X_{i+1} according to $\mathbb{P}^{a_i}$.

Averaged per unit of time risk-sensitive objective function To correctly quantify the total cost, we need to aggregate all (random) costs and then apply some performance criterion. In the averaged (per unit of time) risk-sensitive stochastic control one considers the aggregated sum of all costs and the entropic utility measure imposed on the accumulated cost which lead to the objective function of the form

$$J_\gamma(x,\pi) := \liminf_{n\to\infty} \tfrac{1}{n}\tfrac{1}{\gamma} \ln \mathbb{E}_x^\pi \left[e^{\gamma \sum_{i=0}^n c(X_i,a_i)} \right] \tag{7.1}$$

where $x \in E$ corresponds to the process stating point, $\pi \in \Pi$ is the control strategy, $\mathbb{E}_x^\pi$ is the expectation under $\mathbb{P}^\pi$ under additional condition that $X_0 = x$ (the process had deterministic start in x), and $\gamma < 0$ is the risk-aversion coefficient. The usage of the $\liminf$ in (7.1) is linked to the fact that, without additional ergodicity or irreducibility assumptions imposed on the controlled transition kernels, the entropic utility for accumulated cost need not converge. Under standard assumptions, such as bounded costs and the existence of a stabilizing stationary policy inducing an ergodic Markov chain, the limit exists, is finite, and is in fact independent of the initial state, see [DS99] for details.

The popularity of the objective criterion of the form (7.1) is motivated by the unique properties of the entropic utility measure linked e.g. to cash-additivity and strong time consistency which also indirectly ensure existence of solutions.

Bellman equation and the usage of entropic utility Under additional ergodic conditions imposed on the underlying transition matrices, the optimal solution to problem (7.1) could be encoded in a single risk-sensitive Bellman equation given by

$$w(x) = \sup_{a\in U} \left[c(x,a) - \lambda + \frac{1}{\gamma} \ln \int_{y\in E} e^{\gamma w(y)} \mathbb{P}^a(x,y) \right], \tag{7.2}$$

in which we look for a pair (w, λ), that is, the function $w \in E \to \mathbb{R}$ (defined up to additive constant) and the constant $\lambda \in \mathbb{R}$. The equation (7.2) imposes an ergodic equilibrium on the system, and there are many techniques that allow us to find solutions to it. Assuming that such a pair exists, one can show that $\lambda = \sup_\pi J_\gamma(x,\pi)$, that is, λ recovers the optimal problem value, and we can reconstruct optimal solution by iterating equation (7.2) for solution w. Note the constant λ admits an interpretation as the optimal long-run certainty equivalent cost per unit of time.

The logarithmic–exponential structure in (7.2) is not incidental. It is precisely the cash-additivity and multiplicative aggregation of the entropic utility that ensure strong time consistency of the corresponding dynamic risk measure. For generic convex or coherent risk measures, an analogous stationary Bellman equation typically fails to exist, reflecting the dynamic inconsistency phenomena discussed in Section 6.1.

Of course, one can also replace entropic utility by standard expectation (by considering limit case $\gamma = 0$) and consider the risk-neutral formulation of the problem. While the detailed analysis of the risk-sensitive stochastic control problems is out of scope of these lecture notes, it is worth noting that the entropic utility measure is an integral part of this framework and many results could be shown by utilizing the normative-properties of the entropic utility, see also [Ste99] for applications in the portfolio optimization context.

7.4 Example: Acceptability good-deal bounds and market efficiency evaluation

In this example we illustrate how risk and performance measures can be used not only to quantify the risk of random payoffs, but also to induce economically meaningful *pricing bounds* in incomplete markets. The key idea is that prices should be compatible with the risk preferences encoded by a chosen risk or performance measure. This leads to a natural definition of bid and ask prices and provides a quantitative notion of market efficiency.

Let $\alpha : L^\infty \to \mathbb{R}$ be a fixed coherent acceptability index admitting representation

$$\alpha(X) = \sup\left\{\gamma \in \mathbb{R}_+ : \inf_{\mathbb{Q}\in\mathcal{Q}_\gamma} \mathbb{E}_{\mathbb{Q}}[X] \geq 0\right\}, \quad X \in L^\infty,$$

for a family of probability measures $(\mathcal{Q}_\gamma)_{\gamma\in\mathbb{R}_+}$, see (4.4) for details. Throughout this example we interpret random variables as discounted future profit–and–loss (P&L) values, consistently with the convention used in the previous sections. Let $X \in L^\infty$ denote a random payoff. We interpret X as the future P&L generated by holding one unit of a financial position.

Let us assume that investors on the market are interested in financial positions for which acceptability exceeds some predefined level $\gamma_0 > 0$; recall that acceptability indices are generalizations of Share Ratio, so we can say that investors want to take only positions that are attractive for them.

Ask and bid prices induced by an acceptability index. The *ask price* (seller's price) of X is defined as the smallest price $m \in \mathbb{R}$ such that the investment position $X - m$ is acceptable, that is, it creates an investment opportunity and we have

$$\pi^{\text{ask}}(X) := \inf\{m \in \mathbb{R} : \alpha(X - m) \geq \gamma\} \tag{7.3}$$

Similarly, the *bid price* (buyer's price) of X is defined as the largest price m that an agent is willing to pay while keeping the investment attractive

$$\pi^{\text{bid}}(X) := \sup\{m \in \mathbb{R} : \alpha(m - X) \geq \gamma\} \tag{7.4}$$

These definitions depend only on the properties of the performance measure α and do not rely on replication or completeness assumptions. In particular, the resulting prices generally form an bid-ask price interval $[\pi^{\text{bid}}(X), \pi^{\text{ask}}(X)]$. As α ia a coherent acceptability inded, we immediately get that

$$\pi^{\text{ask}}(X) = \sup_{Q\in\mathcal{Q}_\gamma} \mathbb{E}_Q[X], \qquad \pi^{\text{bid}}(X) = \inf_{Q\in\mathcal{Q}_\gamma} \mathbb{E}_Q[X],$$

so that we can interpret the prices and the expectation under the worst and best scenario from the set $\mathcal{Q}_\gamma$. These bounds are commonly referred to as *good–deal bounds*, and they exclude prices that would correspond to non-attractive investment opportunities relative to performance encoded by α. The size of the interval between bid and ask prices reflects the level of model uncertainty and risk aversion: larger sets $\mathcal{Q}_\gamma$ lead to wider bounds, while tighter bounds correspond to stronger pricing discipline; for complete markets the set $\mathcal{Q}_\gamma$ reduced to a singleton that denoted the pricing measure. In practice, the prices are also defined in terms of allowable (risk) hedges that can further increase investor performance, see [MC10].

Acceptability indices as market efficiency measures. The construction above generalizes classical performance-based pricing ideas. In particular, if α is chosen as a Sharpe-type ratio, then the condition $\alpha(X - m) \geq \gamma$ corresponds to requiring that the secured position achieves a minimal risk-adjusted return. In this sense, good–deal bounds exclude trades with excessively large but finite Sharpe ratios, extending the classical no-arbitrage paradigm which only rules out strictly positive payoffs.

Suppose that a market quotes a price $p(X)$ for the position X. We say that this price is compatible with the acceptability criterion α at level γ if

$$\pi^{\text{bid}}(X) \leq p(X) \leq \pi^{\text{ask}}(X).$$

Prices outside this interval correspond to investment opportunities that are either too attractive or insufficiently attractive relative to the chosen performance threshold. Consequently, market efficiency becomes a quantitative, measure-dependent notion: different choices of α and γ lead to different admissible price ranges, reflecting heterogeneous preferences or regulatory constraints. Note that by observing bid and ask prices we can also try to infer the level of acceptability on the market.

Acceptability indices and market arbitrage A fundamental structural link between acceptability indices and market viability is also provided by their behavior on non-negative positions. For coherent acceptability indices it holds that

$$X \geq 0 \text{ a.s. and } \mathbb{P}[X > 0] > 0 \quad \Longrightarrow \quad \alpha(X) = +\infty.$$

Consequently, any trading strategy (be it on buy or on sell side) that produces a non-negative discounted P&L which is strictly positive with positive probability corresponds to an investment opportunity with infinite acceptability. In this sense, coherent acceptability indices detect classical arbitrage through the explosion of performance. In an arbitrage-free market, all attainable discounted P&L positions necessarily have finite acceptability. Good–deal bounds can therefore be interpreted as a quantitative refinement of the no-arbitrage principle: while no-arbitrage excludes positions with infinite acceptability, performance-based pricing excludes positions whose acceptability exceeds a prescribed finite level. This provides a natural bridge between classical viability conditions and performance-based valuation in incomplete markets. Formally, if $\mathcal{X}$ denotes the set of attainable discounted P&L positions (e.g. from self-financed strategies), then the no-arbitrage condition implies $\sup_{X\in\mathcal{X}} \alpha(X) < \infty$, while the good–deal constraint imposes $\sup_{X\in\mathcal{X}} \alpha(X) \leq \gamma_0$. For details, see [CM09].

References

[Ace02] C. Acerbi. “Spectral measures of risk: a coherent representation of subjective risk aversion”. In: *Journal of Banking & Finance* 26.7 (2002), pp. 1505–1518.

[Aic+25] M. Aichele et al. “Coherent estimation of risk measures”. In: *arXiv preprint arXiv:2510.05809* (2025).

[Ale09] C. Alexander. *Market Risk Analysis*. Vol. 1-4. John Wiley & Sons, 2009.

[Art+97] P. Artzner et al. “Thinking coherently”. In: *Risk* 10 (1997), pp. 68–71.

[Art+99] P. Artzner et al. “Coherent measures of risk”. In: *Math. Finance* 9.3 (1999), pp. 203–228.

[BB15] F. Bellini and V. Bignozzi. “On elicitable risk measures”. In: *Quantitative Finance* 15.5 (2015), pp. 725–733.

[BCB96] BCBS. *Supervisory framework for the use of 'backtesting' in conjunction with the internal models approach to market risk capital requirements*. Tech. rep. Bank for International Settlements, 1996.

[BCP17] T. R. Bielecki, I. Cialenco, and M. Pitera. “A survey of time consistency of dynamic risk measures and dynamic performance measures in discrete time: LM-measure perspective”. In: *Probability, Uncertainty and Quantitative Risk* 2.1 (2017), p. 3.

[Ber01] J. Berkowitz. “Testing density forecasts, with applications to risk management”. In: *Journal of Business & Economic Statistics* 19.4 (2001), pp. 465–474.

[BL00] A. Bernardo and O. Ledoit. “Gain, Loss, and Asset Pricing”. In: *J. of Polit. Econ.* 108 (2000), pp. 144–172.

[BM14] T. Bjork and A. Murgoci. “A theory of Markovian time-inconsistent stochastic control in discrete time”. In: *Finance and Stochastics* 18.3 (2014), pp. 545–592.

[BP03] T. R. Bielecki and S. R. Pliska. “Economic Properties of the Risk Sensitive Criterion for Portfolio Management”. In: *Review of Accounting and Finance* 2 (2003), pp. 3–17.

[BP99] T. R. Bielecki and S. R. Pliska. “Risk-sensitive dynamic asset management”. In: *Applied Mathematics & Optimization* 39.3 (1999), pp. 337–360.

[BR11] N. Bäuerle and U. Rieder. *Markov Decision Processes with Applications to Finance*. Springer Science & Business Media, 2011.

[CGM01] P. Carr, H. Geman, and D. Madan. “Pricing and hedging in incomplete markets”. In: *J. Fin. Econ.* 62.1 (2001), pp. 131–167.

[CH09] P. Cogneau and G. Hübner. “The 101 ways to measure portfolio performance”. In: *Available at SSRN 1326076* (2009).

[CK11] P. Cheridito and M. Kupper. “Composition of time-consistent dynamic monetary risk measures in discrete time”. In: *International Journal of Theoretical and Applied Finance* 14.01 (2011), pp. 137–162.

[CK12] P. Cheridito and E. Kromer. “Reward-risk ratios”. In: *Journal of Investment Strategies* (2012).

[CM09] A. S. Cherny and D. B. Madan. “New measures for performance evaluation”. In: *The Review of Financial Studies* 22.7 (2009), pp. 2571–2606.

[Cui+22] X.-Y. Cui et al. “Survey on multi-period mean–variance portfolio selection model”. In: *Journal of the operations research society of china* 10.3 (2022), pp. 599–622.

[Dav16] M. Davis. “Verification of internal risk measure estimates”. In: *Statistics & Risk Modeling* 33 (2016), pp. 67–93.

[Del02] F. Delbaen. “Coherent risk measures on general probability spaces”. In: *Advances in finance and stochastics*. Springer, 2002, pp. 1–37.

[DM02] F. X. Diebold and R. S. Mariano. “Comparing predictive accuracy”. In: *Journal of Business & economic statistics* 20.1 (2002), pp. 134–144.

[DS99] G. B. Di Masi and Ł. Stettner. “Risk-sensitive control of discrete-time Markov processes with infinite horizon”. In: *SIAM Journal on Control and Optimization* 38.1 (1999), pp. 61–78.

[DZ06] J. Danielsson and J. P. Zigrand. “On time-scaling of risk and the square-root-of-time rule”. In: *Journal of Banking & Finance* 30.10 (2006), pp. 2701–2713.

[EBA23] EBA. *Final report – Draft RTS on the assessment methodology under which competent authorities verify an institution’s compliance with the internal model approach as per Article 325az(8) of Regulation (EU) No 575/2013 (Capital Requirements Regulation 2 - CRR2)*. Tech. rep. EBA/RTS/2023/05. European Banking Authority, 2023.

[EBA25] EBA. *EBA report results from the 2024 Market Risk benchmarking exercise – Part 1 - IMA*. Tech. rep. EBA/REP/2025/11. European Banking Authority, 2025.

[ECB18] ECB. *ECB guide to internal models, Risk-type-specific chapters*. Tech. rep. European Central Bank, 2018.

[Emb+21] P. Embrechts et al. “Bayes risk, elicitability, and the Expected Shortfall”. In: *Mathematical Finance* 31.4 (2021), pp. 1190–1217.

[EU24] EU. *Commission Delegated Regulation (EU) 2024/397 of 20 October 2023 supplementing Regulation (EU) No 575/2013 of the European Parliament and of the Council with regard to regulatory technical standards on the calculation of the stress scenario risk measure*. OJ L 2024/397, 5 February 2024. Official Journal of the European Union. 2024.

[EW15] P. Embrechts and R. Wang. “Seven proofs for the subadditivity of expected shortfall”. In: *Dependence Modeling* 3.1 (2015), pp. 126–140.

[Fis34] R. A. Fisher. “Statistical methods for research workers.” In: (1934).

[Fis88] P. C. Fishburn. *Utility theory*. Wiley Online Library, 1988.

[FP06] H. Föllmer and I. Penner. “Convex risk measures and the dynamics of their penalty functions”. In: *Statist. Decisions* 24.1 (2006), pp. 61–96.

[FS02a] H. Föllmer and A. Schied. *Stochastic finance: An introduction in discrete time*. de Gruyter Studies in Mathematics 27, 2002.

[FS02b] H. Föllmer and A. Schied. “Convex measures of risk and trading constraints”. In: *Finance and stochastics* 6.4 (2002), pp. 429–447.

[FZG15] T. Fissler, J. F. Ziegel, and T. Gneiting. “Expected shortfall is jointly elicitable with value at risk - implications for backtesting”. In: *Risk magazine* December (2015).

[Gne11] T. Gneiting. “Making and evaluating point forecasts”. In: *Journal of the American Statistical Association* 106.494 (2011), pp. 746–762.

[HE14] H. Holzmann and M. Eulert. “The role of the information set for forecasting—with applications to risk management”. In: *Annals of Applied Statistics* 8.1 (2014), pp. 595–621.

[HF96] R. J. Hyndman and Y. Fan. “Sample Quantiles in Statistical Packages”. In: *The American Statistician* 50.4 (1996), pp. 361–365.

[HS01] L. P. Hansen and T. J. Sargent. “Robust control and model uncertainty”. In: *American Economic Review* 91.2 (2001), pp. 60–66.

[JB80] C. M. Jarque and A. K. Bera. “Efficient tests for normality, homoscedasticity and serial independence of regression residuals”. In: *Economics Letters* 6.3 (1980), pp. 255–259.

[Jor07] P. Jorion. *Value at risk: the new benchmark for managing financial risk*. Vol. 2. McGraw-Hill New York, 2007.

[JST06] E. Jouini, W. Schachermayer, and N. Touzi. “Law invariant risk measures have the Fatou property”. In: *Advances in mathematical economics*. Springer, 2006, pp. 49–71.

[Kre88] D. M. Kreps. *Notes on the Theory of Choice*. Vol. 2. Westview press Boulder, 1988.

[KS09] M. Kupper and W. Schachermayer. “Representation results for law invariant time consistent functions”. In: *Mathematics and Financial Economics* 2.3 (2009), pp. 189–210.

[Kus01] S. Kusuoka. “On Law Invariant Coherent Risk Measures”. In: *Adv. Math. Econ.* (3 2001), pp. 83–95.

[MC10] D. B. Madan and A. Cherny. “Markets as a counterparty: an introduction to conic finance”. In: *International Journal of Theoretical and Applied Finance* 13.08 (2010). Publisher: World Scientific, pp. 1149–1177.

[MFE10] A. J. McNeil, R. Frey, and P. Embrechts. *Quantitative risk management: concepts, techniques, and tools*. Princeton university press, 2010.

[MP19] F. Moldenhauer and M. Pitera. “Backtesting expected shortfall: a simple recipe?” In: *Journal of Risk* 22.1 (2019), pp. 17–42.

[MTZ11] L. MacLean, E. Thorp, and W. Ziemba. *The Kelly capital growth investment criterion: Theory and practice*. Vol. 3. world scientific, 2011.

[NZC14] S. Nadarajah, B. Zhang, and S. Chan. “Estimation methods for expected shortfall”. In: *Quantitative Finance* 14.2 (2014), pp. 271–291.

[OR85] K. Osband and S. Reichelstein. “Information-eliciting compensation schemes”. In: *Journal of Public Economics* 27.1 (1985), pp. 107–115.

[PR24] M. Pitera and M. Rásonyi. “Utility-based acceptability indices”. In: *SIAM Journal on Financial Mathematics* 15.2 (2024), SC28–SC40.

[PS18] M. Pitera and T. Schmidt. “Unbiased estimation of risk”. In: *Journal of Banking & Finance* 91 (2018), pp. 133–145.

[PS23] M. Pitera and Ł. Stettner. “Discrete-Time Risk Sensitive Portfolio Optimization with Proportional Transaction Costs”. In: *Mathematical Finance* 33.4 (2023), pp. 1287–1313.

[PSS23] M. Pitera, T. Schmidt, and Ł. Stettner. “A novel scaling approach for unbiased adjustment of risk estimators”. In: *arXiv:2312.05655* (2023).

[Rig24] M. B. Righi. “Star-shaped acceptability indexes”. In: *Insurance: Mathematics and Economics* 117 (2024), pp. 170–181.

[Rok49] V. A. Rokhlin. “On the fundamental ideas of measure theory”. In: *Matematicheskii Sbornik* 67.1 (1949), pp. 107–150.

[Ros56] M. Rosenblatt. “Remarks on some nonparametric estimates of a density function”. In: *The Annals of Mathematical Statistics* 27.3 (1956), pp. 832–837.

[RS06] A. Ruszczyński and A. Shapiro. “Conditional risk mappings”. In: *Mathematics of Operations Research* 31.3 (2006), pp. 544–561.

[RS07] B. Roorda and J. M. Schumacher. “Time consistency conditions for acceptability measures, with an application to Tail Value at Risk”. In: *Insurance: Mathematics and Economics* 40.2 (2007), pp. 209–230.

[RS13] E. Rosazza Gianin and C. Sgarra. "Acceptability indexes via g-expectations: an application to liquidity risk". In: *Mathematics and financial economics* 7.4 (2013), pp. 457–475.

[Sha09] A. Shapiro. "On a time consistency concept in risk averse multistage stochastic programming". In: *Operations Research Letters* 37.3 (2009), pp. 143–147.

[Sha66] W. Sharpe. "Mutual Fund Performance". In: *Journal of Business* 39 (1966), pp. 119–139.

[SP94] F. Sortino and L. Price. "Performance Measurement in a Downside Risk Framework". In: *The Journal of Investing* 3.3 (1994), pp. 59–64.

[Ste99] Ł. Stettner. "Risk sensitive portfolio optimization". In: *Mathematical Methods of Operations Research* 50.3 (1999), pp. 463–474.

[Wan96] S. Wang. "Premium calculation by transforming the layer premium density". In: *ASTIN Bulletin: The Journal of the IAA* 26.1 (1996), pp. 71–92.

[Whi90] P. Whittle. *Risk-sensitive optimal control*. Wiley New York, 1990.