Entropic risk measures: coherence vs. convexity, model ambiguity, and robust large deviations

Hans Föllmer[1] Thomas Knispel[2]

Abstract

We study a coherent version of the entropic risk measure, both in the law-invariant case and in a situation of model ambiguity. In particular, we discuss its behavior under the pooling of independent risks and its connection with a classical and a robust large deviations bound.

Key words: risk measures; premium principles; model ambiguity; large deviations.

AMS Subject Classification (2010): 60F10, 62P05, 91B16, 91B30

1 Introduction

A monetary risk measure specifies the capital which should be added to a given financial position to make that position acceptable. If the monetary outcome of a financial position is described by a bounded random variable on some probability space $(\Omega, \mathcal{F}, P)$, then a monetary risk measure is given by a monotone and translation invariant functional ρ on $L^\infty(\Omega, \mathcal{F}, P)$. In the *law-invariant* case the value $\rho(X)$ only depends on the distribution of X under P. Typical examples are *Value at Risk* (VaR), *Average Value at Risk* (AVaR), also called *Conditional Value at Risk* (CVaR) or *Tail Value at Risk* (TVaR), and the *entropic risk measure* defined by

$$\begin{aligned} e_\gamma(X) :=& \tfrac{1}{\gamma} \log E_P[e^{-\gamma X}] \\ =& \sup_Q \{E_Q[-X] - \tfrac{1}{\gamma} H(Q|P)\} \end{aligned}$$

for parameters $\gamma \in [0, \infty)$, where $e_0(X) := E_P[-X]$ and $H(Q|P)$ denotes the *relative entropy* of Q with respect to P. VaR is the one which is used most widely, but it has various deficiencies; in particular it is not convex and may thus penalize a desirable diversification. AVaR is a *coherent* risk measure, i. e., *convex* and also positively homogeneous. As shown by Kusuoka [13] in the coherent and by Kunze [12] and Frittelli & Rosazza Gianin [8] in the general convex case, AVaR is a basic building block for any law-invariant convex risk measure.

The entropic risk measures e_γ are convex, and they are additive for independent positions. From an actuarial point of view, however, this property may not be desirable. Indeed, if $e_\gamma(X_1 + \ldots + X_n)$ is viewed as the total premium for a homogeneous portfolio of i. i. d. random variables $X_1, \ldots, X_n$, then the premium per contract would simply be $e_\gamma(X_1)$, no matter how large n is. Thus the pooling of independent risks

[1]Humboldt-Universität zu Berlin, Institut für Mathematik, Unter den Linden 6, 10099 Berlin, Germany, email: foellmer@math.hu-berlin.de

[2]Leibniz Universität Hannover, Institut für Mathematische Stochastik, Welfengarten 1, 30167 Hannover, Germany, email: knispel@stochastik.uni-hannover.de

does not have the effect that the premium per contract decreases to the "fair premium", i. e., to the expected loss from a single contract, as the number of contracts increases.

In this note we focus on a fourth example, namely on a *coherent* version of the entropic risk measure defined by

$$\rho_c(X) := \sup_{Q:H(Q|P)\leq c} E_Q[-X].$$

In Section 3 we clarify the connection between the coherent entropic risk measures ρ_c and the convex entropic risk measures e_γ. In Section 4 we show that the capital requirements computed in terms of ρ_c have the desired behavior under the pooling of independent risks $X_1, \ldots, X_n$. In fact it turns out that the asymptotic analysis of $\rho_c(X_1+\ldots+X_n)$ simply amounts to a reformulation, in the language of risk measures, of Cramér's classical proof of the upper bound for large deviations of the average loss.

In Section 5 we extend the discussion beyond the law-invariant case by taking model ambiguity into account. Instead of fixing a probability measure P we consider a whole class $\mathcal{P}$ of probabilistic models. We define corresponding robust versions $e_{\mathcal{P},\gamma}$ and $\rho_{\mathcal{P},c}$ of the entropic risk measures and derive some of their basic properties. In particular, we show that the pooling of risks has the desired effect if premia are computed in terms of $\rho_{\mathcal{P},c}$, and that this corresponds to a robust version of Cramér's theorem for large deviations.

2 Notation and definitions

Let $\mathfrak{X}$ be the linear space of bounded measurable functions on some measurable space $(\Omega, \mathcal{F})$. Consider a set $\mathcal{A} \subseteq \mathfrak{X}$ such that $\emptyset \neq \mathcal{A} \cap \mathbb{R} \neq \mathbb{R}$ and

$$X \in \mathcal{A}, Y \in \mathfrak{X}, Y \geq X \quad \Rightarrow \quad Y \in \mathcal{A}.$$

Then the functional $\rho : \mathfrak{X} \to \mathbb{R}$ defined by

$$\rho(X) := \inf\{m \in \mathbb{R} | X + m \in \mathcal{A}\} \tag{1}$$

is

i) *monotone*, i. e., $\rho(X) \leq \rho(Y)$ if $X \geq Y$,

and

ii) *cash-invariant*, i. e., $\rho(X + m) = \rho(X) - m$ for $X \in \mathfrak{X}$ and $m \in \mathbb{R}$.

Definition 2.1. *A functional $\rho : \mathfrak{X} \to \mathbb{R}$ with properties* i) *and* ii) *is called a* monetary risk measure.

Any monetary risk measure is of the form (1) with $\mathcal{A}_\rho := \{X \in \mathfrak{X} | \rho(X) \leq 0\}$.

If $X \in \mathfrak{X}$ is interpreted as the uncertain monetary outcome of a financial position and $\mathcal{A}$ as a class of "acceptable positions", then $\rho(X)$ can be regarded as a *capital requirement*, i. e., as the minimal capital which should be added to the position to make it acceptable.

Definition 2.2. *A monetary risk measure is called a* convex risk measure *if it is quasi-convex, i. e.,*

$$\rho(\lambda X + (1-\lambda)Y) \leq \max\{\rho(X), \rho(Y)\}$$

for $X, Y \in \mathfrak{X}$ and $\lambda \in [0,1]$. In that case $\mathcal{A}_\rho$ is convex, and this implies that ρ is a convex functional on $\mathfrak{X}$; cf. Föllmer & Schied [7], Proposition 4.6. A convex risk measure is called coherent *if it is positively homogeneous, i. e.,*

$$\rho(\lambda X) = \lambda \rho(X)$$

for $X \in \mathfrak{X}$ and $\lambda \geq 0$.

Remark 2.1. *Convex risk measures are closely related to actuarial premium principles; cf., e. g., Kaas et al. [11]. For example, it is shown in Deprez & Gerber [2], that a convex premium principle H is of the form $H(X) = \rho(-X)$ for some convex risk measure ρ if it satisfies the "no rip-off" condition $H(X) \leq \sup X$.*

Typically, a convex risk measure admits a *robust representation* of the form

$$\rho(X) = \sup_{Q \in \mathcal{M}_1} \{E_Q[-X] - \alpha(Q)\}, \tag{2}$$

where $\mathcal{M}_1$ denotes the class of all probability measures on $\mathfrak{X}$, and where the penalty function $\alpha : \mathcal{M}_1 \to (-\infty, \infty]$ is defined by

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

cf., e. g., Artzner et al. [1], Delbaen [4], Frittelli & Rosazza Gianin [8], and Föllmer & Schied [7], Chapter 4 for criteria and examples. In the coherent case we have $\alpha(Q) \in \{0, \infty\}$, and (2) reduces to

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

where $\mathcal{Q} := \{Q \in \mathcal{M}_1 | \alpha(Q) = 0\}$.

Now suppose that P is a probability measure on $(\Omega, \mathcal{F})$ and that $\rho(X) = \rho(Y)$ if $X = Y$ P-a. s.. Then ρ can be regarded as a convex risk measure on $L^\infty := L^\infty(\Omega, \mathcal{F}, P)$. In this case the representation (2) holds if ρ is *continuous from above*, i. e., $\rho(X_n) \nearrow \rho(X)$ whenever X_n decreases to X in $\mathfrak{X}$, and $\mathcal{M}_1$ can be replaced by the class $\mathcal{M}_1(P) := \{Q \in \mathcal{M}_1 | Q \ll P\}$.

A monetary risk measure ρ is called *law-invariant* if $\rho(X)$ only depends on the distribution of X under the given probability measure P. For a convex risk measure which is continuous from above, this is the case if and only if the penalty $\alpha(Q)$ of $Q \in \mathcal{M}_1(P)$ only depends on the law of $\frac{dQ}{dP}$ under P; cf., e. g., [7], Theorem 4.54.

A large class of examples arises if acceptability is defined in terms of expected utility, i. e., if

$$\mathcal{A} = \{X \in L^\infty | E_P[u(X)] \geq u(0)\}$$

for some concave increasing function u. In this case the resulting risk measure is convex and law-invariant, and its penalty function can be computed in terms of the conjugate function of u; cf. [7], Theorem 4.106.

Let us now take an exponential utility of the form $u(x) = 1 - e^{-\gamma x}$ for some $\gamma > 0$. In that case the corresponding risk measure is given by

$$e_\gamma(X) = \tfrac{1}{\gamma} \log E_P[e^{-\gamma X}], \quad X \in L^\infty, \tag{3}$$

and its robust representation takes the form

$$e_\gamma(X) = \sup_{Q \in \mathcal{M}_1} \{E_Q[-X] - \tfrac{1}{\gamma} H(Q|P)\},$$

where

$$H(Q|P) = \begin{cases} E_Q[\log \frac{dQ}{dP}] & \text{if } Q \ll P \\ +\infty & \text{otherwise} \end{cases}$$

denotes the *relative entropy* of Q with respect to P.

Definition 2.3. *The convex risk measure e_γ defined by (3) is called the* (convex) entropic risk measure *with parameter γ.*

It is easy to see that $e_\gamma(X)$ is increasing in γ and strictly increasing as soon as X is not constant P-a. s.. Moreover,

$$\lim_{\gamma \downarrow 0} e_\gamma(X) = E_P[-X] \quad \text{and} \quad \lim_{\gamma \uparrow \infty} e_\gamma(X) = \operatorname{ess\,sup}(-X); \tag{4}$$

cf., e. g., [11], Theorem 1.3.2.

As noted already by de Finetti [3], the entropic risk measures can be characterized as the only monetary risk measures ρ which are, up to a change of sign, also a *certainty equivalent*, i. e.,

$$u(-\rho(X)) = E_P[u(X)]$$

for some strictly increasing concave utility function u. In this case the utility function is exponential, and $\rho = e_\gamma$ for some $\gamma \in [0, \infty)$.

The actuarial premium principle $H(X) = e_\gamma(-X)$ corresponding to the entropic risk measure is usually called the *exponential principle*; cf. Deprez & Gerber [2] and Gerber [9].

3 Coherent entropic risk measures

In this section we focus on the following *coherent* version of an entropic risk measure.

Definition 3.1. *For each $c > 0$, the risk measure ρ_c defined by*

$$\rho_c(X) := \sup_{Q \in \mathcal{M}_1 : H(Q|P) \le c} E_Q[-X], \quad X \in L^\infty, \tag{5}$$

will be called the coherent entropic risk measure *at level c.*

Clearly, ρ_c is a coherent risk measure. It is also law-invariant; this follows from Theorem 4.54 in Föllmer & Schied [7], and also from the representation (8) in Proposition 3.1 below; cf. Corollary 3.1.

For $X \in L^\infty$ we denote by

$$\mathcal{Q}_{P,X} = \{Q_\gamma | \gamma \in \mathbb{R}\} \tag{6}$$

the exponential family induced by P and $-X$, i. e.,

$$\frac{dQ_\gamma}{dP} = e^{-\gamma X} E_P[e^{-\gamma X}]^{-1}.$$

If $p(X) := P[X = \operatorname{ess\,inf} X] > 0$, then we include as limiting case the measure $Q_\infty := \lim_{\gamma \uparrow \infty} Q_\gamma = P[\cdot | X = \operatorname{ess\,inf} X]$.

The following proposition shows that the supremum in (5) is attained by some probability measure in the exponential family $\mathcal{Q}_{P,X}$.

Proposition 3.1. *For $c \in (0, -\log p(X))$ we have*

$$\rho_c(X) = \max_{Q \in \mathcal{M}_1 : H(Q|P) \le c} E_Q[-X] = E_{Q_{\gamma_c}}[-X], \tag{7}$$

where $Q_{\gamma_c} \in \mathcal{Q}_{P,X}$ and $\gamma_c > 0$ is such that $H(Q_{\gamma_c}|P) = c$, and

$$\rho_c(X) = \min_{\gamma > 0} \{ \tfrac{c}{\gamma} + e_\gamma(X) \} = \tfrac{c}{\gamma_c} + e_{\gamma_c}(X). \tag{8}$$

If $p(X) > 0$ and $c \ge -\log p(X)$, then

$$\rho_c(X) = E_{Q_\infty}[-X] = \operatorname{ess\,sup}(-X). \tag{9}$$

Proof. We exclude the trivial case, where X is P-a. s. constant.

1) Assume that $0 < c < -\log p(X)$ and take Q such that $H(Q|P) \le c$. For any $\gamma > 0$ and for $Q_\gamma \in \mathcal{Q}_{P,X}$,

$$H(Q|P) = H(Q|Q_\gamma) + \gamma E_Q[-X] - \log E_P[e^{-\gamma X}] \tag{10}$$

with $H(Q|Q_\gamma) \ge 0$ and $H(Q|Q_\gamma) = 0$ iff $Q = Q_\gamma$. Thus

$$E_Q[-X] \le \tfrac{c}{\gamma} + e_\gamma(X), \tag{11}$$

and

$$\begin{aligned} \rho_c(X) = & \sup_{Q \in \mathcal{M}_1 : H(Q|P) \le c} E_Q[-X] \\ & \le \inf_{\gamma > 0} \{ \tfrac{c}{\gamma} + e_\gamma(X) \}. \end{aligned}$$

Both the supremum and the infimum are attained, and they coincide. Indeed, we can choose $\gamma_c > 0$ such that $H(Q_{\gamma_c}|P) = c$, and then we get equality in (11) for $Q = Q_\gamma$ and $\gamma = \gamma_c$. Such a $\gamma_c > 0$ exists and is unique. Indeed,

$$H(Q_\gamma|P) = \gamma E_{Q_\gamma}[-X] - \log E_p[e^{-\gamma X}]$$

is continuous and strictly increasing in γ, and $\lim_{\gamma \uparrow \infty} H(Q_\gamma|P) = -\log p(X)$ since $\lim_{\gamma \uparrow \infty} Q_\gamma[A_a] = 1$ for $A_a := \{-X > a\}$ and any $a < \operatorname{ess\,sup}(-X)$.

2) If $p(X) > 0$, then Q_∞ satisfies $H(Q_\infty|P) = -\log p(X)$ and $E_{Q_\infty}[-X] = \operatorname{ess\,sup}(-X)$. For $c \ge -\log p(X)$ we thus obtain $\rho_c(X) \ge \operatorname{ess\,sup}(-X)$, hence (9), since the converse inequality in (9) is clear. $\square$

Remark 3.1. *Conversely, the convex entropic risk measure e_γ can be expressed in terms of the coherent entropic risk measures ρ_c as follows:*

$$e_\gamma(X) = \min_{c > 0} \{ \rho_c(X) - \tfrac{c}{\gamma} \} = \rho_{c_\gamma}(X) - \tfrac{c_\gamma}{\gamma},$$

where $c_\gamma := H(Q_\gamma|P)$; this follows immediately from (8).

Remark 3.2. *Note that the parameter γ_c in (7) depends both on c and on X. For a fixed value $\gamma > 0$, the resulting functional*

$$\rho(X) := E_P[(-X)e^{-\gamma X}](E_P[e^{-\gamma X}])^{-1}$$

is neither coherent nor convex, but the corresponding actuarial premium principle $H(X) = \rho(-X)$ is well-known as the Esscher principle; cf., e. g., Deprez & Gerber [2].

Corollary 3.1. *The coherent risk measure ρ_c is law-invariant, continuous from above, and even continuous from below. Moreover, ρ_c is increasing in c and satisfies*

$$\lim_{c\downarrow 0} \rho_c(X) = E_P[-X] \quad \textit{and} \quad \lim_{c\uparrow\infty} \rho_c(X) = \operatorname{ess\,sup}(-X). \tag{12}$$

Proof. Law-invariance follows from (8) since each e_γ is law-invariant. Continuity from below follows from (7), i. e., from the representation

$$\rho_c(X) = \max_{Q\in\mathcal{Q}} E_Q[-X]$$

with $\mathcal{Q} = \{Q \in \mathcal{M}_1 | H(Q|P) \leq c\}$; cf. [7], Corollary 4.35. In particular, ρ_c is continuous from above; cf. [7], Corollary 4.35 together with Theorem 4.31. The convergence in (12) follows easily from Proposition 3.1. Indeed, (8) implies $\lim_{c\downarrow 0} \rho_c(X) \leq e_\gamma(X)$ for each $\gamma > 0$, hence the first equality in (12), due to (4). As to the second equality, it is enough to consider the measures $Q = P[\cdot|A_a]$ for the sets $A_a := \{-X > a\}$ with $a < \operatorname{ess\,sup}(-X)$. □

Let us now compare the coherent entropic risk measures ρ_c to the familiar risk measures "Value at Risk" and "Average Value at Risk" defined by

$$\mathrm{VaR}_\alpha(X) := \inf\{m \in \mathbb{R} | P[X + m < 0] \leq \alpha\}$$

and

$$\mathrm{AVaR}_\alpha(X) := \tfrac{1}{\alpha} \int_0^\alpha \mathrm{VaR}_\lambda(X)\, d\lambda \geq \mathrm{VaR}_\alpha(X)$$

for any $\alpha \in (0, 1)$. Recall that VaR_α is a monetary risk measure which is positively homogeneous but not convex, while AVaR_α is a coherent risk measure which can also be written as

$$\mathrm{AVaR}_\alpha(X) = \tfrac{1}{\alpha} E_P[(q_\alpha - X)^+] - q_\alpha$$

for any α-quantile q_α of X; cf., e. g., [7], Section 4.4. Note also that $\mathrm{VaR}_\alpha(X)$ is decreasing and right-continuous in α with left limits

$$\mathrm{VaR}_{\alpha-}(X) := \inf\{m \in \mathbb{R} | P[X + m \leq 0] \leq \alpha\}.$$

Proposition 3.2. *For any $\alpha \in (0, 1)$ and any $X \in L^\infty$,*

$$\mathrm{VaR}_\alpha(X) \leq \mathrm{VaR}_{\alpha-}(X) \leq \mathrm{AVaR}_\alpha(X) \leq \rho_{c(\alpha)}(X), \tag{13}$$

where $c(\alpha) := -\log \alpha > 0$.

Proof. Clearly, we have $\mathrm{VaR}_\alpha(X) \leq \mathrm{VaR}_{\alpha-}(X) \leq \mathrm{AVaR}_\alpha(X)$. In view of Corollary 3.1 it is enough to verify the inequality $\mathrm{VaR}_\alpha(X) \leq \rho_{c(\alpha)}(X)$, since AVaR_α is the smallest law-invariant coherent risk measure which is continuous from above and dominates VaR_α; cf. [7], Theorem 4.61. For any $\gamma > 0$,

$$P[X + m \leq 0] \leq e^{-\gamma m} E_P[e^{-\gamma X}],$$

and the right-hand side is $\leq \alpha$ if $-\gamma m + \log E_P[e^{-\gamma X}] \leq \log \alpha$, i. e., if

$$m \geq \tfrac{c(\alpha)}{\gamma} + e_\gamma(X).$$

Thus, by Proposition 3.1,

$$\mathrm{VaR}_{\alpha-}(X) \leq \inf_{\gamma>0} \{\tfrac{c(\alpha)}{\gamma} + e_\gamma(X)\} = \rho_{c(\alpha)}(X).$$

Alternatively, we can check directly the last inequality in (13), using the robust representation

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

with $\mathcal{Q}_\alpha := \{Q \in \mathcal{M}_1 | Q \ll P, \frac{dQ}{dP} \leq \frac{1}{\alpha}\}$; cf. [7], Lemma 4.46 and Theorem 4.47. Indeed, any $Q \in \mathcal{Q}_\alpha$ satisfies $\log \frac{dQ}{dP} \leq c(\alpha)$, hence $H(Q|P) \leq c(\alpha)$. □

4 Capital requirements for i. i. d. portfolios

Consider a homogeneous portfolio of n insurance contracts whose uncertain outcomes are described as i. i. d. random variables $X_1, \ldots, X_n$ on some probability space $(\Omega, \mathcal{F}, P)$. Let μ denote the distribution of X_1 under P. To keep this exposition simple, we assume that μ is non-degenerate and has bounded support; in fact it would be enough to require finite exponential moments $\int e^{-\gamma x} \mu(dx)$ for any $\gamma \in \mathbb{R}$.

If ρ is a monetary risk measure, then $\rho(X_1 + \ldots + X_n)$ can be viewed as the smallest monetary amount which should be added to make the portfolio acceptable. This suggests to equate $\rho(X_1 + \ldots + X_n)$ with the portfolio's total premium, and to use the fraction

$$\pi_n = \tfrac{1}{n}\rho(X_1 + \ldots + X_n)$$

as a premium for each individual contract X_i, $i = 1, \ldots, n$.

For any $\gamma > 0$, the entropic risk measure e_γ satisfies

$$e_\gamma(X_1 + \ldots + X_n) = n e_\gamma(X_1).$$

If e_γ is used to calculate the premium π_n, it yields

$$\pi_n = \tfrac{1}{n} e_\gamma(X_1 + \ldots + X_n) = e_\gamma(X_1) > E_P[-X_1].$$

Thus the exponential premium principle based on the convex entropic risk measure does not have the desirable property that the "risk premium" $\pi_n - E_P[-X_1]$ decreases to 0 as n tends to ∞; cf., e. g., [16], Example 12.5.1 and Remark 12.5.2.

For the *coherent* risk measure ρ_c, however, the pooling of risks does have the desired effect.

Corollary 4.1. *The premium*

$$\pi_{c,n} := \tfrac{1}{n}\rho_c(X_1 + \ldots + X_n) \tag{14}$$

computed in terms of the coherent entropic risk measure ρ_c satisfies $\pi_{c,n} > E_P[-X_1]$ and

$$\lim_{n \uparrow \infty} \pi_{c,n} = E_P[-X_1].$$

Proof. In view of (4) we can choose for any $\epsilon > 0$ some $\delta > 0$ such that $e_\delta(X_1) \leq E_P[-X_1] + \epsilon$. Thus, by (8),

$$\begin{aligned} \pi_{c,n} &= \tfrac{1}{n} \inf_{\gamma > 0} \{ \tfrac{c}{\gamma} + e_\gamma(X_1 + \ldots + X_n) \} \\ &= \inf_{\gamma > 0} \{ \tfrac{c}{\gamma n} + e_\gamma(X_1) \} \\ &\leq \tfrac{c}{\delta n} + E_P[-X_1] + \epsilon, \end{aligned}$$

hence $\overline{\lim}_{n\uparrow\infty} \pi_{c,n} \leq E_P[-X_1]$. Since $\pi_{c,n} \geq E_P[-X_1]$, the conclusion follows. In fact we have $\pi_{c,n} > E_P[-X_1]$ since the distribution μ of X_1 is non-degenerate. □

Let us now describe the decay of the risk premium $\pi_{c,n} - E_P[-X_1]$ more precisely. Note that for any $Q_\gamma \in \mathcal{Q}_{P,X_1+\ldots+X_n}$ we have

$$\tfrac{dQ_\gamma}{dP} = e^{-\gamma\sum_{i=1}^n X_i}(E_P[e^{-\gamma X_1}])^{-n}$$

and

$$H(Q_\gamma|P) = nH(\mu_\gamma|\mu),$$

where μ_γ is the distribution on $\mathbb{R}$ with density

$$\tfrac{d\mu_\gamma}{d\mu}(x) := e^{-\gamma x}(\int e^{-\gamma x}\,\mu(dx))^{-1},$$

mean $m(\gamma) := \int(-x)\mu_\gamma(dx)$, and variance $\sigma^2(\gamma) := \int(x+m(\gamma))^2\,\mu_\gamma(dx)$. We denote by $\sigma_P^2(X_1) := \sigma^2(0)$ the variance of X_1 under P.

Proposition 4.1. *For a given level $c > 0$, the premium $\pi_{c,n}$ defined by (14) is given by $\pi_{c,n} = m(\gamma_{c,n})$, where $\gamma_{c,n}$ is such that $H(\mu_{\gamma_{c,n}}|\mu) = \frac{c}{n}$, and we have*

$$\lim_{n\uparrow\infty} \sqrt{n}(\pi_{c,n} - E_P[-X_1]) = \sqrt{2c}\,\sigma_P(X_1).$$

Proof. Recall from Proposition 3.1 that

$$\rho_c(X_1+\ldots+X_n) = E_{Q_{\gamma_{c,n}}}[-(X_1+\ldots+X_n)],$$

where $Q_{\gamma_{c,n}} \in \mathcal{Q}_{P,X_1+\ldots+X_n}$, and where the parameter $\gamma_{c,n} > 0$ is taken such that $H(Q_{\gamma_{c,n}}|P) = c$. Thus,

$$c = H(Q_{\gamma_{c,n}}|P) = nH(\mu_{\gamma_{c,n}}|\mu), \tag{15}$$

and the individual premium $\pi_{c,n}$ can be rewritten as

$$\pi_{c,n} = \tfrac{1}{n}\rho_c(X_1+\ldots+X_n) = m(\gamma_{c,n}). \tag{16}$$

The smooth function f defined by $f(\gamma) := \log E_P[e^{-\gamma X_1}] = \log\int e^{-\gamma x}\,\mu(dx)$ satisfies $f'(\gamma) = m(\gamma)$ and $f''(\gamma) = \sigma^2(\gamma)$, and so we have

$$H(\mu_\gamma|\mu) = \gamma m(\gamma) - f(\gamma) = \tfrac{1}{2}f''(\widetilde{\gamma})\gamma^2$$

for some $\widetilde{\gamma} \in [0,\gamma]$. The condition

$$\tfrac{c}{n} = H(\mu_{\gamma_{c,n}}|\mu) = \tfrac{1}{2}f''(\widetilde{\gamma}_{c,n})\gamma_{c,n}^2$$

clearly implies $\lim_{n\uparrow\infty}\gamma_{c,n} = 0$, hence

$$\lim_{n\uparrow\infty} n\gamma_{c,n}^2 = 2c(f''(0))^{-1}. \tag{17}$$

Since

$$\begin{aligned} f(\gamma_{c,n}) &= f(0) + f'(0)\gamma_{c,n} + \tfrac{1}{2}f''(\widehat{\gamma}_{c,n})\gamma_{c,n}^2 \\ &= m(0)\gamma_{c,n} + \tfrac{1}{2}f''(\widehat{\gamma}_{c,n})\gamma_{c,n}^2 \end{aligned}$$

for some $\widehat{\gamma}_{c,n} \in [0,\gamma_{c,n}]$, we have

$$\begin{aligned} \pi_{c,n} = m(\gamma_{c,n}) &= \tfrac{1}{\gamma_{c,n}}(H(\mu_{\gamma_{c,n}}|\mu) + f(\gamma_{c,n})) \\ &= \tfrac{1}{\gamma_{c,n}}(\tfrac{c}{n} + m(0)\gamma_{c,n} + \tfrac{1}{2}f''(\widehat{\gamma}_{c,n})\gamma_{c,n}^2). \end{aligned}$$

Due to (17), we finally obtain

$$\lim_{n\uparrow\infty} \sqrt{n}(\pi_{c,n} - E_P[-X_1]) = \lim_{n\uparrow\infty} \frac{c}{\sqrt{n}\gamma_{c,n}} \left(1 + \frac{1}{2} f''(\widehat{\gamma}_{c,n}) \frac{\gamma_{c,n}^2 n}{c}\right)$$
$$= \sqrt{2c}\,\sigma_P(X_1).$$

□

Let us now fix a premium π such that $E_P[-X_1] < \pi < \operatorname{ess\,sup}(-X_1)$, and let us determine the maximal tolerance level

$$c_{\pi,n} := \max\{c > 0 | \tfrac{1}{n}\rho_c(X_1 + \ldots + X_n) \leq \pi\}$$

at which the portfolio $X_1, \ldots, X_n$ is made acceptable by the total premium $n\pi$.

Corollary 4.2. *Take $\gamma(\pi) > 0$ such that $m(\gamma(\pi)) = \pi$. Then*

$$c_{\pi,n} = nH(\mu_{\gamma(\pi)}|\mu).$$

Proof. At level $c_{\pi,n}$ we have

$$\tfrac{1}{n}\rho_{c_{\pi,n}}(X_1 + \ldots + X_n) = m(\gamma(\pi)).$$

In view of (15) and (16) this is the case iff

$$c_{\pi,n} = nH(\mu_{\gamma(\pi)}|\mu).$$

□

Remark 4.1. *Due to (13), Corollary 4.2 implies*

$$\mathrm{VaR}_{\alpha_{\pi,n}-}(X_1 + \ldots + X_n) \leq n\pi$$

for $\alpha_{\pi,n} := \exp(-c_{\pi,n})$. But this translates into

$$P[-\tfrac{1}{n}(X_1 + \ldots + X_n) \geq \pi] \leq \alpha_{\pi,n},$$

and so we obtain

$$\tfrac{1}{n}\log P[-\tfrac{1}{n}(X_1 + \ldots + X_n) \geq \pi] \leq -\tfrac{c_{\pi,n}}{n} = -H(\mu_{\gamma(\pi)}|\mu).$$

In other words, the combination of Corollary 4.2 with the estimate (13) simply amounts to a reformulation, in the language of risk measures, of the classical proof of Cramér's upper bound for the large deviations of the averages $-\frac{1}{n}(X_1 + \ldots + X_n)$; see, e. g., Dembo & Zeitouni [5].

5 Model ambiguity and robust large deviations

So far we have fixed a probability measure P which is assumed to be known. Let us now consider a situation of model ambiguity where P is replaced by a whole class $\mathcal{P}$ of probability measures on $(\Omega, \mathcal{F})$.

Assumption 5.1. *We assume that all measures $P \in \mathcal{P}$ are equivalent to some reference measure R on $(\Omega, \mathcal{F})$, and that the family of densities*

$$\Phi_{\mathcal{P}} := \{\tfrac{dP}{dR}|P \in \mathcal{P}\}$$

is convex and weakly compact in $L^1(R)$.

For a probability measure Q on $(\Omega, \mathcal{F})$, the extent to which it differs from the measures in the class $\mathcal{P}$ will be measured by the *relative entropy of Q with respect to the class $\mathcal{P}$*, defined as

$$H(Q|\mathcal{P}) := \inf_{P \in \mathcal{P}} H(Q|P).$$

Our assumption implies that for each Q such that $H(Q|\mathcal{P}) < \infty$ there is a unique measure $P_Q \in \mathcal{P}$, called the *reverse entropic projection* of Q on $\mathcal{P}$, such that $H(Q|P_Q) = H(Q|\mathcal{P})$; cf. [6], Remark 2.10 and Proposition 2.14.

Let us denote by $\mathcal{M}_1(R)$ the class of all probability measures on $(\Omega, \mathcal{F})$ which are absolutely continuous with respect to R. From now on we write $L^\infty = L^\infty(\Omega, \mathcal{F}, R)$; note that $L^\infty \subseteq L^\infty(\Omega, \mathcal{F}, P)$ for any $P \in \mathcal{M}_1(R)$. We also use the notation $e_{P,\gamma}$ and $\rho_{P,c}$ for the convex and the coherent entropic risk measures defined in terms of the specific measure P.

In this context of model ambiguity, we define the robust version $e_{\mathcal{P},\gamma}$ of the (convex) entropic risk measure by

$$e_{\mathcal{P},\gamma}(X) := \sup_{P \in \mathcal{P}} e_{P,\gamma}(X) = \tfrac{1}{\gamma} \sup_{P \in \mathcal{P}} \log E_P[e^{-\gamma X}], \quad X \in L^\infty.$$

Assumption 5.1 implies that the supremum is actually attained. Clearly, $e_{\mathcal{P},\gamma}$ is again a convex risk measure, and its robust representation takes the form

$$e_{\mathcal{P},\gamma}(X) = \sup_{Q \in \mathcal{M}_1} \{E_Q[-X] - \tfrac{1}{\gamma} H(Q|\mathcal{P})\}, \quad X \in L^\infty.$$

Lemma 5.1. *We have*

$$e_{\mathcal{P},\gamma}(X) \geq \max_{P \in \mathcal{P}} E_P[-X],$$

and $e_{\mathcal{P},\gamma}(X)$ is increasing in γ with

$$\lim_{\gamma \downarrow 0} e_{\mathcal{P},\gamma}(X) = \max_{P \in \mathcal{P}} E_P[-X]. \tag{18}$$

Proof. The functions

$$f_\gamma(\varphi_P) := e_{P,\gamma}(X) - E_P[-X]$$

with $\varphi_P := \frac{dP}{dR}$ are weakly continuous on $\Phi_{\mathcal{P}}$ and they decrease pointwise to 0, due to (4). Since $\Phi_{\mathcal{P}}$ is weakly compact, the convergence is uniform by Dini's lemma, and this implies (18). Note that the maximum in (18) is actually attained since $\varphi_P \mapsto E_R[(-X)\varphi_P]$ is continuous on the weakly compact set $\Phi_{\mathcal{P}}$. □

From now on we focus on the robust extension $\rho_{\mathcal{P},c}$ of the *coherent* entropic risk measure defined by

$$\rho_{\mathcal{P},c}(X) := \sup_{Q \in \mathcal{M}_1 : H(Q|\mathcal{P}) \leq c} E_Q[-X] \tag{19}$$

for any $X \in L^\infty$.

Lemma 5.2. *The supremum in (19) is attained, i. e., for any $X \in L^\infty$ there is a pair $(Q_c, P_c) \in \mathcal{M}_1(R) \times \mathcal{P}$ such that $H(Q_c|P_c) \leq c$ and*

$$\rho_{\mathcal{P},c}(X) = E_{Q_c}[-X].$$

In particular,

$$\rho_{\mathcal{P},c}(X) = \max_{P \in \mathcal{P}} \rho_{P,c}(X) \geq \max_{P \in \mathcal{P}} E_P[-X]. \tag{20}$$

Proof. Since any Q such that $H(Q|\mathcal{P}) < \infty$ admits a reverse entropic projection $P_Q \in \mathcal{P}$, we can write

$$\begin{aligned}\rho_{\mathcal{P},c}(X) &= \sup_{Q\in\mathcal{M}_1(R):H(Q|\mathcal{P})\le c} E_Q[-X] \\ &= \sup_{Q\in\mathcal{M}_1(R),P\in\mathcal{P}:H(Q|P)\le c} E_Q[-X] \\ &= \sup_{(\varphi,\psi)\in\mathcal{C}_c} E_R[(-X)\varphi],\end{aligned}$$

where we define

$$\mathcal{C}_c := \{(\varphi,\psi)\in\Phi\times\Phi_{\mathcal{P}}|E_R[h(\varphi,\psi)]\le c\}$$

with $\Phi := \{\frac{dQ}{dP}|Q\in\mathcal{M}_1(R)\}$ and $h(x,y) := x\log\frac{x}{y}$ for $y>0$ and $x\ge 0$, $h(0,0) := 0$, and $h(x,0) = \infty$ for $x>0$. The functional $F(\varphi,\psi) := E_R[(-X)\varphi]$ is weakly continuous on $\Phi\times\Phi_{\mathcal{P}}$, and the set $\mathcal{C}_c$ is weakly compact in $L^1(R)\times L^1(R)$; cf. the proof of Lemma 2.9 in Föllmer & Gundel [6]. This shows that the supremum in (19) is actually attained and that (20) holds. □

Recall that for $X\in L^\infty$ and $P\in\mathcal{P}$ we denote by $\mathcal{Q}_{P,X}$ the exponential family introduced in (6).

Proposition 5.1. *For*

$$c < -\log\max_{P\in\mathcal{P}} P[X = \operatorname{ess\,inf} X]$$

we have

$$\begin{aligned}\rho_{\mathcal{P},c}(X) &= \max_{P\in\mathcal{P}}\min_{\gamma>0}\{\tfrac{c}{\gamma} + e_{P,\gamma}(X)\} \\ &= \tfrac{c}{\gamma_c} + e_{P_c,\gamma_c}(X) \\ &= E_{Q_c}[-X],\end{aligned}$$

where Q_c denotes the measure in the exponential family $\mathcal{Q}_{P_c,X}$ with parameter γ_c, and $\gamma_c > 0$ is such that

$$H(Q_c|\mathcal{P}) = H(Q_c|P_c) = c. \tag{21}$$

If $c \ge -\log\max_{P\in\mathcal{P}} P[X = \operatorname{ess\,inf} X]$, then

$$\rho_{\mathcal{P},c}(X) = \operatorname{ess\,sup}(-X). \tag{22}$$

Proof. 1) If $c \ge -\log P[X = \operatorname{ess\,inf} X]$ for some $P\in\mathcal{P}$, then $\rho_{P,c}(X) = \operatorname{ess\,sup}(-X)$ due to Proposition 3.1, and this implies (22).

2) The proof of Lemma 5.2 shows that for any $X\in L^\infty$ there exists a pair $(Q_c,P_c)\in\mathcal{M}_1(R)\times\mathcal{P}$ such that $H(Q_c|P_c)\le c$ and

$$\rho_{\mathcal{P},c}(X) = E_{Q_c}[-X].$$

Let us first show that Q_c belongs to the exponential family $\mathcal{Q}_{P_c,X}$, and that $H(Q_c|P_c) = c$. To this end, we take $\gamma_c > 0$ such that $H(Q_{P_c,\gamma_c}|P_c) = c$, and we show that $Q_c = Q_{P_c,\gamma_c}$. Indeed, Q_{P_c,γ_c} satisfies the constraint

$$H(Q_{P_c,\gamma_c}|\mathcal{P}) \le H(Q_{P_c,\gamma_c}|P_c) = c,$$

and as in the proof of Proposition 3.1 we see that

$$\begin{aligned}E_{Q_c}[-X] &= \frac{H(Q_c|P_c) - H(Q_c|Q_{P_c,\gamma_c})}{\gamma_c} + e_{P_c,\gamma_c}(X) \\ &\leq \tfrac{c}{\gamma_c} + e_{P_c,\gamma_c}(X) \\ &= E_{Q_{P_c,\gamma_c}}[-X].\end{aligned}$$

But $E_{Q_c}[-X]$ is maximal under the constraint $H(Q|\mathcal{P}) \leq c$, and so we must have equality. This implies $H(Q_c|Q_{P_c,\gamma_c}) = 0$, hence $Q_c = Q_{P_c,\gamma_c}$ and $H(Q_c|P_c) = H(Q_{P_c,\gamma_c}|P_c) = c$.

3) We have

$$\rho_{\mathcal{P},c}(X) = \rho_{P_c,c}(X) = \sup_{Q:H(Q|P_c)\leq c} E_Q[-X].$$

Indeed, "$\leq$" is clear since $P_c \in \mathcal{P}$. Conversely, if $H(Q|\mathcal{P}) \leq c$ then, since $Q_c \in \mathcal{Q}_{P_c,X}$ and $H(Q_c|P_c) = c$, Proposition 3.1 implies

$$E_Q[-X] \leq E_{Q_c}[-X] = \rho_{P_c,c}(X),$$

and this yields "$\leq$".

4) In view of 2) and Proposition 3.1, we have

$$\begin{aligned}\sup_{Q:H(Q|\mathcal{P})\leq c} E_Q[-X] = \rho_{\mathcal{P},c}(X) &= \sup_{P\in\mathcal{P}} \rho_{P,c}(X) \\ &= \sup_{P\in\mathcal{P}} \inf_{\gamma>0} \{ \tfrac{c}{\gamma} + e_{P,\gamma}(X)\}.\end{aligned}$$

The argument in part 1) shows that the second and the third supremum are attained by $P = P_c$, the first by $Q = Q_c$, and the infimum by $\gamma = \gamma_c$.

5) If (21) does not hold, then we have $H(Q_c|P) < c$ for some $P \in \mathcal{P}$. But then the proof of Proposition 3.1 shows that there exists some Q such that $H(Q|\mathcal{P}) \leq H(Q|P) = c$ and $E_Q[-X] > E_{Q_c}[-X]$, contradicting the definition of Q_c. □

Corollary 5.1. *The robust versions*

$$\mathrm{VaR}_{\mathcal{P},\alpha}(X) := \sup_{P\in\mathcal{P}} \mathrm{VaR}_{P,\alpha}(X) = \inf\{m \in \mathbb{R} | \sup_{P\in\mathcal{P}} P[X + m < 0] \leq \alpha\},$$

$$\mathrm{VaR}_{\mathcal{P},\alpha-}(X) := \sup_{P\in\mathcal{P}} \mathrm{VaR}_{P,\alpha-}(X) = \inf\{m \in \mathbb{R} | \sup_{P\in\mathcal{P}} P[X + m \leq 0] \leq \alpha\},$$

and

$$\mathrm{AVaR}_{\mathcal{P},\alpha}(X) := \sup_{P\in\mathcal{P}} \mathrm{AVaR}_{P,\alpha}(X)$$

of Value at Risk and Average Value at Risk with respect to the class of prior models $\mathcal{P}$ satisfy

$$\mathrm{VaR}_{\mathcal{P},\alpha}(X) \leq \mathrm{VaR}_{\mathcal{P},\alpha-}(X) \leq \mathrm{AVaR}_{\mathcal{P},\alpha}(X) \leq \rho_{\mathcal{P},c(\alpha)}(X) \tag{23}$$

with $c(\alpha) := -\log\alpha > 0$.

Proof. This follows from Proposition 5.1 and Proposition 3.2. □

Let us now look at the asymptotic behavior of the *robust premium*

$$\pi_{c,n} := \tfrac{1}{n}\rho_{\mathcal{P},c}(X_1 + \ldots + X_n) \tag{24}$$

for a portfolio which satisfies the following

Assumption 5.2. *For any $P \in \mathcal{P}$, the random variables $X_1, \ldots, X_n$ are i. i. d. and non-degenerate under P.*

Thus, model ambiguity only appears in the multiplicity of the distributions μ_P of X_1 under the various measures $P \in \mathcal{P}$. As in Section 4 we assume for simplicity that X_1 belongs to L^∞.

Corollary 5.2. *The robust premium $\pi_{c,n}$ defined by (24) satisfies*

$$\pi_{c,n} \geq \max_{P \in \mathcal{P}} E_P[-X_1]$$

and

$$\lim_{n \uparrow \infty} \pi_{c,n} = \max_{P \in \mathcal{P}} E_P[-X_1]. \tag{25}$$

Proof. Due to (18) we can choose $\delta > 0$ such that

$$e_{\mathcal{P},\delta}(X_1) \leq \max_{P \in \mathcal{P}} E_P[-X_1] + \epsilon$$

for a given $\epsilon > 0$. In analogy to the proof of Corollary 4.1, Proposition 5.1 yields the estimate

$$\pi_{c,n} \leq \tfrac{c}{\delta n} + \max_{P \in \mathcal{P}} E_P[-X_1] + \epsilon,$$

and this implies (25). □

Remark 5.1. *While the pooling of risks has the desired effect if premiums are computed in terms of $\rho_{\mathcal{P},c}$, this is not the case if we use the robust version $e_{\mathcal{P},\gamma}$ of the convex entropic risk version. Indeed, it is easy to check that the above homogeneity Assumption 5.2 implies*

$$e_{\mathcal{P},\gamma}(X_1 + \ldots + X_n) = n e_{\mathcal{P},\gamma}(X_1).$$

We conclude with the robust extension of Corollary 4.2 and Remark 4.1.

Proposition 5.2. *For a fixed premium π such that*

$$\max_{P \in \mathcal{P}} E_P[-X_1] < \pi < \operatorname{ess\,sup}(-X_1),$$

the corresponding tolerance level

$$c_{\pi,n} := \max\{c > 0 | \tfrac{1}{n} \rho_{\mathcal{P},c}(X_1 + \ldots + X_n) \leq \pi\} \tag{26}$$

is given by

$$c_{\pi,n} = n I_{\mathcal{P}}(\pi) = n \Lambda^*_{\mathcal{P}}(\pi), \tag{27}$$

where

$$I_{\mathcal{P}}(\pi) := \min_{Q : E_Q[-X_1] = \pi} H(Q|\mathcal{P}) \tag{28}$$

and

$$\Lambda^*_{\mathcal{P}}(\pi) = \sup_{\gamma > 0} \{\gamma \pi - \sup_{P \in \mathcal{P}} \log E_P[e^{-\gamma X_1}]\}. \tag{29}$$

In particular, $I_{\mathcal{P}}$ coincides with the convex conjugate $\Lambda^*_{\mathcal{P}}$ of the convex function $\Lambda_{\mathcal{P}}$ defined by

$$\Lambda_{\mathcal{P}}(\gamma) := \sup_{P \in \mathcal{P}} \log E_P[e^{-\gamma X_1}], \quad \gamma > 0.$$

Proof. 1) Let us first show the identity $I_{\mathcal{P}} = \Lambda^*_{\mathcal{P}}$. Indeed, for any $Q \in \mathcal{M}_1$ such that $E_Q[-X_1] \geq \pi$, for any $P \in \mathcal{P}$, and for any $\gamma > 0$, (10) implies

$$H(Q|P) \geq \gamma\pi - \log E_P[e^{-\gamma X_1}]$$

with equality iff $Q = Q_\gamma \in \mathcal{Q}_{P,X}$ and $\gamma > 0$ is such that $E_{Q_\gamma}[-X] = \pi$. This yields the classical identity

$$I_P(\pi) := \min_{Q: E_Q[-X_1]=\pi} H(Q|P) = \Lambda^*_P(\pi),$$

where Λ^*_P denotes the convex conjugate of the convex function Λ_P defined by $\Lambda_P(\gamma) := \log E_P[e^{-\gamma X_1}]$; cf., e. g., [7], Theorem 3.28. Thus

$$\begin{aligned} I_{\mathcal{P}}(\pi) = \min_{Q: E_Q[-X_1]=\pi} H(Q|\mathcal{P}) &= \inf_{P \in \mathcal{P}} \min_{Q: E_Q[-X_1]=\pi} H(Q|P) \\ &= \inf_{P \in \mathcal{P}} \Lambda^*_P(\pi). \end{aligned}$$

In order to identify the right-hand side with $\Lambda^*_{\mathcal{P}}(\pi)$, we apply a minimax theorem, for example Terkelsen [17], Corollary 2, to the function f on $\Phi_{\mathcal{P}} \times (0, \infty)$ defined by $f(\frac{dP}{dR}, \gamma) = \gamma\pi - \log E_R[e^{-\gamma X_1} \frac{dP}{dR}]$. This yields

$$\begin{aligned} \inf_{P \in \mathcal{P}} \Lambda^*_P(\pi) = \inf_{P \in \mathcal{P}} \sup_{\gamma > 0} \{\gamma\pi - \Lambda_P(\gamma)\} &= \sup_{\gamma > 0} \inf_{P \in \mathcal{P}} \{\gamma\pi - \Lambda_P(\gamma)\} \\ &= \sup_{\gamma > 0} \{\gamma\pi - \Lambda_{\mathcal{P}}(\gamma)\} \\ &= \Lambda^*_{\mathcal{P}}(\pi), \end{aligned}$$

hence $I_{\mathcal{P}}(\pi) = \Lambda^*_{\mathcal{P}}(\pi)$.

2) In order to verify the first equality in (27), recall from Proposition 5.1 that

$$\tfrac{1}{n}\rho_{\mathcal{P},c}(X_1 + \ldots + X_n) = \tfrac{1}{n} E_{Q_c}[-\sum_{i=1}^n X_i] = E_{Q_c}[-X_1],$$

where $Q_c := Q_{\gamma_c} \in \mathcal{Q}_{P_c, X_1 + \ldots + X_n}$ and γ_c is such that $c = H(Q_c|P_c) = H(Q_c|\mathcal{P})$. Our assumptions imply that $\rho_{\mathcal{P},c}(X_1 + \ldots + X_n)$ is strictly increasing and continuous in c. Thus the tolerance level $c_{\pi,n}$ is determined by

$$\tfrac{1}{n}\rho_{\mathcal{P}, c_{\pi,n}}(X_1 + \ldots + X_n) = \pi,$$

i. e., by the two conditions

$$c_{\pi,n} = H(Q_{c_{\pi,n}}|\mathcal{P}) = H(Q_{c_{\pi,n}}|P_{c_{\pi,n}}) \quad \text{and} \quad E_{Q_{c_{\pi,n}}}[-X_1] = \pi.$$

Using part 1), we thus see that

$$\begin{aligned} c_{\pi,n} = \inf_{P \in \mathcal{P}} H(Q_{c_{\pi,n}}|P) &= \inf_{P \in \mathcal{P}} \{E_{Q_{c_{\pi,n}}}[-\gamma_{c_{\pi,n}} \sum_{i=1}^n X_i] - n \log E_P[\exp(-\gamma_{c_{\pi,n}} X_1)]\} \\ &= n \inf_{P \in \mathcal{P}} \{\gamma_{c_{\pi,n}}\pi - \log E_P[\exp(-\gamma_{c_{\pi,n}} X_1)]\} \\ &= n(\gamma_{c_{\pi,n}}\pi - \log \sup_{P \in \mathcal{P}} E_P[\exp(-\gamma_{c_{\pi,n}} X_1)]) \\ &\leq n\Lambda^*_{\mathcal{P}}(\pi) = nI_{\mathcal{P}}(\pi). \end{aligned}$$

On the other hand, the same arguments applied to $P_{c_{\pi,n}}$ yield

$$\begin{aligned}c_{\pi,n} = H(Q_{c_{\pi,n}}|P_{c_{\pi,n}}) &= n(\gamma_{c_{\pi,n}}\pi - \log E_{P_{c_{\pi,n}}}[\exp(-\gamma_{c_{\pi,n}}X_1)]) \\ &= n\sup_{\gamma>0}\{\gamma\pi - \log E_{P_{c_{\pi,n}}}[e^{-\gamma X_1}]\} \\ &= nI_{P_{c_{\pi,n}}}(\pi) \\ &\geq nI_{\mathcal{P}}(\pi)\end{aligned}$$

since the supremum in the second line is attained by $\gamma = \gamma_{c_{\pi,n}}$. □

As in Remark 4.1, (27) translates into the upper bound of the following extension of Cramér's theorem to our present context of model ambiguity. For related results on robust large deviations we refer to Sadowsky [15], Pandit & Meyn [14], and Hu [10].

Corollary 5.3. *For any $\pi > \max_{P\in\mathcal{P}} E_P[-X_1]$ we have*

$$\tfrac{1}{n}\log(\sup_{P\in\mathcal{P}} P[-\tfrac{1}{n}(X_1+\ldots+X_n)\geq\pi]) \leq -I_{\mathcal{P}}(\pi),$$

and

$$\lim_{n\uparrow\infty}\tfrac{1}{n}\log(\sup_{P\in\mathcal{P}} P[-\tfrac{1}{n}(X_1+\ldots+X_n)>\pi]) = -I_{\mathcal{P}}(\pi), \tag{30}$$

where the rate function $I_{\mathcal{P}}$ is given by (28) and coincides with (29).

Proof. For $\alpha_{\pi,n} := \exp(-c_{\pi,n})$, (23) and (26) imply

$$\mathrm{VaR}_{\mathcal{P},\alpha_{\pi,n}-}(X_1+\ldots+X_n) \leq \rho_{\mathcal{P},c_{\pi,n}}(X_1+\ldots+X_n) \leq n\pi,$$

i. e.,

$$\sup_{P\in\mathcal{P}} P[X_1+\ldots+X_n+n\pi\leq 0] \leq \alpha_{\pi,n},$$

hence

$$\tfrac{1}{n}\log(\sup_{P\in\mathcal{P}} P[-\tfrac{1}{n}(X_1+\ldots+X_n)\geq\pi]) \leq -I_{\mathcal{P}}(\pi).$$

In order to verify (30), simply recall that Cramér's theorem yields

$$\varliminf_{n\uparrow\infty}\tfrac{1}{n}\log P[-\tfrac{1}{n}(X_1+\ldots+X_n)>\pi] \geq -\min_{Q:E_Q[-X_1]=\pi} H(Q|P)$$

for any $P\in\mathcal{P}$, hence

$$\begin{aligned}\varliminf_{n\uparrow\infty}\tfrac{1}{n}\log(\sup_{P\in\mathcal{P}} P[-\tfrac{1}{n}(X_1+\ldots+X_n)>\pi]) &\geq \sup_{P\in\mathcal{P}}(-\min_{Q:E_Q[-X_1]=\pi} H(Q|P)) \\ &= -\min_{Q:E_Q[-X_1]=\pi} H(Q|\mathcal{P}) \\ &= -I_{\mathcal{P}}(\pi).\end{aligned}$$

□

Acknowledgements

Thanks are due to Shigeo Kusuoka for a stimulating discussion and to an anonymous referee for several helpful comments. In part the results were obtained while the first author was visiting LMU Munich, funded by LMU Excellent; this support is gratefully acknowledged. For both authors, it is a special pleasure to dedicate this paper to Peter Imkeller on the occasion of his 60th birthday.

References

[1] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, Coherent measures of risk, *Math. Finance* **9** (1999) 203-228.

[2] O. Deprez and H. U. Gerber, On convex principles of premium calculation, *Insurance: Mathematics and Economics* **4** (1985) 179-189.

[3] B. de Finetti, Sul Concetto di Media, *Giornale Dell'Instituto Italiano Degli Attuari* **2** (1931) 369-396.

[4] F. Delbaen, Coherent risk measures, *Cattedra Galileiana, Scuola Normale Superiore, Classe di Scienze, Pisa* (2000).

[5] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd edn. (Springer, 1998).

[6] H. Föllmer and A. Gundel, Robust projections in the class of martingale measures, *Illinois J. Math.* **50** (2006) 439-472.

[7] H. Föllmer and A. Schied, *Stochastic Finance - An Introduction in Discrete Time*, 2nd edn. (de Gruyter, 2004).

[8] M. Frittelli and E. Rosazza Gianin, Putting order in risk measures, *J. Banking & Finance* **26** (2002) 1473-1486.

[9] H. U. Gerber, On additive principles of premium calculation, *Insurance: Mathematics and Economics* **4** (1985) 249-251.

[10] F. Hu, On Cramér's theorem for capacities, *Comptes Rendus Mathematique* **348** (2010) 1009-1013.

[11] R. Kaas, M. Goovaerts, J. Dhaene, and M. Denuit, *Modern Actuarial Risk Theory - Using R*, 2nd edn. (Springer, 2008).

[12] M. Kunze, Verteilungsinvariante konvexe Risikomaße, Diploma thesis, Humboldt-Universität zu Berlin (2003).

[13] S. Kusuoka, On law-invariant coherent risk measures, *Adv. Math. Econ.* **3** (2001) 83-95.

[14] C. Pandit and S. Meyn, Worst-case large deviations asymptotics with application to queuing and information theory, *Stochastic Process. Appl.* **116** (2006) 724-756.

[15] J. S. Sadowsky, Robust large deviations performance analysis for large sample detectors, *IEEE Trans. Inform. Theory* **35** (1989) 917-920.

[16] K. D. Schmidt, *Versicherungsmathematik*, 3rd edn. (Springer, 2009).

[17] F. Terkelsen, Some minimax theorems, *Math. Scand.* **31** (1972) 405-413.