Differential equations connecting VaR and CVaR

Alejandro Balbás*, **Beatriz Balbás**** **and Raquel Balbás*****

* University Carlos III of Madrid. C/ Madrid, 126. 28903 Getafe (Madrid, Spain). alejandro.balbas@uc3m.es.

** University of Alcalá de Henares. Pl. de la Victoria, 2. 28802 Alcalá de Henares (Madrid, Spain). beatriz.balbas@uah.es.

*** University Complutense of Madrid. Somosaguas. 28223 Pozuelo de Alarcón (Madrid, Spain). raquel.balbas@ccee.ucm.es.

Abstract. The Value at Risk (VaR) is a very important risk measure for practitioners, supervisors and researchers. Many practitioners draw on VaR as a critical instrument in Risk Management and other Actuarial/Financial problems, while supervisors and regulators must deal with VaR due to the Basel Accords and Solvency II, among other reasons. From a theoretical point of view VaR presents some drawbacks overcome by other risk measures such as the Conditional Value at Risk (CVaR). VaR is neither differentiable nor sub-additive because it is neither continuous nor convex. On the contrary, CVaR satisfies all of these properties, and this simplifies many analytical studies if VaR is replaced by CVaR. In this paper several differential equations connecting both VaR and CVaR will be presented. They will allow us to address several important issues involving VaR with the help of the CVaR properties. This new methodology seems to be very efficient. In particular, a new VaR Representation Theorem may be found, and optimization problems involving VaR or probabilistic constraints always have an equivalent differentiable optimization problem. Applications in VaR, marginal VaR, CVaR and marginal CVaR estimates will be addressed as well. An illustrative actuarial numerical example will be given.

Key words VaR and CVaR, Differential Equations, VaR Representation Theorem, Risk Optimization and Probabilistic Constraints, Risk and Marginal Risk Estimation.

A.M.S. Classification. 91B06, 91B30, 91G99, 90C90, 34A30.
J.E.L. Classification. C65, G11, G12, G22.

1 Introduction

The Value at Risk (VaR) is a very important risk measure for several reasons. Firstly, regulation (Basel, Solvency, etc.) often deals with VaR. Secondly, many practitioners use VaR in risk management problems because this risk measure has a simple and intuitive economic interpretation and provides us with adequate capital requirements preventing the arrival of bad news and unfavorable scenarios.

From a theoretical point of view VaR presents several drawbacks motivating the use of alternative risk measures such as the Conditional Value at Risk

($CVaR$). In particular, $CVaR$ is continuous and convex and can be represented by means of a linear optimization problem (Rockafellar *et al,* 2006). Consequently, applications of risk measurement in Asset Allocation (Stoyanov *et al*, 2007), Optimal Reinsurance (Centeno and Simoes, 2009) and other classical issues of Actuarial and/or Financial Mathematics become much easier if VaR is replaced by $CVaR$. Nevertheless, there are also reasons justifying the convenience of VaR in some theoretical approaches. Indeed, VaR reflects some kind of consistency (Goovaerts *et al*, 2004), VaR makes sense for risks with unbounded expectations (Chavez-Demoulin *et al,* 2006), the lack of sub-additivity of VaR may be useful in some actuarial applications (Dhaene *et al*, 2008), some pricing and/or hedging methods are easy to interpret if VaR is involved (Goovaerts and Laeven, 2008, or Assa and Karai, 2013), etc.

The main purpose of this paper is to recover analytical properties for problems involving VaR. This seems to be an interesting objective if VaR must be used in theoretical approaches but it is neither differentiable nor convex because it is nether continuous nor sub-additive. In a recent paper Balbás *et al* (2017) showed that VaR can be interpreted in terms of $CVaR$ sensitivity with respect to the level of confidence. These authors dealt with this property in order to develop new methods in VaR optimization. In Section 3 we will draw on the Balbás *et al* result in order to prove that VaR and $CVaR$ are related by a first order linear differential equation with constant coefficients. On the other hand, as said above, the $CVaR$ representation implies that $CVaR$ is the optimal value of a linear optimization problem, and hence, standard results in Mathematical Programming guarantee that the first derivative of the $CVaR$ with respect to the level of confidence is given by a Lagrange multiplier. As a consequence, bearing in mind both the differential equation and the Lagrange multiplier interpretation, in Lemma 4 and Theorem 8 we will give a new representation result for VaR. Indeed, VaR will equal $CVaR$ plus this Lagrange multiplier. To the best of our knowledge, Theorem 8 will be the first VaR Representation Theorem available in the literature.

The rest of the paper will be devoted to explore applications of the representation above. In Section 4 we will give new Karush-Kuhn-Tucker-like and Fritz John-like necessary optimality conditions for optimization problems involving VaR in the objective function. Actually, the VaR representation will allow us to transform the original optimization problem into an equivalent differentiable one. This is important because the minimization of VaR is traditionally a very complex problem due to the lack of continuity and differentiability (Rockafellar and Uryasev, 2000**,** Larsen *et al*, 2002, etc.). Our approach makes the problem differentiable and provides an alternative methodology to that proposed in Balbás *et al* (2017), who used their sensitivity result in order to illustrate that VaR is in the limit the difference of two convex functions. Wozabal (2012) had shown that VaR equals the difference of two convex functions under adequate assumptions and for probability spaces only containing finitely many scenarios.

In Section 5 the VaR Representation Theorem (Theorem 8) will allow us to give Karush-Kuhn-Tucker and Fritz John multipliers for optimization problems involving VaR in the constraints and optimization problems with probabilistic constraints, also called "probabilistically constrained problems" (Lejeune, 2012, Lejeune and Shen, 2016, etc.). These problems are becoming very important in Applied Mathematics and Operations Research. The main reason is that rigid constraints sometimes invalidate many potential solutions and lead to "optimal decisions" with a poor realization of the objective function. However, if some constraints are relaxed, in the sense that they only have to hold with a high enough but lower than one probability, then the optimal objective value reflects a significant improvement.

Section 6 will present further applications of Lemma 4 and Theorem 8. Under appropriate (but general) assumptions, a second order differential equation will also relate VaR and $CVaR$, and this will enable us to establish new methods in VaR, marginal VaR, $CVaR$ and marginal $CVaR$ estimation.

Section 7 will present a practical example illustrating all the findings of previous sections. We will focus on the Optimal Reinsurance Problem. Since this is a classical topic in Actuarial Mathematics, a complete solution is complex and obviously beyond our scope. We will only attempt to point out how the results above may be useful in the study of this problem and other classical topics of Actuarial and/or Financial Mathematics.

Section 8 will conclude the paper.

2 Preliminaries and notations

Consider the probability space $(\Omega, \mathcal{F}, \mathbb{P}_0)$ composed of the set Ω, the $\sigma-$algebra $\mathcal{F}$ and the probability measure $\mathbb{P}_0$. For $1 \leq p < \infty$ the Banach space L^p (also denoted by $L^p(\mathbb{P}_0)$, $L^p(\Omega)$ or $L^p(\Omega, \mathcal{F}, \mathbb{P}_0)$) is composed of the real-valued random variables y such that $\mathbb{E}(|y|^p) < \infty$, $\mathbb{E}()$ representing the mathematical expectation. If $p = \infty$ then L^∞ is the Banach space of essentially bounded random variables. The usual norm of L^p is $\|y\|_p := (\mathbb{E}(|y|^p))^{1/p}$ for $1 \leq p < \infty$ and $\|y\|_\infty := Ess\_Sup\ \{|y|\}$ for $p = \infty$, Ess_Sup denoting "essential supremum". The inclusion $L^p \supset L^P$ holds for $1 \leq p \leq P \leq \infty$. If $1 \leq p < \infty$, $1 < q \leq \infty$ and $1/p + 1/q = 1$ then L^q is the dual space of L^p (Riesz Representation Theorem, Kopp,1984). Moreover L^q may be endowed with the (weak*) topology $\sigma(L^q, L^p)$, which is weaker than the norm topology. Every $\sigma(L^q, L^p)-$closed and bounded subset of L^q is $\sigma(L^q, L^p)-$compact (Alaoglu's Theorem, Kopp,1984). If Ω is a finite set then L^p and L^q become finite-dimensional spaces, and the two given topologies of L^q coincide. Lastly, the space L^0 contains every real-valued and $\mathcal{F}-$measurable random variable. L^0 is a metric (but not Banach) space whose natural distance is given by

$$d(x, y) := \mathbb{E}(Max\ \{|y - x|, 1\}).$$

This metric generates the classical convergence in probability. Further details about metric, Banach or Hilbert spaces of random variables and some other topological results may be found in Kelly (1955), Rudin (1973), Kopp (1984) or Zeidler (1995).

Let us fix a confidence level $1-\alpha \in (0,1)$. As usual, for a random variable $y \in L^0$ the Value at Risk $VaR_{1-\alpha}(y)$ of y is given by[1]

$$VaR_{1-\alpha}(y) := -Inf\{x \in \mathbb{R};\ \mathbb{P}_0(y \le x) > \alpha\}. \tag{2}$$

For $y \in L^1$ the Conditional Value at Risk $CVaR_{1-\alpha}(y)$ of y is

$$CVaR_{1-\alpha}(y) := \frac{1}{\alpha}\int_0^\alpha VaR_{1-t}(y)\,dt. \tag{3}$$

Furthermore, according to the $CVaR$ Representation Theorem (Rockafellar *et al,* 2006), we have that

$$CVaR_{1-\alpha}(y) = Max\ \{-\mathbb{E}(yz);\ z \in \Delta_\alpha\}, \tag{4}$$

where the sub-gradient Δ_α of $CVaR_{1-\alpha}$ is given by

$$\Delta_\alpha = \{z \in L^\infty;\ \mathbb{E}(z) = 1\ and\ 0 \le z \le 1/\alpha\}. \tag{5}$$

Bearing in mind some ideas summarized above, $\Delta_\alpha \subset L^q$ for every $1 < q \le \infty$, and Δ_α is convex and $\sigma(L^q, L^p)-$compact.

Fix $y \in L^1$. The function

$$(0,1) \ni t \to \psi(t) = VaR_{1-t}(y) \in \mathbb{R} \tag{6}$$

is non-increasing, right-continuous and Lebesgue integrable in $(0,1)$. Thus, if one considers the function

$$(0,1) \ni \alpha \to \varphi_y(\alpha) := \alpha CVaR_{1-\alpha}(y) \in \mathbb{R}, \tag{7}$$

which satisfies (see (3))

$$\varphi_y(\alpha) = \int_0^\alpha VaR_{1-t}(y)\,dt, \tag{8}$$

then the First Fundamental Theorem of Calculus guarantees that

$$\varphi_y'^{+}(\alpha) = VaR_{1-\alpha}(y) \tag{9}$$

for every $\alpha \in (0,1)$, $\varphi_y'^{+}$ denoting the right-hand side first order derivative of φ_y.

[1] (2) is the usual definition of $VaR_{1-\alpha}(y)$ if y represents a future random wealth (or income). In many actuarial and financial applications y represents random losses, in which case (2) is replaced by

$$VaR_{1-\alpha}(y) := Sup\ \{x \in \mathbb{R};\ \mathbb{P}_0(y \le x) < 1-\alpha\}. \tag{1}$$

Throughout this paper we will deal with (2), but a parallel analysis holds for (1).

3 VaR Representation Theorem

Representation theorems may play a very important role in risk analysis. This is the reason why several authors have provided representation results for a vast family of risk measures (Rockafellar *et al*, 2006, for expectation bounded risk measures, Kupper and Svindland, 2011, for monotone convex risk measures, etc.). Let us deal with several links between VaR and $CVaR$ in order to give a representation result applying for VaR. To the best of our knowledge, this is the first VaR representation result available in the literature.

A simple change of variable and Expression (9) allow one to prove a main result for this paper.

Lemma 1 *(Ordinary Linear Differential Equation). Fix $y \in L^1$ and the functions*

$$h(\mu) = VaR_{1-e^{\mu}}(y) \quad \text{and} \quad H(\mu) = CVaR_{1-e^{\mu}}(y) \tag{10}$$

for every $-\infty < \mu < 0$. Then,

$$h(\mu) = H(\mu) + H'^{+}(\mu). \tag{11}$$

Consequently, if H has first order derivative at $\mu \in (-\infty, 0)$ then

$$h(\mu) = H(\mu) + H'(\mu). \tag{12}$$

Lastly, if the function φ_y of (7) or (8) has a first order derivative for every $\alpha \in (0,1)$ then (12) and

$$H(\mu) = H(\nu) + \int_{\nu}^{\mu} h(s)\, e^{s-\mu} ds \tag{13}$$

hold for every $-\infty < \mu, \nu < 0$.

Proof. Consider the composition of functions

$$(-\infty, 0) \ni \mu \to \alpha = e^{\mu} \in (0,1) \to CVaR_{1-\alpha}(y) = H(\mu).$$

Obviously,

$$H'^{+}(\mu) = \frac{d\left(CVaR_{1-\alpha}(y)\right)}{d\alpha^{+}} e^{\mu} \tag{14}$$

because $(-\infty, 0) \ni \mu \to \alpha = e^{\mu} \in (0,1)$ is increasing. Expressions (7) and (9), along with $\alpha = e^{\mu}$, lead to

$$h(\mu) = VaR_{1-\alpha}(y) = \varphi_y'^{+}(\alpha) = CVaR_{1-\alpha}(y) + \alpha \frac{d\left(CVaR_{1-\alpha}(y)\right)}{d\alpha^{+}},$$

i.e. (see (14)), $h(\mu) = H(\mu) + H'^{+}(\mu)$. Besides, (12) becomes obvious, and so does (13) because this is the general solution of the ordinary differential equation (12) (Coddington and Levinson, 1955). □

Remark 2 *Fix* $1 \leq p < \infty$ *and* $y \in L^p$*. (4), (5) and (10) imply that* $-H(\mu) = -CVaR_{1-e^{\mu}}(y)$ *is the optimal value of Problem*

$$Min \; \mathbb{E}(yw) \begin{cases} \mathbb{E}(w) = 1 \\ 0 \leq w \leq e^{-\mu} \end{cases} \tag{15}$$

$w \in L^{\infty}$ *being the decision variable. According to Balbás et al (2016) and others, the optimal solution of (15) and the Karush-Kuhn-Tucker multiplier* $(\lambda, \lambda_m, \lambda_M) \in \mathbb{R} \times L^p \times L^p$ *of (15) exist and are characterized by System*

$$\begin{cases} \mathbb{E}(w) = 1 \\ y = \lambda_m - \lambda_M - \lambda \\ (e^{-\mu} - w)\lambda_M = w\lambda_m = 0 \\ 0 \leq w \leq e^{-\mu}, \; 0 \leq \lambda_m, \; 0 \leq \lambda_M \end{cases} \tag{16}$$

In general, the uniqueness of the multiplier $(\lambda, \lambda_m, \lambda_M)$ *cannot be guaranteed. Similarly, the uniqueness of the solution* w *of (15) is false in general.* □

Definition 3 *Fix* $1 \leq p < \infty$*,* $y \in L^p$ *and* $-\infty < \mu_0 < 0$*.* (y, p, μ_0) *will be said to be regular if there exists a function*

$$(-\infty, 0) \ni \mu \to (\lambda(y,p,\mu), \lambda_m(y,p,\mu), \lambda_M(y,p,\mu)) \in \mathbb{R} \times L^p \times L^p$$

such that $(\lambda(y,p,\mu), \lambda_m(y,p,\mu), \lambda_M(y,p,\mu))$ *is a multiplier of (15) for every* $-\infty < \mu < 0$ *and*

$$(-\infty, 0) \ni \mu \to \mathbb{E}(\lambda_M(y,p,\mu)) \in \mathbb{R} \tag{17}$$

is continuous at μ_0*.*[2] □

Lemma 4 *Suppose that* $1 \leq p < \infty$*,* $y \in L^p$*,* $-\infty < \mu_0 < 0$ *and* (y, p, μ_0) *is regular. Consider the functions* h *and* H *of (10) and a multiplier*

$$(\lambda(y,p,\mu), \lambda_m(y,p,\mu), \lambda_M(y,p,\mu))$$

of (15) such that (17) is continuous at μ_0*.Then,*

$$\begin{cases} H'(\mu_0) = -e^{-\mu_0}\mathbb{E}(\lambda_M(y,p,\mu_0)) \\ h(\mu_0) = H(\mu_0) - e^{-\mu_0}\mathbb{E}(\lambda_M(y,p,\mu_0)) \end{cases} \tag{18}$$

Proof. According to classical results about sensitivity analysis in Convex Optimization (Luenberger, 1969, Theorem 1, *pp* 222), if (y, p, μ_0) is regular and $(\lambda, \lambda_m, \lambda_M)$ is the given multiplier of (15) at μ_0, then $-\mathbb{E}(\lambda_M)$ is the first order derivative of $(-\infty, 0) \ni \mu \to -CVaR_{1-e^{\mu}}(y) = -H(\mu) \in \mathbb{R}$ with respect to $e^{-\mu}$ at $e^{-\mu_0}$. Consequently,

$$H' = \frac{dH}{d\mu} = \frac{dH}{d(e^{-\mu})}\frac{d(e^{-\mu})}{d\mu} = \frac{dH}{d(e^{-\mu})}\left(-e^{-\mu}\right) = -e^{-\mu}\mathbb{E}(\lambda_M)$$

holds at $\mu = \mu_0$, and Lemma 1 implies (18). □

[2] Obviously, (y, p, μ_0) will be regular if $(-\infty, 0) \ni \mu \to \lambda_M(y,p,\mu) \in L^p$ is continuous at μ_0. Notice also that the regularity of (y, p, μ_0) implies the regularity of (y, P, μ_0) for every $P \in [1, p]$ and, more generally, for every P such that $y \in L^P$.

Definition 5 *If* $-\infty < \mu < 0$ *then the* $VaR_{1-e^{\mu}}$ *Representation Set* $\delta_{e^{\mu}} \subset L^{\infty} \times L^{1} \times L^{1}$ *will be composed of those* $(w, \lambda_m, \lambda_M)$ *such that* $\mathbb{E}(w) = 1$, $(e^{-\mu} - w)\lambda_M = w\lambda_m = 0$, $0 \leq w \leq e^{-\mu}$, $0 \leq \lambda_0$ *and* $0 \leq \lambda_M$. $\square$

Lemma 6 *Suppose that* $y \in L^1$ *and* $-\infty < \mu < 0$. *Then,*

$$\lambda \geq h(\mu) = VaR_{1-e^{\mu}}(y)$$

holds for every $\lambda \in \mathbb{R}$ *and every* $(w, \lambda_m, \lambda_M) \in \delta_{e^{\mu}}$ *such that* $y = \lambda_m - \lambda_M - \lambda$.

Proof. If $\lambda \in \mathbb{R}$ and $(w, \lambda_m, \lambda_M) \in \delta_{e^{\mu}}$ satisfy the given condition then $\lambda - \mathbb{E}(w(\lambda_m - \lambda_M)) - e^{-\mu}\mathbb{E}(\lambda_M) = \lambda - \mathbb{E}(w\lambda_m) - \mathbb{E}(\lambda_M(e^{-\mu} - w)) = \lambda$ because $w\lambda_m = \lambda_M(e^{-\mu} - w) = 0$. Besides, Remark 2 and (16) show that w solves (15) and $(\lambda, \lambda_m, \lambda_M)$ is a multiplier of this problem. Consequently,

$$\begin{array}{l} \lambda = \lambda - \mathbb{E}(w(\lambda_m - \lambda_M)) - e^{-\mu}\mathbb{E}(\lambda_M) = \\ \mathbb{E}(\lambda w) - \mathbb{E}(w(\lambda_m - \lambda_M)) - e^{-\mu}\mathbb{E}(\lambda_M) = \\ -\mathbb{E}(w(\lambda_m - \lambda_M - \lambda)) - e^{-\mu}\mathbb{E}(\lambda_M) = \\ -\mathbb{E}(wy) - e^{-\mu}\mathbb{E}(\lambda_M) = \\ CVaR_{1-e^{\mu}}(y) - e^{-\mu}\mathbb{E}(\lambda_M) = H(\mu) - e^{-\mu}\mathbb{E}(\lambda_M), \end{array}$$

where H is given by (10). Hence (Lemma 1 and (11)), it is sufficient to see that

$$-e^{-\mu}\mathbb{E}(\lambda_M) \geq H'^{+}(\mu). \tag{19}$$

As in the proof of Lemma 4, classical results about sensitivity analysis in Convex Optimization (Luenberger, 1969, Corollary 1, *pp* 219) apply on Problem (15) and lead to

$$\mathbb{E}(y(z - w)) \geq -\mathbb{E}(\lambda_M(z - w)) \tag{20}$$

for every $z \in L^{\infty}$ such that $0 \leq z$ and $\mathbb{E}(z) = 1$. Since $(w, \lambda_m, \lambda_M) \in \delta_{e^{\mu}}$ we have that $\lambda_M(e^{-\mu} - w) = 0$ and therefore

$$\lambda_M w = \lambda_M e^{-\mu}. \tag{21}$$

Since $\lambda_M \geq 0$, if $\varepsilon > 0$ is small enough and $0 \leq z_{\varepsilon} \leq e^{-(\mu+\varepsilon)}$ one has that $\lambda_M z \leq \lambda_M e^{-(\mu+\varepsilon)}$, and (21) implies that

$$\begin{array}{l} \mathbb{E}(\lambda_M(z_{\varepsilon} - w)) \leq \mathbb{E}\left(\lambda_M e^{-(\mu+\varepsilon)}\right) - \mathbb{E}(\lambda_M w) \\ = \mathbb{E}\left(\lambda_M e^{-(\mu+\varepsilon)}\right) - \mathbb{E}(\lambda_M e^{-\mu}) = \left(e^{-(\mu+\varepsilon)} - e^{-\mu}\right)\mathbb{E}(\lambda_M). \end{array}$$

Consequently, if $\mathbb{E}(z_{\varepsilon}) = 1$ then (20) implies that

$$\mathbb{E}(y(z_{\varepsilon} - w)) \geq \left(e^{-\mu} - e^{-(\mu+\varepsilon)}\right)\mathbb{E}(\lambda_M). \tag{22}$$

Suppose that z_{ε} is the solution of (15) once $e^{-\mu}$ is replaced by $e^{-(\mu+\varepsilon)}$ for $\varepsilon > 0$ and small enough. (10) and (22) imply that

$$-H(\mu + \varepsilon) + H(\mu) \geq \left(e^{-\mu} - e^{-(\mu+\varepsilon)}\right)\mathbb{E}(\lambda_M),$$

i.e.,

$$H(\mu+\varepsilon) - H(\mu) \leq \left(e^{-(\mu+\varepsilon)} - e^{-\mu}\right) \mathbb{E}(\lambda_M).$$

Hence, since $e^{-(\mu+\varepsilon)} - e^{-\mu} < 0$,

$$\frac{H(\mu+\varepsilon) - H(\mu)}{e^{-(\mu+\varepsilon)} - e^{-\mu}} \geq \mathbb{E}(\lambda_M)$$

and therefore

$$\mathbb{E}(\lambda_M) \leq$$

$$Lim_{\varepsilon\to 0^+}\left(\frac{H(\mu+\varepsilon) - H(\mu)}{e^{-(\mu+\varepsilon)} - e^{-\mu}}\right) = H'^{+}(\mu)\, Lim_{\varepsilon\to 0^+}\left(\frac{\varepsilon}{e^{-(\mu+\varepsilon)} - e^{-\mu}}\right),$$

which trivially leads to

$$\mathbb{E}(\lambda_M) \leq -e^{\mu} H'^{+}(\mu),$$

and (19) holds. □

Next let us consider a particular probability space $(\Omega, \mathcal{F}, \mathbb{P}_0)$ and let us give a first VaR representation result.

Lemma 7 *Suppose that $\Omega = (0,1)$, $\mathcal{F}$ is the Borel $\sigma-$algebra of $(0,1)$ and $\mathbb{P}_0$ is the Lebesgue probability measure on $\mathcal{F}$. Suppose that $-\infty < \mu < 0$ and $y \in L^1$ is a right-continuous non-decreasing real valued function on $(0,1)$. Then, $VaR_{1-e^{\mu}}(y)$ is the optimal value of the solvable Problem*

$$Min\ \lambda \begin{cases} y = \lambda_m - \lambda_M - \lambda \\ \lambda \in \mathbb{R} \\ (w, \lambda_m, \lambda_M) \in \delta_{e^{\mu}} \end{cases} \tag{23}$$

$(\lambda, w, \lambda_m, \lambda_M)$ being the decision variable.

Proof. Bearing in mind the properties of y, it is obvious that

$$VaR_{1-e^{\mu}}(y) = -y(e^{\mu}) \tag{24}$$

for every $\mu \in (-\infty, 0)$. Besides, $w(\mu) = e^{-\mu}\mathcal{X}_{(0,e^{\mu})} \in L^{\infty}$ solves (15) for every $\mu \in (-\infty, 0)$, $\mathcal{X}_{(0,e^{\mu})}$ denoting the indicator of $(0, e^{\mu})$. Moreover, it is lastly obvious that

$$\begin{cases} \lambda(\mu) = -y(e^{\mu}) \\ \lambda_m(\mu) = \begin{cases} 0, & on \quad (0, e^{\mu}) \\ y - y(e^{\mu}), & on \quad (e^{\mu}, 1) \end{cases} \\ \lambda_M(\mu) = \begin{cases} y(e^{\mu}) - y, & on \quad (0, e^{\mu}) \\ 0, & on \quad (e^{\mu}, 1) \end{cases} \end{cases} \tag{25}$$

along with $w(\mu)$ above satisfy (16), and therefore $(\lambda(\mu), \lambda_m(\mu), \lambda_M(\mu))$ is a multiplier of (15) and $(\lambda(\mu), w(\mu), \lambda_m(\mu), \lambda_M(\mu))$ is (23)-feasible for every $\mu \in (-\infty, 0)$. Obviously, (24) and (25) imply that $\lambda(\mu) = VaR_{1-e^{\mu}}(y)$, and the result trivially follows from Lemma 6. □

Next we will prove a VaR representation result for a general probability space $(\Omega, \mathcal{F}, \mathbb{P}_0)$.

Theorem 8 *(VaR Representation Theorem) Suppose that* $1 \le p < \infty$, $-\infty < \mu < 0$ *and* $y \in L^p$. $VaR_{1-e^{\mu}}(y)$ *is the optimal value of the solvable Problem (23). Moreover, the result remains true if* $\delta_{e^{\mu}} \subset L^{\infty} \times L^1 \times L^1$ *is replaced by* $\delta_{e^{\mu}} \subset L^{\infty} \times L^p \times L^p$ *in Definition 5.*

Proof. Suppose that we prove the equality between $VaR_{1-e^{\mu}}(y)$ and the optimal value of (23). Then, $\delta_{e^{\mu}} \subset L^{\infty} \times L^1 \times L^1$ may be obviously replaced by $\delta_{e^{\mu}} \subset L^{\infty} \times L^p \times L^p$ because the constraints of $\delta_{e^{\mu}}$ and $y = \lambda_m - \lambda_M - \lambda$ imply that

$$\lambda_m = \begin{cases} y + \lambda & w = 0 \\ 0, & otherwise \end{cases}$$

and

$$\lambda_M = \begin{cases} -y - \lambda & w = e^{-\mu} \\ 0, & otherwise \end{cases}$$

and consequently $(\lambda_m, \lambda_M) \in L^p \times L^p$.

Let us prove the equality between $VaR_{1-e^{\mu}}(y)$ and the optimal value of (23). According to Lemma 6, it is sufficient to see that there exists a (23)-feasible $(\lambda, w, \lambda_m, \lambda_M)$ such that

$$\lambda = VaR_{1-e^{\mu}}(y). \tag{26}$$

Consider the function $\psi(t) = VaR_{1-t}(y)$ of (6). As said in Section 2, ψ is non-increasing, right-continuous and Lebesgue integrable in $(0,1)$. Thus, $-\psi$ is non-decreasing, right-continuous and Lebesgue integrable in $(0,1)$. Bearing in mind that $VaR_{1-e^{\mu}}(y) = VaR_{1-e^{\mu}}(-\psi)$ obviously holds (see (2)), Lemma 7 implies the existence of

$$\left(\lambda, \tilde{w}, \tilde{\lambda}_m, \tilde{\lambda}_M\right) \in \mathbb{R} \times L^{\infty}(0,1) \times L^1(0,1) \times L^1(0,1)$$

satisfying the obvious conditions and such that (26) holds. Denote by F the cumulative distribution function of y, and for $\omega \in \Omega$ define

$$\begin{cases} \lambda_m(\omega) = \tilde{\lambda}_m(F(y(\omega))) \\ \lambda_M(\omega) = \tilde{\lambda}_M(F(y(\omega))) \\ w(\omega) = \tilde{w}(F(y(\omega))) \end{cases}$$

It is easy to see that $(w, \lambda_m, \lambda_M) \in L^{\infty} \times L^1 \times L^1$. Besides,

$$\lambda_m(\omega) = \tilde{\lambda}_m(F(y(\omega))) \ge 0,$$

$$\lambda_M(\omega) = \tilde{\lambda}_M(F(y(\omega))) \ge 0,$$

$$w(\omega) = \tilde{w}(F(y(\omega))) \ge 0,$$

$$w(\omega) = \tilde{w}(F(y(\omega))) \le e^{-\mu},$$

$$w(\omega)\lambda_m(\omega) = \tilde{w}(F(y(\omega)))\tilde{\lambda}_m(F(y(\omega))) = 0$$

and

$$\left(e^{-\mu}-w\left(\omega\right)\right)\lambda_M\left(\omega\right)=\left(e^{-\mu}-\tilde{w}\left(F\left(y\left(\omega\right)\right)\right)\right)\tilde{\lambda}_M\left(F\left(y\left(\omega\right)\right)\right)=0.$$

The fulfillment of (26) implies that the proof will be complete if one shows that

$$y=\lambda_m-\lambda_M-\lambda.$$

Since $-\psi=\tilde{\lambda}_m-\tilde{\lambda}_M-\lambda$, it is sufficient to see that $y\left(\omega\right)=-\psi\left(F\left(y\left(\omega\right)\right)\right)$ almost surely (*i.e.*, out of a $\mathbb{P}_0-$null set), but (2) and (6) show that $x=-\psi\left(F\left(x\right)\right)$ holds almost surely if $\mathbb{R}$ and its Borel $\sigma-$algebra are endowed with the probability measure $y\left(\mathbb{P}_0\right)\left(B\right):=\mathbb{P}_0\left(y\in B\right)$ for every Borel set $B\subset\mathbb{R}$. □

Corollary 9 *Suppose that* $1\leq p<\infty$, $-\infty<\mu<0$ *and* $y\in L^p$*. If the the multiplier* $\left(\lambda,\lambda_m,\lambda_M\right)$ *of Problem (15) is unique then* $VaR_{1-e^{\mu}}\left(y\right)=\lambda$.[3] □

4 VaR optimization

The minimization of risk measures beyond the standard deviation frequently arises in practical applications of risk analysis such as portfolio choice (Gaivoronski and Pflug, 2005, Stoyanov *et al*, 2007, Zakamouline and Koekebbaker, 2009, etc.), pricing and hedging (Assa and Karai, 2013) optimal reinsurance (Cai and Tan, 2007, Cai *et al*, 2016, etc.), etc. Balbás *et al* (2010) gave a general method to optimize convex risk measures and later Balbás *et al* (2017) dealt with their previous findings and Expression (9) in order to provide new methods in VaR optimization. Next let us address again the VaR optimization problem with a different approach. In particular, despite the fact that $L^p\ni y\rightarrow VaR_{1-\alpha}\left(y\right)\in\mathbb{R}$ is not continuous for $1\leq p<\infty$ (Balbás *et al,* 2017, among many others), Theorem 10 below will allow us to transform the minimization of this non-continuous function in the equivalent optimization problem (28) with differentiable (and "almost linear") objective and constraints.

Theorem 10 *Consider* $1\leq p<\infty$, $Y\subset L^p$ *and* $-\infty<\mu<0$*. Consider also the optimization problems*

$$Min\ \left\{VaR_{1-e^{\mu}}\left(y\right);\ y\in Y\right\} \tag{27}$$

and

$$\left\{\begin{array}{ll} Min & \lambda \\ & y=\lambda_m-\lambda_M-\lambda \\ & \lambda\in\mathbb{R},\ y\in Y,\ \ \left(w,\lambda_m,\lambda_M\right)\in\delta_{e^{\mu}} \end{array}\right. \tag{28}$$

$y\in L^p$ *and* $\left(y,w,\lambda,\lambda_m,\lambda_M\right)\in L^p\times L^{\infty}\times\mathbb{R}\times L^p\times L^p$ *being the decision variables, respectively. Then,* $y^*\in Y$ *solves (27) if and only if there exists*

[3] Obviously, if w solves (15), then

$$CVaR_{1-e^{\mu}}\left(y\right)=-\mathbb{E}\left(wy\right)=\lambda-\mathbb{E}\left(w\left(\lambda_m-\lambda_M\right)\right)$$

holds even if the multiplier of (15) is not unique.

$(w^*,\lambda^*,\lambda_m^*,\lambda_M^*) \in L^\infty \times \mathbb{R} \times L^p \times L^p$ such that $(y^*,w^*,\lambda^*,\lambda_m^*,\lambda_M^*)$ solves (28). If so, the optimal values of both problems coincide,

$$CVaR_{1-e^\mu}(y^*) = \lambda^* - \mathbb{E}\left(w^*\left(\lambda_m^* - \lambda_M^*\right)\right) \tag{29}$$

and the right-hand side derivative of H in (10) if $y = y^*$ satisfies

$$\frac{d\left(CVaR_{1-e^\mu}(y^*)\right)}{d\mu^+} = -e^{-\mu}\mathbb{E}\left(\lambda_M^*\right). \tag{30}$$

Moreover, $\delta_{e^\mu} \subset L^\infty \times L^1 \times L^1$ may replaced by $\delta_{e^\mu} \subset L^\infty \times L^p \times L^p$ in Definition 5.

Proof. Let us prove (29) and (30) because the rest of properties are obvious implications of Theorem 8. Since $(y^*,w^*,\lambda^*,\lambda_m^*,\lambda_M^*)$ must satisfy the constraints of (28), Remark 2 implies that w^* solves (15) and $(\lambda^*,\lambda_m^*,\lambda_M^*)$ is a multiplier of this problem when $y = y^*$. Hence, (4) implies that

$$\begin{aligned} &CVaR_{1-e^\mu}(y^*) = -\mathbb{E}\left(y^* w^*\right) = -\mathbb{E}\left(\left(\lambda_m^* - \lambda_M^* - \lambda^*\right)w^*\right) \\ &= \lambda^* - \mathbb{E}\left(w^*\left(\lambda_m^* - \lambda_M^*\right)\right). \end{aligned}$$

Thus, Lemma 1 implies that

$$H'^+(\mu) = h(\mu) - H(\mu) = \lambda^* - \left(\lambda^* - \mathbb{E}\left(w^*\left(\lambda_m^* - \lambda_M^*\right)\right)\right) = \mathbb{E}\left(w^*\left(\lambda_m^* - \lambda_M^*\right)\right).$$

Since $w^*\lambda_m^* = 0$ and $w^*\lambda_M^* = e^{-\mu}\lambda_M^*$ we have $H'^+(\mu) = -e^{-\mu}\mathbb{E}\left(\lambda_M^*\right)$. □

Remark 11 *(Karush-Kuhn-Tucker (KKT) and Fritz John (FJ) optimality conditions). As said above, despite the fact that Problem (27) is neither differentiable nor convex, the equivalent Problem (28) becomes differentiable. Consequently, the KKT or FJ necessary optimality conditions (Mangasarian and Fromovitz, 1967) of (28) will allow us to provide (27) with optimality conditions too. In order to make that possible, we need to represent the constraint $y \in Y$ by means of equalities/inequalities. Thus, suppose that B is a Banach space, B' is its dual, $A \subset L^p$ is an open set, $\varphi : A \to B$ is a continuously Frechet differentiable function and*

$$Y = \left\{y \in A \subset L^p;\ \varphi(y) = 0\right\}.^4 \tag{31}$$

In order to select appropriate multipliers we have to analyze the Banach spaces where the constraints of (28) are valued. Bearing in mind Definition 5, (28) is

[4] A totally similar analysis would apply if Y were represented by inequality constraints.

the same problem as

$$
\left\{
\begin{array}{ll}
Min & \lambda \\
 & \mathbb{E}\left(w\right)=1 \\
 & \varphi\left(y\right)=0 \\
 & y=\lambda_m-\lambda_M-\lambda \\
 & w\lambda_m=0 \\
 & \left(e^{-\mu}-w\right)\lambda_M=0 \\
 & w\leq e^{-\mu} \\
 & w\geq 0 \\
 & \lambda_m\geq 0 \\
 & \lambda_M\geq 0
\end{array}
\right. \tag{32}
$$

The first constraint of (32) is valued on $\mathbb{R}$*, the second one on* B *and the rest of constraints are valued in* L^p*. Accordingly, the related multipliers should satisfy (Luenberger, 1969)*

$$
\begin{array}{l}
\left(\Gamma_1,b',\Gamma_2,\Gamma_3,\Gamma_4,\Gamma_5,\Gamma_6,\Gamma_7,\Gamma_8\right)\in \\
\mathbb{R}\times B'\times L^q\times L^q\times L^q\times L^q\times L^q\times L^q\times L^q
\end{array}
$$

and the components associated with inequality constraints must be non-negative. The Lagrangian function becomes (Luenberger, 1969)

$$
\begin{array}{l}
\mathcal{L}\left(y,w,\lambda,\lambda_m,\lambda_M,\Gamma_1,b',\Gamma_2,\Gamma_3,\Gamma_4,\Gamma_5,\Gamma_6,\Gamma_7,\Gamma_8\right)= \\
\lambda+\Gamma_1\left(\mathbb{E}\left(w\right)-1\right)+\left\langle\varphi\left(y\right),b'\right\rangle \\
+\mathbb{E}\left(\Gamma_2\left(y-\lambda_m+\lambda_M+\lambda\right)\right)+\mathbb{E}\left(\Gamma_3\lambda_m w\right)+\mathbb{E}\left(\Gamma_4\lambda_M\left(e^{-\mu}-w\right)\right) \\
+\mathbb{E}\left(\Gamma_5\left(w-e^{-\mu}\right)\right)-\mathbb{E}\left(\Gamma_6 w\right)-\mathbb{E}\left(\Gamma_7\lambda_m\right)-\mathbb{E}\left(\Gamma_8\lambda_M\right).
\end{array} \tag{33}
$$

Consequently, denoting by $\left\langle\varphi'\left(y\right),b'\right\rangle$ *the composition of* b' *with the Frechet differential* $\varphi'\left(y\right)$ *of* φ *at* y*, the* KKT *optimality conditions for (32) (and (27)) become,*[5]

$$
\left\{
\begin{array}{ll}
\left\langle\varphi'\left(y\right),b'\right\rangle+\Gamma_2=0 & (y) \\
\Gamma_1+\Gamma_3\lambda_m-\Gamma_4\lambda_M+\Gamma_5-\Gamma_6=0 & (w) \\
1+\mathbb{E}\left(\Gamma_2\right)=0 & (\lambda) \\
-\Gamma_2+\Gamma_3 w-\Gamma_7=0 & (\lambda_m) \\
\Gamma_2+\Gamma_4\left(e^{-\mu}-w\right)-\Gamma_8=0 & (\lambda_M) \\
\varphi\left(y\right)=0 & (b') \\
\mathbb{E}\left(w\right)=1 & (\Gamma_1)
\end{array}
\right.
\left\{
\begin{array}{ll}
y=\lambda_m-\lambda_M-\lambda & (\Gamma_2) \\
w\lambda_m=0 & (\Gamma_3) \\
\left(e^{-\mu}-w\right)\lambda_M=0 & (\Gamma_4) \\
0\leq w & (\Gamma_5) \\
e^{-\mu}-w\geq 0 & (\Gamma_6) \\
\lambda_m\geq 0 & (\Gamma_7) \\
\lambda_M\geq 0 & (\Gamma_8)
\end{array}
\right. \tag{34}
$$

The FJ *multipliers arise if one slightly modifies (33) and adds the new*

[5] The partial differential with respect to several variables must vanish. These variables are indicated on the right hand side. Notice that, as usual with the KKT multipliers, those partial differentials with respect to the multipliers related to inequalities (Γ_5, Γ_6, Γ_7 and Γ_8) must be replaced by the associated inequalities themselves.

multiplier $\Gamma_0 \in \mathbb{R}$. The Lagrangian function becomes

$$\begin{array}{l} \mathcal{L}\left(y, w, \lambda, \lambda_m, \lambda_M, \Gamma_0, \Gamma_1, b', \Gamma_2, \Gamma_3, \Gamma_4, \Gamma_5, \Gamma_6, \Gamma_7, \Gamma_8\right) = \\ \Gamma_0 \lambda + \Gamma_1 \left(\mathbb{E}\left(w\right) - 1\right) + \left\langle \varphi\left(y\right), b' \right\rangle \\ + \mathbb{E}\left(\Gamma_2\left(y - \lambda_m + \lambda_M + \lambda\right)\right) + \mathbb{E}\left(\Gamma_3 \lambda_m w\right) + \mathbb{E}\left(\Gamma_4 \lambda_M \left(e^{-\mu} - w\right)\right) \\ + \mathbb{E}\left(\Gamma_5\left(w - e^{-\mu}\right)\right) - \mathbb{E}\left(\Gamma_6 w\right) - \mathbb{E}\left(\Gamma_7 \lambda_m\right) - \mathbb{E}\left(\Gamma_8 \lambda_M\right). \end{array}$$

and consequently, the FJ optimality conditions of both (27) and (32) are similar to (34) once the third equality $1 + \mathbb{E}\left(\Gamma_2\right) = 0$ is replaced by $\Gamma_0 + \mathbb{E}\left(\Gamma_2\right) = 0$. □

5 Optimization with VaR-constraints and probabilistic constraints

Let us deal with optimization problems with $VaR-$linked constraints. These problems are usual in applications of risk analysis such as portfolio choice (Zhao and Xiao, 2016), optimal reinsurance (Centeno and Simoes, 2009), combinations of both optimal investment and optimal reinsurance (Peng and Wang, 2016), etc. Moreover, as will be seen, probabilistic constraints are a particular case of $VaR-$linked constraints.

Suppose that $-\infty < \mu < 0$, $1 \leq p < \infty$, $\tilde{B}$ is a Banach space, $Y \subset A \subset \tilde{B}$ and A is an open set, and $\Psi : A \rightarrow \mathbb{R}$ and $\phi : A \rightarrow L^p$ are continuously differentiable functions. Consider Problem,[6]

$$Min\ \Psi\left(y\right) \left\{ \begin{array}{l} VaR_{1-e^{\mu}}\left(\phi\left(y\right)\right) \leq 0 \\ y \in Y \end{array} \right. \tag{35}$$

If $c \in \mathbb{R}$ then notice that $VaR_{1-e^{\mu}}\left(\phi\left(y\right)\right) \leq c$ is equivalent to

$$VaR_{1-e^{\mu}}\left(\phi\left(y\right) + c\right) \leq 0$$

because $VaR_{1-e^{\mu}}$ is translation invariant (Rockafellar and Uryasev, 2000), and therefore (35) still applies if 0 is replaced by a different number.

Theorem 8 also allows us to find a differentiable optimization problem equivalent to (35).

Theorem 12 *Under the notations above, fix $\mu \in (-\infty, 0)$. Consider Problem*

$$Min\ \Psi\left(y\right) \left\{ \begin{array}{l} \phi\left(y\right) = \lambda_m - \lambda_M - \lambda \\ \lambda \leq 0 \\ \lambda \in \mathbb{R},\ y \in Y,\ \ \left(w, \lambda_m, \lambda_M\right) \in \delta_{e^{\mu}} \end{array} \right. \tag{36}$$

$\left(y, \lambda, w, \lambda_m, \lambda_M\right) \in \tilde{B} \times \mathbb{R} \times L^{\infty} \times L^p \times L^p$ being the decision variable. Then, $y^* \in Y$ solves (35) if and only if there exists $\left(\lambda^*, w^*, \lambda_m^*, \lambda_M^*\right) \in \mathbb{R} \times L^{\infty} \times L^p \times L^p$ such

[6] As in (31), Y may be given by

$$Y = \{y \in A \subset B;\ \varphi\left(y\right) = 0\} \quad or \quad Y = \{y \in A \subset B;\ \varphi\left(y\right) \leq 0\}$$

B being a (ordered, if necessary) Banach space and $\varphi : A \rightarrow B$ being a continuously Frechet differentiable function.

that $(y^*, w^*, \lambda^*, \lambda_m^*, \lambda_M^*)$ *solves (36), and the optimal values of both problems coincide. Moreover,*

$$CVaR_{1-e^{\mu}}(\phi(y^*)) = \lambda^* - \mathbb{E}(w^*(\lambda_m^* - \lambda_M^*))$$

holds. Lastly, if $(\lambda^*, \lambda_m^*, \lambda_M^*)$ *is the unique multiplier of (15) when* y *is replaced by* $\phi(y^*)$, *then*

$$\frac{d(CVaR_{1-e^{\mu}}(\phi(y^*)))}{d\mu^+} = -e^{-\mu}\mathbb{E}(\lambda_M^*)$$

holds too.

Proof. Bearing in mind Corollary 9, the proof of this result is analogous to the proof of Theorem 10. □

Optimization problems with probabilistic constraints, also called "probabilistically constrained optimization problems" (Lejeune, 2012, Lejeune and Shen, 2016, etc.), are becoming very important in Applied Mathematics and Operations Research. The main reason is that rigid constraints sometimes invalidate many potential solutions and lead to "optimal decisions" with a poor realization of the objective function. However, if some constraints are relaxed, in the sense that they only have to hold with a high enough but lower than one probability, then the optimal objective value reflects a significant improvement. Let us show that probabilistic restrictions are equivalent to $VaR-$linked restrictions (Proposition 13 below), and consequently every problem with probabilistic constraints may be solved with the aid of Theorem 12.

Suppose that $-\infty < \mu < 0$, $1 \le p < \infty$, $\tilde{B}$ is a Banach space, $Y \subset A \subset \tilde{B}$ and A is an open set, and $\Psi : A \to \mathbb{R}$ and $\phi : A \to L^p$ are continuously differentiable functions. Consider Problem,[7]

$$Min\ \Psi(y) \begin{cases} \mathbb{P}_0(\phi(y) < 0) \le e^{\mu} \\ y \in Y \end{cases} \tag{37}$$

Notice that Constraint $\mathbb{P}_0(\phi(y) < c) \le e^{\mu}$ is equivalent to $\mathbb{P}_0(\phi(y) - c < 0) \le e^{\mu}$, and therefore (37) still applies if 0 is replaced by a different number. Notice also that $\mathbb{P}_0(\phi(y) < 0) \le e^{\mu}$ is equivalent to $\mathbb{P}_0(\phi(y) \ge 0) \ge 1 - e^{\mu}$ (or $\mathbb{P}_0(-\phi(y) \le 0) \ge 1 - e^{\mu}$), and consequently it contains every restriction imposing an inequality with a "high" probability.

Proposition 13 *Under the given notations,* $\mathbb{P}_0(\phi(y) < 0) \le e^{\mu}$ *holds if and only if* $VaR_{1-e^{\mu}}(\phi(y)) \le 0$ *holds, and therefore (35) and (37) are exactly the same problem.*

Proof. According to (2), $VaR_{1-e^{\mu}}(\phi(y)) \le 0$ holds if and only if the implication $(\mathbb{P}_0(\phi(y) \le x) > e^{\mu}) \Longrightarrow x \ge 0$ holds. Hence, if $VaR_{1-e^{\mu}}(\phi(y)) \le 0$ then for $n = 1, 2, 3, ...$ we have the inequality $\mathbb{P}_0\left(\phi(y) \le -\frac{1}{n}\right) \le e^{\mu}$, and therefore

$$\mathbb{P}_0(\phi(y) < 0) = Lim_{n \to \infty}\left(\mathbb{P}_0\left(\phi(y) \le -\frac{1}{n}\right)\right) \le e^{\mu}.$$

[7] The footnote of Problem (35) about the definition of Y still applies.

Conversely, suppose that $\mathbb{P}_0(\phi(y) < 0) \leq e^{\mu}$. If $\mathbb{P}_0(\phi(y) \leq x) > e^{\mu}$ then $x \geq 0$. Hence, $Inf\{x;\ \mathbb{P}_0(\phi(y) \leq x) > e^{\mu}\} \geq 0$, and (2) leads to the inequality $VaR_{1-e^{\mu}}(\phi(y)) \leq 0$. $\square$

Remark 14 *Theorem* 12 *and Proposition* 13 *above enable us to provide the non differentiable Problems (35) and (37) with KKT and FJ optimality conditions. Since the analysis is analogous to that of Remark* 11*, we will skip details and shorten the exposition.* $\square$

6 Further applications

Lemmas 1 and 4, along with the differential equations (11) and (12), may have further applications, some of them beyond optimization problems or VaR representation. In order to illustrate this fact, in this section we will select two examples related to the estimation and sensitivity of both VaR and $CVaR$. A third application will be devoted to heavy tailed risks.

6.1 Marginal CVaR with respect to the confidence level

Expression (11) trivially implies that $H'^{+}(\mu) = h(\mu) - H(\mu)$ for $-\infty < \mu < 0$, and therefore, if $\alpha = e^{\mu}$

$$\frac{d\left(CVaR_{1-\alpha}(y)\right)}{d\alpha^{+}} = \frac{d\left(CVaR_{1-\alpha}(y)\right)}{d\mu^{+}}\frac{d\mu}{d\alpha}$$
$$= \left(h(\mu) - H(\mu)\right)\frac{1}{\alpha} = \frac{VaR_{1-\alpha}(y) - CVaR_{1-\alpha}(y)}{\alpha}.$$

In other words, if $y \in L^1$ and one estimates both $VaR_{1-\alpha}(y)$ and $CVaR_{1-\alpha}(y)$, then the right hand side marginal $CVaR$ with respect to the level of confidence is given by the proportion $\left(VaR_{1-\alpha}(y) - CVaR_{1-\alpha}(y)\right)/\alpha$. If (6) is continuous in α then $H'(\mu)$ does exist and this marginal value of the $CVaR$ also holds on the left hand side, *i.e.*,

$$\frac{d\left(CVaR_{1-\alpha}(y)\right)}{d\alpha} = \frac{VaR_{1-\alpha}(y) - CVaR_{1-\alpha}(y)}{\alpha}. \tag{38}$$

6.2 The second order differential equation

Consider $-\infty \leq a < b \leq \infty$, $0 \leq c < d \leq 1$ and $y \in L^1$, and suppose that $F : (a,b) \to (c,d)$ and $f : (a,b) \to (0,\infty)$ are the restrictions to (a,b) of the cumulative distribution function and the density function of y, respectively. Suppose that $F^{-1} : (c,d) \to (a,b)$ is the inverse of F and it is differentiable. It is known that (2) leads to $VaR_{1-\alpha}(y) = -F^{-1}(\alpha)$ for $c < \alpha = e^{\mu} < d$. This equality implies the continuity of (6) in (c,d), and then $H'(\mu)$ does exist and (12) holds for $Log(c) < \mu < Log(d)$.[8] Consequently, (12) leads

[8] If necessary, take $Log(0) = -\infty$.

to $H + H' = -F^{-1}(e^{\mu})$, *i.e.*, H' is differentiable and $F(-H - H') = e^{\mu}$. Hence, computing derivatives, $f(-H - H')(-H' - H'') = e^{\mu}$, which implies that $f(-H - H') \neq 0$ and

$$H'' = -H' - \frac{e^{\mu}}{f(-H - H')}. \tag{39}$$

This second order ordinary differential equation may be easily transformed into a system of two first order differential equations with the usual change of variable $(H_1, H_2) = (H, H')$. We will obtain

$$\left\{ \begin{array}{l} H_1' = H_2 \\ H_2' = -H_2 - \dfrac{e^{\mu}}{f(-H_1 - H_2)} \end{array} \right. \tag{40}$$

Both (39) and (40) hold for $Log(c) < \mu < Log(d)$ if F^{-1} exists on (c, d) and both F and F^{-1} are differentiable on (a, b) and (c, d) respectively.

6.3 VaR and CVaR estimation and sensitivity

Under appropriate conditions one can go beyond Expression (38). If we know the $CVaR$ and the marginal $CVaR$ for a specific level of confidence e^{μ_0} then we can give $CVaR$, marginal $CVaR$, VaR and marginal VaR for every level of confidence. Indeed, suppose that the assumptions guaranteeing the fulfillment of (40) hold. Suppose that we are able to estimate both $H(\mu_0) = CVaR_{1-e^{\mu_0}}(y)$ and $H'(\mu_0)$ for some $y \in L^1$ and some $Log(c) < \mu_0 < Log(d)$. Then, the classical Picard–Lindelöf Theorem (Coddington and Levinson, 1955) about the existence and uniqueness of solution for ordinary differential equations will enable us to give approximations of the global solution $(H_1, H_2) = (H, H')$ of (40). In other words, we will have approximations of both the $CVaR_{1-e^{\mu}}$ and the marginal $CVaR_{1-e^{\mu}}$ of y for a confidence level such that $Log(c) < \mu < Log(d)$. Moreover, $h(\mu) = VaR_{1-e^{\mu}}(y)$ and the marginal $VaR_{1-e^{\mu}}$ of y will also be easily estimated. Indeed, for every $\mu \in (Log(c), Log(d))$ we have that $h(\mu)$ will be given by (12), and the marginal VaR, $h'(\mu) = H'(\mu) + H''(\mu)$, will be easily obtained from (39).

In order to reach an approximation of the global solution of (40) we have several alternatives. Firstly, we can transform (40) into an integral equation and then iterate the integral operator in order to obtain approximations of its fixed point (Picard–Lindelöf Theorem, Coddington and Levinson, 1955). Secondly, we can deal with standards in Numerical Analysis such as the Eüler method or the Runge - Kutta method and its classical extensions (Butcher, 2008).

Though for most of them there are closed formulas, this general methodology firstly applies for many classical continuous distributions (normal, log-normal, Weibull, Student's t, Fisher–Snedecor's F, Gamma, Beta, Pareto, etc.). Secondly, this methodology also applies for some compositions/mixtures and other practical distributions related to the examples above. Lastly, and maybe more importantly, the methodology also applies for the solutions of many optimization problems involving VaR or $CVaR$. For instance, if one solves Problem (32) in order to solve (27) (respectively, one solves (36) in order to solve (35) or (37)), and the solution y^* of (27) (respectively, $\phi(y^*)$, where y^* solves (35)) satisfies the conditions guaranteeing the fulfillment of (39) and (40),[9] then Theorem 10 (respectively, Theorem 12) will provide us with the $CVaR$ and marginal $CVaR$ of y^* (respectively, $\phi(y^*)$) at a given confidence level, and therefore we will have the instruments so as to compute $CVaR$, marginal $CVaR$, VaR and marginal VaR of y^* (respectively, $\phi(y^*)$) at every confidence level. Analogously, in $CVaR$ minimization problems, *i.e.*, if $CVaR$ replaces VaR in (27), the arguments of Lemma 4 and Expression (18) still apply, and consequently, if the solution y^* satisfies the conditions of Section 6.2, (12), (39) and (40) will allow us to estimate $VaR(y)$, $CVaR(y)$ and their marginal values for infinitely many levels of confidence.

6.4 VaR representation for heavy tailed risks

As indicated in the introduction, VaR makes sense for heavy tailed risks with unbounded expectations. Though some other risk measures may also apply for some of these risks (Kupper and Svindland, 2011), the most important ones ($CVaR$, the weighted $CVaR$ of Rockafellar *et al,* 2006, etc.) will become infinite for risks beyond L^1. Actually, to the best of our knowledge, $VaR_{1-e^\mu}(y)$ is the unique "popular" risk measure applying for every $y \in L^0$, and for that reason several authors have proposed the use of VaR_{1-e^μ} in their analytical studies of problems involving heavy tailed risks. For instance, Chavez-Demoulin *et al* (2006) dealt with VaR_{1-e^μ} in order to address some Operational Risk topics.

The ideas above may justify the interest of extensions of Theorem 8 and (23) beyond L^1. If $-\infty < \mu < 0$, $y \in L^0$, $VaR_{1-e^\mu}(y) > 0$, $\theta_1 \geq 0$, and one considers

$$y_{\theta_1} = \begin{cases} \theta_1, & \theta_1 \leq y \\ y, & -\theta_1 < y < \theta_1 \\ -\theta_1, & y \leq -\theta_1 \end{cases} \tag{41}$$

then it is easy to see that $VaR_{1-e^\mu}(y) = VaR_{1-e^\mu}(y_{\theta_1})$ for

$$-\theta_1 < -VaR_{1-e^\mu}(y) - 1.$$

Furthermore, $y_{\theta_1} \in L^p$ for every $1 \leq p \leq \infty$. Since $\theta_1 \to VaR_{1-e^\mu}(y_{\theta_1})$ is obviously a non-decreasing function, one has that

$$VaR_{1-e^\mu}(y) = Max\ \{VaR_{1-e^\mu}(y_{\theta_1});\ \theta_1 \geq 0\}$$

or, equivalently, $VaR_{1-e^\mu}(y)$ is the optimal value of

$$\begin{cases} Max\ \theta_2 \\ \theta_1 \geq 0 \\ \theta_2 \leq VaR_{1-e^\mu}(y_{\theta_1}) \end{cases} \tag{42}$$

[9]Actually, these conditions could be slightly relaxed (Mohammadi *et al*, 2017).

$(\theta_1, \theta_2) \in \mathbb{R}^2$ being the decision variable. Besides, (23) applies for $VaR_{1-e^\mu}(y_{\theta_1})$, and consequently (42) implies that $VaR_{1-e^\mu}(y)$ equals the optimal value of

$$\left\{ \begin{array}{l} Max\ \theta_2 \\ \theta_1 \geq 0 \\ \theta_2 \leq \lambda, \left\{ \begin{array}{l} \forall (\lambda, w, \lambda_m, \lambda_M) \in \mathbb{R} \times \delta_{e^\mu} \\ with\ y_{\theta_1} = \lambda_m - \lambda_M - \lambda \end{array} \right. \end{array} \right. \tag{43}$$

$(\theta_1, \theta_2) \in \mathbb{R}^2$ being the decision variable. This representation of $VaR_{1-e^\mu}(y)$ remains true if L^1 is replaced by L^p, $1 < p < \infty$, in Definition 5. Moreover, it is obvious that "the definition of y_{θ_1} is not unique", in the sense that many modifications of (41) will still imply the fulfillment of (43).

7 Illustrative actuarial example

Let us present an application showing the potential practical interest of the findings of this paper. In particular, let us deal with the optimal reinsurance problem. The purpose of this section is merely illustrative. Though we will focus on a classical topic in Actuarial Mathematics, a complete solution is complex and is obviously beyond our scope. We will only attempt to point out how the results above may be useful in the analysis of this problem.[10]

Consider the random variable u reflecting claims within a given time period $[0, T]$. The insurer has to select the retained (u_r) and ceded (u_c) risks satisfying $u = u_r + u_c$. In classical literature both u_r and u_c are given by measurable functions of u, *i.e.*, $u_r = g(u)$ and $u_c = u - g(u)$, so without loss of generality one can assume that $\Omega = (0,1)$ and $\mathbb{P}_0$ is the Lebesgue probability measure on $(0,1)$. Moreover, though it is not necessary, in order to simplify the exposition we will accept the existence of $-\infty < \zeta < \Xi < \infty$ such that $u : (0,1) \rightarrow (\zeta, \Xi)$ is a strictly increasing bijection such that $u \in L^2 = L^2(0,1)$, and u and u^{-1} are continuously differentiable.[11]

In order to prevent moral hazard, it is usually imposed that the three random risks, u, u_r and u_c, must be co-monotone.[12] Nevertheless, for the stop-loss contract with deductible $D \in \mathbb{R}$, *i.e.*, $u_r = Max\ \{D, u\}$ and $u_c = (u - D)^+$,

[10] An interesting survey about the State of the Art a few years ago in optimal reinsurance may be found in Centeno and Simoes (2009). Recent approaches may be found in Balbas *et al* (2015) or Cai *et al* (2016), among many others.

[11] If $\zeta < u < \Xi$ holds *a.s.*, then it is sufficient to take $u(\omega) = F^{-1}(\omega)$ for $0 < \omega < 1$, F being the cumulative distribution function of u. If $\zeta = -\infty$ or $\Xi = \infty$ then the analysis below will remain true under appropriate modifications. For instance, if $\zeta = \infty$ the operator J below should become

$$J(x)(t) = \int_{t_0}^{t} x(u(s))\, u'(s)\, ds = \int_{u(t_0)}^{u(t)} x(r)\, dr,$$

where $0 < t_0 < 1$ must be fixed. Nevertheless, since u represents claims, $0 < u$ (and therefore $0 \leq \zeta$) is natural and not at all restrictive, and $\Xi < \infty$ may be justified because global claims cannot be larger than the value of the insured goods.

[12] Recal that the random variables v and w are said to be co-monotone if the probability of the event $(v(\omega_1) - v(\omega_2))(w(\omega_1) - w(\omega_2)) \geq 0$ equals one.

they are co-monotone but the insurer has no incentives to verify the absence of fraud once $u \geq D$ holds, and therefore the reinsurer might be facing moral hazard. In order to prevent this caveat Balbás *et al* (2015) proposed to deal with the sensitivity (or first order derivative $g' = du_r/du$) of u_r with respect to u as the decision variable. We will follow this approach in this example, and therefore we will consider the Lebesgue measure on (ζ, Ξ), the Banach space $L^\infty(\zeta, \Xi)$ and the linear and continuous operator

$$\begin{array}{lcl} J : L^\infty(\zeta, \Xi) & \to & L^2(0,1) \\ x & \to & J(x)(t) = \int_0^t x(u(s))\, u'(s)\, ds. \end{array}$$

Notice that with the change of variable $r = u(s)$, $dr = u'(s)\, ds$, $J(x)$ may be also given by

$$J(x)(t) = \int_0^t x(u(s))\, u'(s)\, ds = \int_\zeta^{u(t)} x(r)\, dr, \tag{44}$$

and therefore the First Fundamental Theorem of Calculus guarantees that $x = d(J(x))/du$ holds almost everywhere in (ζ, Ξ). In particular, if $u_r = J(x)$ is the retained risk then $x = du_r/du$ holds almost everywhere, and therefore x coincides with the decision variable proposed in Balbás *et al* (2015).

For the constant function $x = 1$ we have

$$J(1)(t) = \int_\zeta^{u(t)} dr = u(t) - \zeta, \tag{45}$$

for $0 < t < 1$.

Bearing in mind that $L^2(0,1)$ is its own dual space, if $L^\infty(\zeta, \Xi)^*$ denotes the dual space of $L^\infty(\zeta, \Xi)$, the adjoint of J will be the linear and continuous operator $J^* : L^2(0,1) \to L^\infty(\zeta, \Xi)^*$ characterized by

$$\langle J(x), y \rangle = \langle x, J^*(y) \rangle \tag{46}$$

for every $x \in L^\infty(\zeta, \Xi)$ and every $y \in L^2(0,1)$, $\langle ., . \rangle$ denoting the classical bilinear product. Obviously,

$$\begin{array}{l} \langle J(x), y \rangle = \int_0^1 J(x)(t)\, y(t)\, dt = \int_0^1 y(t) \left(\int_0^t x(u(s))\, u'(s)\, ds \right) dt \\ = \int_0^1 x(u(s))\, u'(s) \left(\int_s^1 y(t)\, dt \right) ds. \end{array}$$

With the change of variable $r = u(s)$, $dr = u'(s)\, ds$, we get

$$\langle J(x), y \rangle = \int_\zeta^\Xi x(r) \left(\int_{u^{-1}(r)}^1 y(t)\, dt \right) dr,$$

and (46) implies that

$$J^*(y)(s) = \int_{u^{-1}(s)}^1 y(t)\, dt \tag{47}$$

for $\zeta < s < \Xi$. It is also possible to obtain a second expression for J^*. Indeed, with the change of variable $t = u^{-1}(r)$, $dt = \dfrac{dr}{u'(u^{-1}(r))}$, Expression (47) implies that

$$J^*(y)(s) = \int_s^{\Xi} \frac{y(u^{-1}(r))}{u'(u^{-1}(r))} dr,$$

and therefore

$$J^*(y)(s) = \int_{u^{-1}(s)}^{1} y(t)\, dt = \int_s^{\Xi} \left(\frac{y(u^{-1}(t))}{u'(u^{-1}(t))} \right) dt \tag{48}$$

for $y \in L^2(0,1)$ and $\zeta < s < \Xi$.

Let us focus on some reinsurer conditions. First, in order to prevent her/his moral hazard, the reinsurer will select a real number $c \geq 0$, $c < 1$, indicating the minimum allowed rate of retained risk. If the reinsurer accepts stop-loss contracts or other ones whose rate may vanish then $c = 0$ may hold, and the classical approach is included in this analysis. Otherwise, the reinsurer will choose $c > 0$. The reinsurer will also select the premium principle Π. As said above, this section only has illustrative purposes, so we will only deal with the Expected Value Premium Principle. Consequently, $\Pi(u_c) = (1+\beta)\,\mathbb{E}(u_c)$, $\beta > 0$ being the loading rate.

If $x \in L^\infty(\zeta, \Xi)$ is the insurer decision variable and $J(x) \in L^2(0,1)$ is the retained risk, then the ceded risk will be (see (45), and recall that J is linear)

$$u_c(t) = u(t) - J(x)(t) = \zeta + J(1)(t) - J(x)(t) = \zeta + J(1-x)(t).$$

If A is the premium paid by the insurer clients, the insurer random wealth at T will be

$$\begin{aligned} W &= A - J(x) - \Pi(u_c) = A - J(x) - (1+\beta)\,\mathbb{E}(u_c) \\ &= A - J(x) - (1+\beta)\zeta - (1+\beta)\,\mathbb{E}(J(1-x)). \end{aligned}$$

Let us assume that the insurer objective is the simultaneous maximization of the expected wealth $\mathbb{E}(W)$ and minimization of the risk $\rho(W)$, where $\rho = CVaR_{1-\alpha}$ or $\rho = VaR_{1-\alpha}$ for some $0 < \alpha < 1$. Then the insurer will look for Pareto solutions, and, according to Nakayama *et al* (1985), most of them can be found by minimizing linear combinations $-\mathbb{E}(W) + U\rho(W)$, $\frac{U}{1+U} > 0$ being the "relative weight" or "relative importance" of the risk $\rho(W)$ with respect to the expected wealth $\mathbb{E}(W)$. Bearing in mind that both J and $\mathbb{E}$ are linear and both $-\mathbb{E}$ and ρ are translation invariant and positively homogeneous (Rockafellar and Uryasev, 2000), the minimization of $-\mathbb{E}(W) + U\rho(W)$ is equivalent to the minimization of $\rho(-J(Vx)) - \mathbb{E}(J(x))$ with $V = \frac{U}{\beta + U + \beta U} \in (0,1)$. Hence, the insurer problem becomes

$$Min\ \left\{ \rho(-J(Vx)) - \mathbb{E}(J(x))\,;\ c \leq x \leq 1 \right\}.^{13} \tag{49}$$

[13] Constraint $x \leq 1$ must hold because u, u_r and u_c must be co-monotone.

Proposition 15 *Consider* $-\infty < \mu_0 < 0$ *and* $x \in L^\infty(\zeta, \Xi)$ *with* $c \leq x \leq 1$. *Denote by* $\mathcal{X}_{(1-e^{\mu_0},1)}$ *the characteristic function of* $(1-e^{\mu_0}, 1) \subset (0,1)$.

a) $CVaR_{1-e^{\mu_0}}(-J(x)) = e^{-\mu_0}\mathbb{E}\left(\mathcal{X}_{(1-e^{\mu_0},1)}J(x)\right)$.

b) $(-J(x), p=2, \mu_0)$ *is regular (see Definition 3).*

c) If $J_{\mu_0}(x) \in \mathbb{R}$ *denotes the value of* $J(x)(t)$ *in (44) for* $t = 1-e^{\mu_0}$, *then* $VaR_{1-e^{\mu_0}}(-J(x)) = J_{\mu_0}(x)$.

Proof. *a)* Expressions (4) and (5) trivially imply that $CVaR_{1-e^{\mu_0}}(-u) = e^{-\mu_0}\mathbb{E}\left(u\mathcal{X}_{(1-e^{\mu_0},1)}\right)$ because $u:(0,1) \rightarrow (\zeta, \Xi)$ is a strictly increasing bijection. Since $J(x)$ and u are co-monotone we have that $J(x):(0,1) \rightarrow (\zeta, \Xi)$ is non decreasing, and therefore (4) again leads to $CVaR_{1-e^{\mu_0}}(-J(x)) = e^{-\mu_0}\mathbb{E}\left(\mathcal{X}_{(1-e^{\mu_0},1)}J(x)\right)$.

b) We have to prove that $(-\infty, 0) \ni \mu \rightarrow \mathbb{E}(\lambda_\mu) \in \mathbb{R}$ is continuous at μ_0 if λ_M is the multiplier of (15) related to Constraint $w \leq e^{-\mu}$ and $y = -J(x)$. (16) indicates that the multipliers of (15) are the solutions of

$$\left\{\begin{array}{l} \mathbb{E}(w) = 1 \\ -J(x) = \lambda_m - \lambda_M - \lambda \\ w\lambda_m = \left(e^{-\mu} - w\right)\lambda_M = 0 \\ 0 \leq w \leq e^{-\mu},\ 0 \leq \lambda_m,\ 0 \leq \lambda_M \end{array}\right. \tag{50}$$

it may be easily verified that

$$\left\{\begin{array}{l} w = e^{-\mu}\mathcal{X}_{(1-e^{-\mu},1)} \\ \\ \lambda = J_\mu(x) \\ \\ \lambda_m = \left\{\begin{array}{ll} -J(x) + J_\mu(x), & (0, 1-e^{\mu}) \\ 0 & (1-e^{\mu}, 1) \end{array}\right\} \\ \\ \lambda_M = \left\{\begin{array}{ll} 0, & (0, 1-e^{\mu}) \\ J(x) - J_\mu(x), & (1-e^{\mu}, 1) \end{array}\right\} \end{array}\right. \tag{51}$$

solves the system above if $J_\mu(x) \in \mathbb{R}$ denotes the value of $J(x)(t)$ in (44) for $t = 1-e^{\mu}$. Hence, we only have to see that $(-\infty, 0) \ni \mu \rightarrow \int_{1-e^{\mu}}^{1} J(x)\,dt - e^{\mu}J_\mu(x) \in \mathbb{R}$ is continuous at μ_0, which is an obvious property.

c) The regularity of $(-J(x), 2, \mu_0)$, Lemma 4 and (18) imply that

$$-e^{-\mu_0}\mathbb{E}(\lambda_M) = \frac{dCVaR_{1-e^{\mu}}(-J(x))}{d\mu}\Big|_{\mu=\mu_0} \tag{52}$$

holds if $(w, \lambda, \lambda_m, \lambda_M) \in L^\infty(\zeta, \Xi) \times L^\infty(0,1) \times \mathbb{R} \times L^2(0,1) \times L^2(0,1)$ satisfy (50) and $\mu = \mu_0$. Then, (51) with $\mu = \mu_0$, and (52) lead to

$$\frac{dCVaR_{1-e^{\mu_0}}(-J(x))}{d\mu}\Big|_{\mu=\mu_0} = -e^{-\mu_0}\mathbb{E}\left(\mathcal{X}_{(1-e^{\mu_0},1)}\left(J(x) - J_{\mu_0}(x)\right)\right). \tag{53}$$

Hence, Statement a) and Lemma 1 imply that

$$\begin{array}{l} VaR_{1-e^{\mu_0}}\left(-J\left(x\right)\right)= \\ e^{-\mu_0}\mathbb{E}\left(\mathcal{X}_{\left(1-e^{\mu_0},1\right)}J\left(x\right)\right)-e^{-\mu_0}\mathbb{E}\left(\mathcal{X}_{\left(1-e^{\mu_0},1\right)}\left(J\left(x\right)-J_{\mu_0}\left(x\right)\right)\right), \end{array}$$

and c) holds. □

Remark 16 *(Optimal reinsurance if $\rho = CVaR_{1-e^{\mu_0}}$) Let us suppose that $\rho = CVaR_{1-e^{\mu_0}}$. Proposition* 15$a$ *implies that the objective function of (49) equals*

$$Ve^{-\mu_0}\mathbb{E}\left(\mathcal{X}_{\left(1-e^{\mu_0},1\right)}J\left(x\right)\right)-\mathbb{E}\left(J\left(x\right)\right)=\mathbb{E}\left(\left(Ve^{-\mu_0}\mathcal{X}_{\left(1-e^{\mu_0},1\right)}-1\right)J\left(x\right)\right).$$

Consequently, (46) leads to the objective function

$$\left\langle x,J^*\left(Ve^{-\mu_0}\mathcal{X}_{\left(1-e^{\mu_0},1\right)}-1\right)\right\rangle = Ve^{-\mu_0}\left\langle x,J^*\left(\mathcal{X}_{\left(1-e^{\mu_0},1\right)}\right)\right\rangle-\left\langle x,J^*\left(1\right)\right\rangle. \quad (54)$$

(48) trivially implies that

$$J^*\left(1\right)\left(s\right)=1-u^{-1}\left(s\right) \quad (55)$$

and

$$J^*\left(\mathcal{X}_{\left(1-e^{\mu_0},1\right)}\right)\left(s\right)=\begin{cases} e^{\mu_0}, & s\leq u\left(1-e^{\mu_0}\right) \\ 1-u^{-1}\left(s\right), & s\geq u\left(1-e^{\mu_0}\right)\end{cases} \quad (56)$$

for $\zeta < s < \Xi$. Hence, the objective function in (54) becomes

$$\begin{array}{l} \rho\left(-J\left(Vx\right)\right)-\mathbb{E}\left(J\left(x\right)\right)=V\int_\zeta^{u\left(1-e^{\mu_0}\right)}x\left(s\right)ds+ \\ Ve^{-\mu_0}\int_{u\left(1-e^{\mu_0}\right)}^{\Xi}x\left(s\right)\left(1-u^{-1}\left(s\right)\right)-\int_\zeta^\Xi x\left(s\right)\left(1-u^{-1}\left(s\right)\right)ds \\ =\int_\zeta^\Xi K\left(s\right)x\left(s\right)ds \end{array}$$

with

$$K\left(s\right)=\begin{cases} u^{-1}\left(s\right)+V-1, & s<u\left(1-e^{\mu_0}\right) \\ \left(1-u^{-1}\left(s\right)\right)\left(Ve^{-\mu_0}-1\right), & s>u\left(1-e^{\mu_0}\right)\end{cases} \quad (57)$$

for $\zeta < s < \Xi$.[14] *Hence, Problem (49) becomes*

$$Min\left\{\int_\zeta^\Xi K\left(s\right)x\left(s\right)ds;\ c\leq x\leq 1\right\}. \quad (58)$$

Since $c \geq 0$, the obvious solution of (58) (and the solution of (49)) will be

$$X^*=\begin{cases} 1, & K\left(s\right)\leq 0 \\ c, & K\left(s\right)>0\end{cases} \quad (59)$$

for $\zeta < s < \Xi$.

[14] Notice that the imposed assumptions imply that $1-u^{-1}\in L^1\left(\zeta,\Xi\right)$.

Proposition 15*a will allow us to compute easily the retained optimal CVaR given by* $e^{-\mu_0}\mathbb{E}\left(\mathcal{X}_{(1-e^{\mu_0},1)}J\left(X^*\right)\right)$*. Proceeding as in the proof of Propositions* 15*b and* 15*c we have that (51) and (53) for* $\mu=\mu_0$ *yield the marginal CVaR of the retained claims* $-J\left(X^*\right)$ *with respect to* μ *at* μ_0*. Moreover, Proposition* 15*c will provide us with the retained VaR, i.e.,* $VaR_{1-e^{\mu_0}}\left(-J\left(X^*\right)\right)$*. Lastly, if the required conditions hold, the second order differential equation (39) will yield the marginal VaR of the retained claims* $-J\left(X^*\right)$ *with respect to* μ *at* μ_0*, and the methodology proposed in Section* 6.3 *will yield the evolution of both* $CVaR_{1-e^{\mu}}\left(-J\left(X^*\right)\right)$ *and* $VaR_{1-e^{\mu}}\left(-J\left(X^*\right)\right)$ *as* μ *is modified. An illustrative numerical example will be shown in Remark* 19. $\square$

Remark 17 *(Optimal reinsurance if* $\rho=VaR_{1-e^{\mu_0}}$*) Proposition* 15*c implies that for* $\rho=VaR_{1-e^{\mu_0}}$ *the objective function of (49) equals*

$$VJ_{\mu_0}(x)-\mathbb{E}\left(J(x)\right)=VJ_{\mu_0}(x)-\langle J(x),1\rangle=VJ_{\mu_0}(x)-\langle x,J^*(1)\rangle\,.$$

Therefore, (44) and (55) lead to the objective function

$$V\int_{\zeta}^{u(1-e^{\mu_0})}x(s)\,ds-\int_{\zeta}^{\Xi}x(s)\left(1-u^{-1}(s)\right)ds=\int_{\zeta}^{\Xi}k(s)\,x(s)\,ds$$

with

$$k(s)=\begin{cases}u^{-1}(s)+V-1, & s<u\left(1-e^{\mu_0}\right)\\ u^{-1}(s)-1, & s>u\left(1-e^{\mu_0}\right)\end{cases}\tag{60}$$

for $\zeta<s<\Xi$*. The solution of (49) for* $\rho=VaR_{1-e^{\mu_0}}$ *becomes*

$$x^*=\begin{cases}1, & k(s)\le 0\\ c, & k(s)>0\end{cases}\tag{61}$$

for $\zeta<s<\Xi$.

As in Remark 16*, Propositions* 15*a and* 15*c give the VaR and CVaR of the retained claims* $-J\left(x^*\right)$*, and Differential Equations (12) and (39) will provide us with marginal values and potential evolutions as the level of confidence evolves too.* $\square$

Remark 18 *Suppose that* $\rho=VaR_{1-e^{\mu_0}}$*. Proposition* 15 *and Remark* 17 *illustrate that Theorem* 10 *is not the unique way to apply Lemmas* 1 *and* 4 *(or Theorem* 8*) in VaR optimization. As shown in Remark* 17*, Proposition* 15 *allows us to solve (49) without using the equivalent differentiable optimization problem (32). Nevertheless, needless to say that Theorem* 10 *is general and also applies to solve (49). Though we will not address this question in order to shorten the exposition, at least we will indicate that the arguments of Theorem* 10 *will imply the equivalence between the non-differentiable problem (49) and the differentiable one*

$$\begin{cases}Min & V\lambda-\mathbb{E}\left(J(x)\right)\\ & -J(x)=\lambda_m-\lambda_M-\lambda\\ & \lambda\in\mathbb{R},\ c\le x\le 1,\ \left(w,\lambda_m,\lambda_M\right)\in\delta_{e^{\mu}}\end{cases}$$

$(x, w, \lambda, \lambda_m, \lambda_M) \in L^\infty(\zeta, \Xi) \times L^\infty(0,1) \times \mathbb{R} \times L^2(0,1) \times L^2(0,1)$ *being the decision variable.* $\square$

Remark 19 *(Numerical experiment) Let us deal with a simple and illustrative numerical example. Suppose that* $\zeta = 0$, $\Xi = 100$ *and* $u(t) = 100t^2$ *for* $0 < t < 1$. *Obviously,*

$$u^{-1}(s) = \frac{\sqrt{s}}{10} \tag{62}$$

for $0 < s < 100$ *is the cumulative distribution function of the total risk* u, *and its derivative* $\left(u^{-1}\right)'(s) = \dfrac{1}{20\sqrt{s}}$ *for* $0 < s < 100$ *is the density function of* u. *Suppose that* $c = \dfrac{1}{2}$, *i.e., the reinsurer will never accept to pay more than* 50% *per claim. If* $-\infty < \mu < 0$, *the kernels (57) and (60) become*

$$0 < s \le u(1 - e^\mu) \Longrightarrow \quad \left\{ K(s) = k(s) = \frac{\sqrt{s}}{10} + V - 1 \right.$$

$$u(1 - e^\mu) < s < 100 \Longrightarrow \quad \left\{ \begin{array}{l} K(s) = \left(1 - \dfrac{\sqrt{s}}{10}\right)(Ve^{-\mu} - 1) \\ k(s) = -\left(1 - \dfrac{\sqrt{s}}{10}\right) \end{array} \right.$$

Thus, bearing in mind that $1 - \dfrac{\sqrt{s}}{10} > 0$ *for* $0 < s < 100$, *(59) and (61) lead to the optimal solutions*

$$\left\{ \begin{array}{l} 0 < s < u(1 - e^\mu) \\ \dfrac{\sqrt{s}}{10} + V < 1 \end{array} \right. \quad \Longrightarrow \quad X^* = x^* = 1$$

$$\left\{ \begin{array}{l} 0 < s < u(1 - e^\mu) \\ \dfrac{\sqrt{s}}{10} + V > 1 \end{array} \right. \quad \Longrightarrow \quad X^* = x^* = \frac{1}{2}$$

$$\left\{ \begin{array}{l} u(1 - e^\mu) < s < 100 \\ Ve^{-\mu} < 1 \end{array} \right. \quad \Longrightarrow \quad X^* = 1$$

$$\left\{ \begin{array}{l} u(1 - e^\mu) < s < 100 \\ Ve^{-\mu} > 1 \end{array} \right. \quad \Longrightarrow \quad X^* = \frac{1}{2}$$

$$u(1 - e^\mu) < s < 100 \Longrightarrow \quad x^* = 1$$

For instance, if $V = 1/2$ *and the level of confidence is* 80% *(i.e.,* $1 - e^\mu = 0.8 \Longrightarrow \mu = Log(0.2) \simeq -1.60944$*) then we get*

$$X^* = \left\{ \begin{array}{ll} 1/2, & 25 < s < 100 \\ 1, & 0 < s < 25 \end{array} \right. \quad \textit{and} \quad x^* = \left\{ \begin{array}{ll} 1/2, & 25 < s < 64 \\ 1, & \textit{otherwise} \end{array} \right. \tag{63}$$

and, bearing in mind that $\mathbb{E}(u) = 100/3 \simeq 33.33333$, *for both risk measures the insurer will retain all the risk up to a given threshold lower than the expected* *claim, and he/she will cede* 50% *of every claim beyond that threshold, though for* $\rho = VaR$ *he/she will recover the full rate beyond a second threshold almost equaling two times the global expected claim.*

Next let us focus on x^* *and illustrate the method proposed in Section* 6.3. *Consider the functions* h *and* H *of (10) for* $y = -J(x^*)$. *Proposition* 15 *(46) and (56) imply that for* $\mu_0 = Log(0.2)$ *we have*

$$H(\mu_0) = CVaR_{80\%}(-J(x^*)) = e^{-\mu_0}\mathbb{E}\left(\mathcal{X}_{(1-e^{\mu_0},1)}J(x^*)\right) = 5\mathbb{E}\left(\mathcal{X}_{(0.8,1)}J(x^*)\right)$$
$$= 5\left\langle x^*, J^*\left(\mathcal{X}_{(0.8,1)}\right)\right\rangle \simeq 61.83333,$$

and

$$h(\mu_0) = VaR_{80\%}(-J(x^*)) = J_{\mu_0}(x^*) = 44.5.$$

Lemma 1 *implies that*

$$H'(\mu_0) = 44.5 - 61.83333 \simeq -17.33333.$$

Bearing in mind (62) and (63), straightforward manipulations enable us to obtain

$$f(s) = \begin{cases} \dfrac{1}{20\sqrt{-s-19.5}}, & -80.5 < s < -44.5 \\ \dfrac{\sqrt{2}}{40\sqrt{25-s}}, & -44.5 < s < -25 \\ \dfrac{1}{20\sqrt{-s}}, & -25 < s < 0 \\ 0 & otherwise \end{cases}$$

as the density function of $-J(x^*)$. *Thus bearing in mind (39) and (40), we have that the vector* $(CVaR, M\arg inal\_CVaR, VaR, M\arg inal\_VaR)$, *represented by*

$$\vec{H} = (H_1 = H, H_2 = H', H_3 = h, H_4 = h')$$

satisfy the initial conditions

$$H_1(\mu_0) = 61.83333$$
$$H_2(\mu_0) = -17.33333$$

along with the system

$$H_1' = H_2$$
$$H_2' = -H_2 - \frac{e^{\mu}}{f(-H_1 - H_2)}$$
$$H_3 = H_1 + H_2$$
$$H_4 = -\frac{e^{\mu}}{f(-H_1 - H_2)}$$

Solving this system with the Eüler method and the step $\delta = -10^{-6}$ for the increment of $\alpha = e^{\mu}$, we have obtained the approximated evolution of $\vec{H}$ as a function of α. A brief summary of the results is reported in the matrix below

$$\begin{pmatrix} \alpha = e^{\mu}, & H_1, & H_2, & H_3, & H_4 \\ 0.2 & 61.83333 & -17.33333 & 44.5 & -32 \\ 0.15 & 66.25 & -13.5 & 52.75 & -25.5 \\ 0.1 & 70.83333 & -9.333333 & 61.5 & -18 \\ 0.05 & 75.58333 & -4.833333 & 70.75 & -9.5 \\ 0.03 & 77.53 & -2.94 & 74.59 & -5.82 \\ 0.02 & 78.51333 & -1.973333 & 76.54 & -3.92 \\ 0.01 & 79.50333 & -0.993333 & 78.51 & -1.98 \end{pmatrix}$$

In particular, though the optimal reinsurance x^* shows a very stable $CVaR$, the VaR stability seems to be lower, and this information could be interesting to the insurer before making the reinsurance purchase decision. This is a simple but illustrative application of Section 6.3. □

8 Conclusion

It has been shown that VaR and $CVaR$ are related by a first order linear differential equation with constant coefficients. This differential equation holds when the independent variable is the level of confidence of both risk measures. Besides, since the $CVaR$ Representation Theorem establishes that the $CVaR$ is the optimal value of a linear optimization problem, standard results in Mathematical Programming guarantee that the first derivative of the $CVaR$ with respect to the level of confidence is given by the Lagrange multiplier associated with the $CVaR$ representation. A new representation of VaR is proved by combining both the differential equation and the Lagrange multiplier interpretation; VaR equals $CVaR$ plus the cited Lagrange multiplier.

There are many analytical properties holding for $CVaR$ and failing for VaR, which implies that there are many interesting consequences of the VaR representation above, because several problems involving VaR may be solved with the help of the $CVaR$ properties. Some of these consequences have been pointed out in this paper. In particular, optimization problems involving VaR and/or probabilistic constraints always have an equivalent differentiable optimization problem, and this allows us to provide $VaR-$linked (or probabilistic constraints-linked) optimization problems with Karush-Kuhn-Tucker-like and Fritz John-like necessary optimality conditions, despite the fact that VaR is neither continuous nor differentiable. Furthermore, under appropriate (but general) assumptions, a second order differential equation also relates VaR and $CVaR$, and this enables us to establish new methods in risk and marginal risk estimates. Practical illustrative examples have been presented.

References

Assa, H. and K.M. Karai, 2013. Hedging, Pareto optimality and good deals. *Journal of Optimization Theory and Applications*, 157, 900-917.

Balbás, A., B. Balbás and R. Balbás, 2010. Minimizing measures of risk by saddle point conditions. *Journal of Computational and Applied Mathematics*, 234, 2924 - 2931.

Balbás, A., B. Balbás and R. Balbás, 2016. Outperforming benchmarks with their derivatives: Theory and empirical evidence. *The Journal of Risk*, 18, 4, 25 - 52.

Balbás, A., B. Balbás and R. Balbás, 2017. VaR as the $CVaR$ sensitivity: Applications in risk optimization. *Journal of Computational and Applied Mathematics*, 309, 175 - 185.

Balbás, A., B. Balbás, R. Balbás and A. Heras, 2015. Optimal reinsurance under risk and uncertainty. *Insurance: Mathematics and Economics*, 60, 61 - 74.

Butcher, J.C., 2008. *Numerical methods for ordinary differential equations*. John Wiley & Sons.

Cai, J., C. Lemieux and F. Liu, 2016. Optimal reinsurance from the perspectives of both an insurer and a reinsurer. *ASTIN Bulletin*, 43, 815 - 849.

Cai, J. and K.S. Tan, 2007. Optimal retention for a stop loss reinsurance under the VaR and CTE risk measures. *ASTIN Bulletin*, 37, 1, 93-112.

Centeno, M.L. and O. Simoes, 2009. Optimal reinsurance. *Revista de la Real Academia de Ciencias, RACSAM*, 103, 2, 387-405.

Chavez-Demoulin, V., P. Embrechts and J. Nešlehová, 2006. Quantitative models for operational risk: Extremes, dependence and aggregation. *Journal of Banking & Finance,* 30, 2635–2658.

Coddington, E. A. and N. Levinson, 1955. *Theory of ordinary differential equations*. McGraw-Hill.

Dhaene, J., R.J. Laeven, S. Vanduffel, G. Darkiewicz and M.J. Goovaerts, 2008. Can a coherent risk measure be too subadditive? *Journal of Risk and Insurance*, 75, 365 - 386.

Gaivoronski, A. and G. Pflug, 2005. Value at risk in portfolio optimization: Properties and computational approach. *The Journal of Risk*, 7,2, 1 - 31.

Goovaerts, M.J., R. Kaas, J. Dhaene and Q. Tang, 2004. A new classes of consistent risk measures. *Insurance: Mathematics and Economics*, 34, 505-516.

Goovaerts, M.J. and R. Laeven, 2008. Actuarial risk measures for financial derivative pricing. *Insurance: Mathematics and Economics*, 42, 540 - 547.

Kelly, J.L., 1955. *General topology*. Springer.

Kopp, P.E., 1984. *Martingales and stochastic integrals*. Cambridge University Press.

Kupper, M. and G. Svindland, 2011. Dual representation of monotone convex functions on L^0. *Proceedings of the American Mathematical Society*, 139, 4073–4086.

Larsen, N., H. Mausser and S. Uryasev, 2002. *Algorithms for optimization of value-at-risk*. In: Pardalos P, Tsitsiringos V (eds) Financial Engineering, e-Commerce and Supply Chain, Kluwer Academic Publishers, Dordrecht, Netherlands, 129–157.

Lejeune, M., 2012. Pattern-based modeling and solution of probabilistically constrained optimization problems. *Operations Research*, 60 (6), 1356–1372 .

Lejeune, M. and S. Shen, 2016. Multi-objective probabilistically constrained programs with variable risk: Models for multi-portfolio financial optimization. *European Journal of Operational Research*, 252, 522 - 539.

Luenberger, D.G., 1969. *Optimization by vector spaces methods*. John Wiley & Sons.

Mangasarian, O.L. and S. Fromovitz, 1967. The Fritz John necessary optimality conditions in the presence of equality and inequality constraints. *Journal of Mathematical Analysis and Applications*, 17, 37 - 47.

Mohammadi, F., C. Arévalo and C. Führer, 2017. Runge–Kutta restarters for multistep methods in presence of frequent discontinuities. *Journal of Computational and Applied Mathematics*, 316, 287 - 297.

Nakayama, H., Y. Sawaragi and T. Tanino, 1985. *Theory of multiobjective optimization*. Academic Press.

Peng, X. and W. Wang, 2016. Optimal investment and risk control for an insurer under inside information. *Insurance: Mathematics and Economics*, 69, 104-116.

Rockafellar, R.T. and S. Uryasev, 2000. Optimization of conditional-value-at-risk. *The Journal of Risk*, 2, 21 – 42.

Rockafellar, R.T., S. Uryasev and M. Zabarankin, 2006. Generalized deviations in risk analysis. *Finance & Stochastics*, 10, 51-74.

Rudin, W., 1973. *Functional analysis*. McGraw-Hill.

Stoyanov, S.V., S.T. Rachev and F.J. Fabozzi, 2007. Optimal financial portfolios. *Applied Mathematical Finance*, 14, 401-436.

Wozabal, D., 2012. Value-at-Risk optimization using the difference of convex algorithm. *OR Spectrum*, 34, 861-883.

Zakamouline, V. and S. Koekebbaker, 2009. Portfolio performance evaluation with generalized Sharpe ratios: Beyond the mean and variance. *Journal of Banking & Finance*, 33, 1242-1254.

Zeidler, E., 1995. *Applied functional analysis: Main principles and their applications*. Springer.

Zhao, P. and Q. Xiao, 2016. Portfolio selection problem with Value-at-Risk constraints under non-extensive statistical mechanics. *Journal of Computational and Applied Mathematics*, 298, 74-91.