J Optim Theory Appl (2012) 155:1124–1128
DOI 10.1007/s10957-012-0014-9

Addendum to: Entropic Value-at-Risk: A New Coherent Risk Measure

A. Ahmadi-Javid

Received: 24 January 2012 / Accepted: 15 February 2012 / Published online: 16 March 2012

Abstract This short addendum consists of two sections. The first provides proofs that were omitted in Ahmadi-Javid (J. Optim. Theory Appl., 2012) for the sake of brevity, and also demonstrates that the dual representation of the entropic value-at-risk, which is given in Ahmadi-Javid (J. Optim. Theory Appl., 2012) for the case of bounded random variables, holds for all random variables whose moment-generating functions exist everywhere. The second section provides a few corrections.

Keywords Chernoff inequality · Coherent risk measure · Conditional value-at-risk (CVaR) · Convex optimization · Cumulant-generating function · Duality · Entropic value-at-risk (EVaR) · g-entropic risk measure · Moment-generating function · Relative entropy · Stochastic optimization · Stochastic programming · Value-at-risk (VaR)

1 Supplementary Proofs

In this section, we begin by providing detailed proofs for some of the statements made in [1], which may be helpful to readers who are less familiar with convex optimization. Then we discuss the dual representation of the entropic value-at-risk (EVaR) for any random variable whose moment-generating function exists everywhere.

The following lemma proves the convexity of cumulant-generating functions, which is used in the proof of Lemma 3.1 of [1].

Lemma 1.1 *The cumulant-generating function* $\ln M_X(t)$ *is convex in* X.

A. Ahmadi-Javid (✉)
Department of Industrial Engineering, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran
e-mail: ahmadi_javid@aut.ac.ir

Proof Without any loss of generality, we can set $t = 1$. It now suffices to show that, for any two random variables X and Y with finite $M_X(1)$, $M_Y(1)$ and any $\lambda \in [0,1]$,

$$E\big(e^{\lambda X+(1-\lambda)Y}\big) \le E\big(e^X\big)^\lambda E\big(e^Y\big)^{1-\lambda}.$$

Defining $W := e^X/E(e^X)$ and $V := e^Y/E(e^Y)$, this inequality is equivalent to $E(W^\lambda V^{1-\lambda}) \le 1$ that immediately follows from the inequality of weighted arithmetic and geometric means for the non-negative random variables W and V, i.e., $W^\lambda V^{1-\lambda} \le \lambda W + (1-\lambda)V$. □

The following lemma proves a statement used in the proof of Theorem 3.1 of [1].

Lemma 1.2 *The function* $\inf_{t>0}\{\kappa_\alpha(X,t)\}$ *is convex in* X *for all* $\alpha \in]0,1]$.

Proof For all $\varepsilon > 0$, and $X, Y \in \mathbf{L}_{M^+}$, the continuity of $\kappa_\alpha(X,t)$ in $t > 0$ implies the existence of $t_1, t_2 > 0$ such that

$$\kappa_\alpha(X,t_1) \le \inf_{t>0}\big\{\kappa_\alpha(X,t)\big\} + \varepsilon, \qquad \kappa_\alpha(Y,t_2) \le \inf_{t>0}\big\{\kappa_\alpha(Y,t)\big\} + \varepsilon.$$

Since $\kappa_\alpha(X,t)$ is convex in (X,t) from Lemma 3.1 of [1], we further find that, for all $\lambda \in [0,1]$,

$$\begin{aligned}&\inf_{t>0}\big\{\kappa_\alpha\big(\lambda X + (1-\lambda)Y, t\big)\big\} \le \kappa_\alpha\big(\lambda X + (1-\lambda)Y, \lambda t_1 + (1-\lambda)t_2\big)\\ &\quad\le \lambda\kappa_\alpha(X,t_1) + (1-\lambda)\kappa_\alpha(Y,t_2) \le \lambda \inf_{t>0}\big\{\kappa_\alpha(X,t)\big\} + (1-\lambda)\inf_{t>0}\big\{\kappa_\alpha(Y,t)\big\} + \varepsilon.\end{aligned}$$

As this holds for all $\varepsilon > 0$, the proof is done by taking the limit $\varepsilon \downarrow 0$. □

The following lemma proves the validity of the last identities used in the proofs of Theorems 3.3 and 5.1 of [1].

Lemma 1.3 *Let* $g : \mathbb{R} \to [0,\infty]$ *be a non-negative closed convex function with* $g(1) = 0$ *and* $\operatorname{dom} g = [\gamma_1, \gamma_2]$ *where* $\gamma_1 < 1 < \gamma_2$ *and* $\gamma_1, \gamma_2 \in [-\infty, \infty]$. *Then the following identity holds for any* $\beta > 0$

$$\inf_{t>0}\Big\{\sup_{Q\ll P}\big\{\mathrm{E}_Q(X) + t\big(\beta - H_g(P,Q)\big)\big\}\Big\} = \sup_{Q\ll P, H_g(P,Q)\le\beta}\big\{\mathrm{E}_Q(X)\big\}.$$

Proof By denoting $Y = \frac{dQ}{dP}$, which is a non-negative random variable with the mean equal to 1, the above identity can be rewritten as

$$\sup_{t\ge0}\big\{L(t)\big\} = \inf_{Y\in S, \mathrm{E}_P(g(Y))\le\beta}\big\{-\mathrm{E}_P(XY)\big\},$$

where $L(t) = \inf_{Y\in S}\{-\mathrm{E}_Q(XY) + t(\mathrm{E}_P(g(Y)) - \beta)\}$ is the Lagrangian associated with the optimization problem on the right-hand side, and $S = \{Y \in \mathbf{L}_1 : \mathrm{E}_P(Y) = 1, \max\{\gamma_1, 0\} \le Y \le \gamma_2 \ a.e.\}$. Hence, it suffices to show that the optimal duality gap for the right-hand side optimization problem is zero. This is possible by showing that the generalized Slater's constraint qualification [2] holds for this problem, i.e., that there exists $\hat{Y} \in \mathbf{L}_1$ satisfying $\mathrm{E}_P(\hat{Y}) = 1$, $\max\{\gamma_1, 0\} < \hat{Y} < \gamma_2$ *a.e.*, $\mathrm{E}_P(g(\hat{Y})) < \beta$.

As we assumed $\gamma_1 < 1 < \gamma_2$, the solution $\widehat{Y} = 1$ *a.e.* fulfills these conditions, and so the proof is complete. □

Remark 1.1 We can also show the validity of the identity in Lemma 1.3 for the case where g is not always non-negative over its domain. If g is not non-negative, then, assuming that $c(x-1)$ is a supporting hyperplane to the epigraph of g at the point $(1,0)$, it is sufficient to replace g with the non-negative function $\tilde{g}(x) = g(x) - c(x-1)$, for which we have $H_{\tilde{g}}(P,Q) = H_g(P,Q)$. Moreover, if $\gamma_1 = 1$ or $\gamma_2 = 1$, then the validity of the identity in Lemma 1.3 is clear, because the constraint $\mathrm{E}_P(\frac{dQ}{dP}) = 1$ implies $\frac{dQ}{dP} = 1$ *a.e.* or, equivalently, $Q = P$. Finally, note that the proof for the case $\beta = 0$ is also straightforward, given that $\gamma_1 < 1 < \gamma_2$ and g is non-negative. Indeed, for this case, the constraint $H_g(P,Q) \le 0$ is feasible if and only if g is zero over the interval $\operatorname{dom} g = [\gamma_1, \gamma_2]$, which implies that the constraint $H_g(P,Q) \le 0$ is redundant since it is equivalent to $\gamma_1 \le \frac{dQ}{dP} \le \gamma_2$ *a.e.* which already exists in the set S.

Here we give the detailed proof of Proposition 4.2 of [1].

Proof of Proposition 4.2 of [1] To use Theorem 4.1 of [1], we first need to reformulate problem (8) of [1] by using the CVaR representation given in (2) of [1] and adding the additional constraint $lC \le -t \le uC$. The resulting problem can be rewritten in the form of problem (6) of [1] as follows:

$$\min_{\boldsymbol{w} \in \mathbf{W},\, lC \le -t \le uC} \left\{ t + \alpha^{-1} \mathrm{E}\left[-\sum_{i=1}^{n} w_i R_i - t \right]_+ \right\} = \min_{\boldsymbol{x} \in \mathbf{X}} \mathrm{E}\big(F(\boldsymbol{x}, \boldsymbol{\xi})\big),$$

where $\boldsymbol{x} = (\boldsymbol{w}^T, t)^T$, $\boldsymbol{\xi} = \boldsymbol{R}$, $\mathbf{X} = \mathbf{W} \times [-uC, -lC]$ and $F(\boldsymbol{x}, \boldsymbol{\xi}) = t + \alpha^{-1}[-\sum_{i=1}^{n} w_i R_i - t]_+$. Then we need to find D and L. In this case, $D = \sup_{\boldsymbol{x}, \boldsymbol{x}' \in \mathbf{X}} \|\boldsymbol{x} - \boldsymbol{x}'\| = C\sqrt{2 + B^2}$, where the maximum value is attained for $\boldsymbol{x} = (C, 0, \ldots, 0, -uC)^T$ and $\boldsymbol{x}' = (0, C, \ldots, 0, -lC)^T$, or other similar pairs of points. To determine the Lipschitz constant L, we have

$$\begin{aligned}
\big|F(\boldsymbol{x}, \mathbf{z}) - F(\boldsymbol{x}', \mathbf{z})\big| &\le |t - t'| + \alpha^{-1} \big| [-\mathbf{z}^T \boldsymbol{w} - t]_+ - [-\mathbf{z}^T \boldsymbol{w}' - t']_+ \big| \\
&\le |t - t'| + \alpha^{-1} \big| [-\mathbf{z}^T (\boldsymbol{w} - \boldsymbol{w}') - (t - t')]_+ \big| \\
&\le |t - t'| + \alpha^{-1} \big| \mathbf{z}^T (\boldsymbol{w} - \boldsymbol{w}') + (t - t') \big| \\
&= |t - t'| + \alpha^{-1} \big| (\mathbf{z}^T, 1)(\boldsymbol{x} - \boldsymbol{x}') \big| \le \big(1 + \alpha^{-1} \|(\mathbf{z}^T, 1)^T\|\big) \|\boldsymbol{x} - \boldsymbol{x}'\| \\
&\le \big(1 + \alpha^{-1} \sqrt{n \max\{u^2, l^2\} + 1}\big) \|\boldsymbol{x} - \boldsymbol{x}'\|,
\end{aligned}$$

where, in the last two inequalities, we used the Cauchy–Schwarz inequality and the fact that $\mathbf{z} \in S_{\boldsymbol{\xi}} = S_{\boldsymbol{R}} = [l, u]^n$. This shows that $L = 1 + \alpha^{-1}\sqrt{n \max\{u^2, l^2\} + 1}$. □

The next theorem proves that the dual representation in Theorem 3.3 of [1], which is given for bounded random variables, also holds for $X \in \mathbf{L}_M$.

Theorem 1.1 (Dual representation of the EVaR) *For $X \in \mathbf{L}_M$ and any $\alpha \in]0,1]$*

$$\mathrm{EVaR}_{1-\alpha}(X) = \sup_{Q\in\Im} \mathrm{E}_Q(X),$$

where $\Im = \{Q \ll P : D_{KL}(Q \parallel P) \le -\ln\alpha\}$.

Proof By virtue of the result given in Sect. 5.4 of [3], for any $X \in \mathbf{L}_M$, we have

$$\ln \mathrm{E}_P\big(e^X\big) = \sup_{Q\in\Im'} \big\{\mathrm{E}_Q(X) - D_{KL}(Q \parallel P)\big\},$$

where

$$\Im' = \left\{Q \ll P : \exists c > 0 : \mathrm{E}_P\left(h\left(c\frac{dQ}{dP}\right)\right) < \infty\right\}$$

$$\text{with} \quad h(x) = \begin{cases} 0, & 0 \le x < 1, \\ x\ln x - x + 1, & 1 \le x. \end{cases}$$

Hence, similarly as in the proofs of Lemma 1.3 and Theorem 3.3 of [1], and together with Remark 1.1, we can show

$$\mathrm{EVaR}_{1-\alpha}(X) = \sup_{Q\in\Im''} \mathrm{E}_Q(X),$$

where $\Im'' = \{Q \ll P : D_{KL}(Q \parallel P) \le -\ln\alpha, \exists c > 0 : \mathrm{E}_P(h(c\frac{dQ}{dP})) < \infty\}$. However, one can see that the constraint $D_{KL}(Q \parallel P) \le -\ln\alpha$ implies $\mathrm{E}_P(h(c\frac{dQ}{dP})) < \infty$ with $c = 1$, because the relative entropy can be rewritten as $D_{KL}(Q \parallel P) = \mathrm{E}_P(e(\frac{dQ}{dP}))$ with $e(x) = x\ln x - x + 1$, $x > 0$. Hence, the two sets $\Im$ and $\Im''$ are actually identical. This completes the proof. □

2 Corrections

This section corrects a few errors found in [1]. Since the numbering system of [1] was changed for publication, there are now two lemmas numbered 3.1: one precedes Theorem 3.1 of [1] and the other precedes Theorem 3.3 of [1]. The former of these lemmas was only used in the proof of Theorem 3.1 of [1], but the latter was used both in the proof of Theorem 3.3 of [1], and in the statement of Lemma 5.1 of [1]. Furthermore, at the end of the proof of the former, there are a few unnecessary sections that should be removed. Here is the corrected proof.

Proof of Lemma 3.1 of [1] We must show that, for all $\lambda \in [0,1]$, $X, Y \in \mathbf{L}_{M^+}$ and $t_1, t_2 > 0$,

$$\lambda\kappa_\alpha(X, t_1) + (1-\lambda)\kappa_\alpha(Y, t_2) \ge \kappa_\alpha\big(\lambda X + (1-\lambda)Y, \lambda t_1 + (1-\lambda)t_2\big),$$

which is equivalent to

$$\begin{aligned}&\lambda t_1 \ln M_X\big(t_1^{-1}\big) + (1-\lambda)t_2 \ln M_Y\big(t_2^{-1}\big) \\ &\quad\ge \big(\lambda t_1 + (1-\lambda)t_2\big) \ln M_{\lambda X+(1-\lambda)Y}\big(\big(\lambda t_1 + (1-\lambda)t_2\big)^{-1}\big).\end{aligned}$$

Denoting $t = \lambda t_1 + (1-\lambda)t_2$ and $w = \lambda t_1/t$, the left-hand side of the above inequality can be expressed as

$$t\big(w \ln M_X\big(t_1^{-1}\big) + (1-w)\ln M_Y\big(t_2^{-1}\big)\big).$$

Then by using the known fact that the cumulant-generating function is convex, it yields

$$\begin{aligned} t\big(w \ln M_X\big(t_1^{-1}\big) + (1-w)\ln M_Y\big(t_2^{-1}\big)\big) &\geq t \ln E\big(e^{wXt_1^{-1}+(1-w)Yt_2^{-1}}\big) \\ &= t \ln E\big(e^{\lambda X t^{-1}+(1-\lambda)Yt^{-1}}\big) \\ &= \big(\lambda t_1 + (1-\lambda)t_2\big)\ln M_{\lambda X+(1-\lambda)Y}\big(\big(\lambda t_1+(1-\lambda)t_2\big)^{-1}\big). \end{aligned}$$

This completes the proof. □

In the first sentence, coming after the proof of Proposition 4.2 of [1], the term "*the feasible set of problem (8)*" must be replaced by "*the feasible set of the problem obtained from problem (8) by using equation (2)*." Note that the problem obtained from problem (8) of [1] by using equation (2) of [1] is as follows:

$$\min_{\boldsymbol{w}\in\mathbf{W},t\in\mathbb{R}}\left\{t+\alpha^{-1}\mathrm{E}\left[-\sum_{i=1}^{n} w_i R_i - t\right]_+\right\}.$$

References

1. Ahmadi-Javid, A.: Entropic value-at-risk: A new coherent risk measure. J. Optim. Theory Appl. (2012), this issue
2. Jeyakumar, V., Wolkowicz, H.: Generalizations of Slater's constraint qualification for infinite convex programs. Math. Program., Ser. B **57**, 85–101 (1992)
3. Cheridito, P., Li, T.: Risk measures on Orlicz hearts. Math. Finance **19**, 189–214 (2009)