Czech Economic Review 7 (2013) 162–177
Acta Universitatis Carolinae Oeconomica

Risk Measures in Optimization Problems via Empirical Estimates

Vlasta Kaňková*

Received 27 January 2013; Accepted 14 October 2013

Abstract Economic and financial activities are often influenced simultaneously by a decision parameter and a random factor. Since mostly it is necessary to determine the decision parameter without knowledge of the realization of the random element, deterministic optimization problems depending on a probability measure often correspond to such situations. In applications the problem has to be very often solved on the data basis. It means that usually the "underlying" probability measure is replaced by empirical one. Great effort has been made to investigate properties of the corresponding (empirical) estimates; mostly under assumptions of "thin" tails and a linear dependence on the probability measure. The aim of this paper is to focus on the cases when these assumptions are not fulfilled. This happens usually just in economic and financial applications (see, e.g., Mandelbort 2003; Pflug and Römisch 2007; Rachev and Römisch 2002; Shiryaev 1999).

Keywords Static stochastic optimization problems, linear and nonlinear dependence, risk measures, thin and heavy tails, Wasserstein metric, $\mathcal{L}_1$ norm, empirical distribution function
JEL classification C44
AMS classification 90C15

1. Introduction

Let $(\Omega, \mathcal{S}, P)$ be a probability space; $\xi\,(:= \xi(\omega) = [\xi_1(\omega), \ldots, \xi_s(\omega)])$ an s-dimensional random vector defined on $(\Omega, \mathcal{S}, P)$; $F\,(:= F(z), z \in R^s)$ the distribution function of ξ; P_F, Z_F the probability measure and a support corresponding to F. Let, moreover, $g_0(:= g_0(x,z))$ be a function defined on $R^n \times R^s$; $X_F \subset R^n$ a nonempty set generally depending on F, $X \subset R^n$ a nonempty "deterministic" set. If E_F denotes the operator of mathematical expectation corresponding to F, then static rather general "classical" stochastic optimization problem can be introduced in the form:

$$\varphi(F, X_F) = \inf\{\mathsf{E}_F g_0(x, \xi) | x \in X_F\}. \tag{1}$$

The objective function in (1) depends linearly on the probability measure P_F. More general problems appeared recently. Some of them can be covered by the following type:

* Academy of Sciences of the Czech Republic, Institute of Information Theory and Automation, Department of Econometrics, Pod Vodárenskou věží 4, 182 08 Praha 8, Czech Republic. Phone: +420 266 052 501, E-mail: kankova@utia.cas.cz

$$\overline{\varphi}(F, X_F) = \inf\{\mathsf{E}_F \overline{g}_0(x, \xi, \mathsf{E}_F h(x, \xi))|x \in X_F\}, \tag{2}$$

where $h(:= h(x,z)) = (h_1(x,z), \dots, h_{m_1}(x,z))$ is an m_1-dimensional vector function defined on $R^n \times R^s$, $\overline{g}_0(:= \overline{g}_0(x,z,y))$ is a real-valued function defined on $R^n \times R^s \times R^{m_1}$.

Let us recall and analyze a relationship between risk measures and problems defined above, employing a few examples.

(i) If $L(:= L(x, z)$ defined on $R^n \times R^s)$ represents a loss function, then $VaR_\alpha(x) := \min\limits_u \{P\{\omega : L(x, \xi) \le u\} \ge \alpha\}$, $\alpha \in (0, 1)$ can be considered as a risk measure, known as "Value–at –Risk" (see, e.g., Dupačová 2009). Setting

$$X_F(:= X_F(u_0, \alpha)) = \{x \in X : [\min_u P\{\omega : L(x, \xi) \le u\} \ge \alpha] \le u_0\}, \tag{3}$$

with $u_0 \in R^1$, we can obtain the problem with risk measure in constraints.

(ii) $CVaR_\alpha(x) = \min\limits_{v \in R}[v + \frac{1}{1-\alpha}\mathsf{E}_F(L(x, \xi) - v)^+]$ is another risk measure known as "Conditional Value–at–Risk" (see, e.g., Dupačová 2009). Evidently, the objective function $CVaR_\alpha(x)$ can be written in the form

$$CVaR_\alpha(x) = \overline{g}_0(x, z, y) \quad \text{with} \quad \overline{g}_0(x, z, y) = \min_{v \in R}[v + \frac{1}{1-\alpha} y],$$

where $y = \mathsf{E}_F h(x, \xi)$ and $h(x, z) (:= (h(x, v, z)) = (L(x, z) - v)^+$.

It could happen that the corresponding optimization problem does not depend linearly on the probability measure. However, considering the case when $X_F = X$ and employing the result of Rockafellar and Uryasev (2002) the problem can be rewritten in the form:

$$\min_{(v,x) \in R^1 \times X} \{v + \frac{1}{1-\alpha} \mathsf{E}_F(L(x, \xi) - v)^+\}, \tag{4}$$

in which the dependence on the probability measure is surely linear.

(iii) Employing Markowitz approach to very simple portfolio problem:

$$\max \sum_{k=1}^{n} \xi_k x_k$$
$$\text{s.t.} \quad \sum_{k=1}^{n} x_k \le 1, \quad x_k \ge 0, \quad k = 1, \dots, n$$

with x_k a fraction of the unit wealth invested in the asset k, ξ_k the return of the asset, we can introduce the Markowitz problem (see, e.g., Dupačová et al. 2002):

$$\begin{aligned} &\varphi^M(F) = \max\{\sum_{k=1}^{n} \mu_k x_k - K\sum_{k=1}^{n}\sum_{j=1}^{n} x_k c_{k,j} x_j\} \\ \text{s. t.} \quad &\sum_{k=1}^{n} x_k \le 1, \quad x_k \ge 0, \quad k = 1, \dots, n, \quad K > 0 \; constant, \end{aligned} \tag{5}$$

where $\mu_k = \mathsf{E}_F \xi_k$, $c_{k,j} = \mathsf{E}_F(\xi_k - \mu_k)(\xi_j - \mu_j)$, $k, j = 1, \dots n$. The dependence on the probability measure in (5) is not linear.

Let us analyze the problem (5). Since

$$\sum_{k=1}^{n} \mu_k x_k - K\sum_{j=1}^{n}\sum_{k=1}^{n} x_k c_{k,j} x_j = \mathsf{E}_F\{\sum_{k=1}^{n} \xi_k x_k - K\sum_{k=1}^{n}\sum_{j=1}^{n}[x_k\xi_k\xi_j x_j - x_k\xi_k \mathsf{E}_F \xi_j x_j]\},$$

setting

$$\begin{aligned} \bar{g}_0(x, \xi, y) &= \sum_{k=1}^{n} \xi_k x_k - K\sum_{k=1}^{n}\sum_{j=1}^{n}[x_k\xi_k\xi_j x_j + x_k\xi_k y_j], \\ h_j(x, z) &= z_j x_j, \quad j = 1, \dots, n \end{aligned}$$

we can see that the problem (5) can be written in the form of the problem (2). However, later we can recognize that the sufficient assumptions (introduced in this paper) guaranteing "good" rate convergence will be fulfilled only in the case when the support Z_F is a bounded set.

Evidently, $\sum_{j=1}^{n} x_k c_{k,j} x_j$ can be considered as a risk measure that can be replaced by another risk measure, for example by $\mathsf{E}_F|\sum_{k=1}^{n}\xi_k x_k - \mathsf{E}_F[\sum_{k}^{n}\xi_k x_k]|$, (see Konno and Yamazaki 1991). The dependence on the probability measure is again nonlinear. However, the above mentioned assumptions will be (for $X_F = X$; X compact) already fulfilled.

In applications we have often to replace the measure P_F by an empirical measure P_{F^N} determined by a random sample corresponding to the measure P_F. Consequently, instead of the problems (1) and (2), the following problems are solved:

$$\varphi(F^N, X_{F^N}) = \inf\{\mathsf{E}_{F^N} g_0(x, \xi) | x \in X_{F^N}\}, \tag{6}$$

$$\bar{\varphi}(F^N, X_{F^N}) = \inf\{\mathsf{E}_{F^N} \bar{g}_0(x, \xi, \mathsf{E}_{F^N} h(x, \xi)) | x \in X_{F^N}\}. \tag{7}$$

By solving (6) and (7), we obtain estimates of the optimal values and optimal solutions. The investigation of these estimates started (for problem (1) with $X_F = X$) in Wets (1974), followed by many papers (e.g., Dai et al. 2000; Dupačová and Wets 1984; Kaniovski et al. 1995; Kaňková 1978, 1994; Pflug 1999; Römisch 2003; Shapiro 2003). The consistency, the convergence rate and asymptotic distributions have been studied therein under the assumptions of "weak" tailed distributions and $X_F = X$. The exception are, e.g., Kaňková (2010), Houda and Kaňková (2012), and Rachev and Römisch (2002). We focus to the problem (2), the case of "heavy" tails and a special type of the set X_F. Especially, we assume that either $X_F = X$ or that there exist real valued functions $g_i (:= g_i(x), x \in R^n)$, $i = 1, \ldots, s$ such that

$$
\begin{aligned}
& X_F (:= X_F(u_0, \alpha)) = \bigcap_{i=1}^{s} \{x \in X : \min_{u^i} \{P[\omega : L_i(x, \xi) \le u^i] \ge \alpha_i\} \le u_0^i\}, \\
& \text{with } u_0 = (u_0^1, \ldots, u_0^s),\ \alpha = (\alpha_1, \ldots, \alpha_s),\ u_0^i > 0,\ \alpha_i \in (0,1),\ i = 1, \ldots, s, \\
& L_i(x, z) = g_i(x) - z_i,\ i = 1, \ldots, s,\ z = (z_1, \ldots, z_s).
\end{aligned}
\tag{8}
$$

$L_i(x, z)$, $i = 1, \ldots, s$ can be considered as loss functions. This type of loss function can appear, e.g., in a connection with an inner problem in two stage stochastic (generally nonlinear) programming problems (for a definition of two-stage problems see, e.g., Birge and Louveaux 1992).

Remark 1.

(i) *Evidently, problem* (2) *covers problem* (1) *with* $\overline{g}_0(x, z, y) := g_0(x, z)$.

(ii) *Distribution functions with heavy tails appear mainly in economic, financial and energetic problems (see, e.g., Kozubowski et al. 2003; Meerchaert and Scheffler 2003). Pareto and Weibull distributions belong to this class of the distributions. However, stable distributions (with an exception of the normal distribution) are their main representatives. (For the definition of the stable distributions see, e.g., Klebanov 2003, or Meerchaert and Scheffler 2003). The relationship between the stable distributions and Pareto distribution can be found in Shiryaev (1999).*

2. Some definitions and auxiliary assertions

First, if F, G are two arbitrary s-dimensional distribution functions for which the problem (2) is well defined, then according to the triangular inequality we obtain

$$
|\overline{\varphi}(F, X_F) - \overline{\varphi}(G, X_G)| \le |\overline{\varphi}(F, X_F) - \overline{\varphi}(G, X_F)| + |\overline{\varphi}(G, X_F) - \overline{\varphi}(G, X_G)|. \tag{9}
$$

Furthermore, if we denote by the symbols F_i, $i = 1, \ldots, s$ one-dimensional marginal distribution functions corresponding to the distribution function F and if we can assume that F_i, $i = 1, \ldots, s$ are absolutely continuous with respect to the Lebesgue measure on R^1, $g_i(x)$, $i = 1, \ldots, s$ are continuous functions on X and the relation (8) is fulfilled, then (under the assumption that F_i, $i = 1, \ldots, s$ are increasing functions in a neighbourhoods of $k_{F_i}(\alpha_i)$) we can obtain

$$
\begin{aligned}
X_F &= \bigcap_{i=1}^{s} \{x \in X : \min_{u^i} \{P[\omega : g_i(x) - u^i \le \xi_i] \ge \alpha_i\} \le u_0^i\} \\
&= \bigcap_{i=1}^{s} \{x \in X : \min_{u^i} [g_i(x) - u^i \le k_{F_i}(\alpha_i)]\} \le u_0^i\} \\
&= \bigcap_{i=1}^{s} \{x \in X : [g_i(x) - u_0^i \le k_{F_i}(\alpha_i)]\},
\end{aligned}
\tag{10}
$$

where $k_{F_i}(\alpha_i) = \sup\{z_i : P_{F_i}\{\omega : z_i \le \xi_i(\omega)\} \ge \alpha_i\}$.

Consequently, setting

$$\overline{X}(v) = \bigcap_{i=1}^{s} \{x \in X : g_i(x) - u_0^i \le v_i\}, \, v = (v_1, \ldots, v_s),$$

we obtain

$$X_F = \overline{X}(k_F(\alpha)), \quad k_F(\alpha) = (k_{F_1}(\alpha_1), \ldots, k_{F_s}(\alpha_s)). \tag{11}$$

Definition 1. (Rockafellar and Wets 1998) *If $X', X'' \subset R^n$ are two non-empty sets, then the Hausdorff distance of these sets $\Delta_n[X', X'']$ is defined by*

$$\begin{aligned} \Delta_n[X', X''] &= \max[\delta_n(X', X''), \delta_n(X'', X')], \\ \delta_n(X', X'') &= \sup_{x' \in X'} \inf_{x'' \in X''} \|x' - x''\|, \end{aligned}$$

where $\|\cdot\| = \|\cdot\|_n^2$ denotes the Euclidean norm in R^n.

Proposition 1. *Let X be a nonempty compact set. If*

(i) *$\hat{g}_0 (:= \hat{g}_0(x), x \in R^n)$ is a Lipschitz function on X with the Lipschitz constant L,*

(ii) *$\overline{X}(v)$, $v \in Z_F$ is a nonempty set for every $v \in Z_F$ and, moreover, there exists a constant $\hat{C} > 0$ such that*

$$\Delta_n[\overline{X}(v(1)), \overline{X}(v(2))] \le \hat{C}\|v(1) - v(2)\|, \quad v(1), v(2) \in Z_F,$$

then

$$|\inf_{x \in \overline{X}(v(1))} \hat{g}_0(x) - \inf_{x \in \overline{X}(v(2))} \hat{g}_0(x)| \le L\hat{C}\,\|v(1) - v(2)\|, \quad v(1), v(2) \in Z_F.$$

Proof. The assertion of Proposition 1 is a little modified assertion of Proposition 1 in Kaňková (1997).

Lemma 1. (Kaňková 1997) *Let X be a nonempty convex compact set. If*

(i) *$g_i(x)$, $i = 1, \ldots, s$ are convex continuous bounded functions on X,*

(ii) *$\overline{X}(v)$ is a nonempty set for every $v \in Z_F$,*

then there exists $C > 0$ such that

$$\Delta_n[\overline{X}(v(1)), \overline{X}(v(2))] \le C\|v(1) - v(2)\| \quad for \quad v(1), v(2) \in Z_F.$$

To recall stability results, let $\mathcal{P}(R^s)$ be the set of Borel probability measures on $R^s, s \ge 1$; $\mathcal{M}_1^1(R^s) = \{P \in \mathcal{P}(R^s) : \int_{R^s} \|z\|_s^1 P(dz) < \infty\}$; $\|\cdot\|_s^1$ denote $\mathcal{L}_1$ norm in R^s. Borrowing notation from Kaňková (1997), we introduce the following assumptions:

B.1 $P_F, P_G \in \mathcal{M}_1(R^s)$, there exist $\varepsilon > 0$ such that

- $\bar{g}_0(x, z, y)$ is for $x \in X(\varepsilon)$, $z \in R^s$ a Lipschitz function of $y \in Y(\varepsilon)$ with a Lipschitz constant L_y; $Y(\varepsilon) = \{y \in R^{m_1} : y = h(x, z)$ for some $x \in X(\varepsilon)$, $z \in R^s\}$, $\mathsf{E}_F h(x, \xi)$, $\mathsf{E}_G h(x, \xi) \in Y(\varepsilon)$;
- for every $x \in X(\varepsilon)$, $y \in Y(\varepsilon)$ there exist finite mathematical expectations, $\mathsf{E}_F \bar{g}_0(x, \xi, \mathsf{E}_F h(x, \xi))$, $\mathsf{E}_F g_0^1(x, \xi, \mathsf{E}_G h(x, \xi))$, $\mathsf{E}_G g_0^1(x, \xi, \mathsf{E}_F h(x, \xi))$, $\mathsf{E}_G g_0^1(x, \xi, \mathsf{E}_G h(x, \xi))$;
- $h_i(x, z)$, $i = 1, \ldots, m_1$ are for every $x \in X(\varepsilon)$ Lipschitz functions of z with the Lipschitz constants L_h^i (corresponding to $\mathcal{L}_1$ norm),
- $\bar{g}_0(x, z, y)$ is for every $x \in X(\varepsilon)$, $y \in Y(\varepsilon)$ a Lipschitz function of $z \in R^s$ with the Lipschitz constant L_z (corresponding to $\mathcal{L}_1$ norm).

($X(\varepsilon)$, $\varepsilon > 0$ denotes ε- neighbourhood of X.)

B.2 $\mathsf{E}_F \bar{g}_0(x, \xi, \mathsf{E}_F h(x, \xi))$ is a continuous function on X.

Remark 2. *Fulfilling at least one of the next assumptions:*

B.3 $\bar{g}_0(x, z, y)$, $h(x, z)$ *are uniformly continuous functions on* $X(\varepsilon) \times R^s \times Y(\varepsilon)$,

B.4 X *is a convex set and there exists* $\varepsilon > 0$ *such that* $\bar{g}_0(x, \xi, \mathsf{E}_F h(x, \xi))$ *is a convex function on* $X(\varepsilon)$,

guarantees fulfilling of Assumption B.2.

Lemma 2. *Let G be an arbitrary s-dimensional distribution function, $\varepsilon > 0$, X be a nonempty compact set. If Assumptions B.1 and B.2 are fulfilled (for $F = G$), then* $\mathsf{E}_G \bar{g}_0(x, \xi, E_G h(x, \xi))$ *is a uniformly continuous function on* $X(\varepsilon)$.

Lemma 3. *Let G be an arbitrary s-dimensional distribution function, $\varepsilon > 0$. Let, moreover, Assumption B.1 be fulfilled (for $F = G$). If $\bar{g}_0(x, z, y)$, $h(x, z)$ are Lipschitz functions on $X \times Y(\varepsilon)$ with the Lipschitz constants L_{g_0} and L_h, then there exists a constant $\bar{C}$ such that $\mathsf{E}_G \bar{g}_0(x, \xi, E_G h(x, \xi))$ is a Lipschitz function on X with the Lipschitz constant $\bar{C}$.*

Proof. The assertion of Lemma 3 follows from the properties of the Lipschitz functions and integrals. □

Since it follows from the Assumption B.1 and the triangular inequality that

$$|\mathsf{E}_F \bar{g}_0(x, \xi, \mathsf{E}_F h(x, \xi)) - \mathsf{E}_G \bar{g}_0(x, \xi, \mathsf{E}_G h(x, \xi))| \le$$
$$L_y \| \mathsf{E}_F h(x, \xi) - \mathsf{E}_G h(x, \xi)\| + |\mathsf{E}_F \bar{g}_0(x, \xi, \mathsf{E}_G h(x, \xi)) - \mathsf{E}_G \bar{g}_0(x, \xi, \mathsf{E}_G h(x, \xi))| \tag{12}$$

we can introduce one essential Proposition 2.

Proposition 2. *Let $P_F, P_G \in \mathcal{M}_1^1(R^s)$, the assumptions B.1, B.2 be fulfilled (for both F, G), then there exist $\hat{C} > 0$ such that the following relation*

$$|\mathsf{E}_F\overline{g}_0(x,\xi,\mathsf{E}_F h(x,\xi)) - \mathsf{E}_G\overline{g}_0(x,\xi,\mathsf{E}_G h(x,\xi))| \le \hat{C}\sum_{i=1}^{s}\int_{-\infty}^{\infty}|F_i(z_i) - G_i(z_i)|dz_i$$

holds for $x \in X$.

If, moreover, X *is a compact set, then also*

$$|\overline{\varphi}(F,X) - \overline{\varphi}(G,X)| \le \hat{C}\sum_{i=1}^{s}\int_{-\infty}^{\infty}|F_i(z_i) - G_i(z_i)|dz_i. \tag{13}$$

Proof. Employing the main idea of the proof of Lemma 3.1 in Kaňková and Houda (2006) we can see that the assertion follows from the relation (12) and assumption B.1 (for more details see also the proof of Proposition 4.1 in Kaňková 2010). □

Remark 3.

- $\hat{C} = L_z$ *in the case of problem* (1).
- *Proposition 2 reduces (from the mathematical point of view) s-dimensional case to one-dimensional. However, the dependence between components of the random vector* ξ *is there neglected. The idea to reduce s-dimensional case to one-dimensional appeared already in Pflug (2001) (see also Šmíd 2009).*
- *It follows from Proposition 2 that the stability of the problem* (2) *can be bounded by* $\sum_{i=1}^{s}\int_{-\infty}^{\infty}|F_i(z_i) - G_i(z_i)|dz_i$. *Consequently, we obtained very similar situation to the problem* (1) *(for more details see, e.g. Houda and Kaňková 2012).*

3. Problem analysis

To employ the assertion of Proposition 2 to empirical estimates we introduce a system of the assumptions:

A.1 $\{\xi^i\}_{i=1}^{\infty}$ is an independent random sequence corresponding to F (we denote by the symbol F^N the empirical distribution function determined by $\{\xi^i\}_{i=1}^{N}$);

A.2 $P_{F_i}, i = 1,\dots s$ are absolutely continuous w. r. t. the Lebesgue measure on R^1;

A.3 For every $i \in \{1,\dots,s\}$ there exist $\delta > 0$ and $\vartheta > 0$ such that $f_i(z_i) > \vartheta$ for $z_i \in Z_{F_i}$, $|z_i - k_{F_i}(\alpha_i)| < 2\delta$ (f_i denote probability densities corresponding to F_i, $i = 1,\dots,s$).

Furthermore we recall some assertions important for our investigation.

Lemma 4. (Shorack and Welner 1986) *Let* $s = 1$, $P_F \in \mathcal{M}_1^1(R^1)$ *and A.1 be fulfilled. Then*

$$P\{\omega : \int_{-\infty}^{\infty}|F(z) - F^N(z)|\,dz \longrightarrow_{N\longrightarrow\infty} 0\} = 1.$$

Lemma 5. *Let* $s = 1$, $\alpha \in (0, 1)$, $t > 0$. *If Assumptions A.1, A.2 and A.3 are fulfilled,* $0 < t < \delta$, *then*

$$P\{\omega : |k_{F^N}(\alpha) - k_F(\alpha)| > t\} \le 2\exp\{-2N(\vartheta t)^2\} \textit{ for } N \in \mathcal{N}.$$

($\mathcal{N}$ denotes the set of natural numbers.)

The proof of Lemma 5 will be given in Appendix.

Corollary 1. *Let* $\varepsilon > 0$, *X be a nonempty convex compact set. If*

(i) $\hat{g}_0(x)$, $x \in R^n$ *is a Lipschitz function on* $X(\varepsilon)$,

(ii) $g_i(x)$, $i = 1, \ldots, s$ *are convex continuous bounded functions on* $X(\varepsilon)$,

(iii) *Assumptions A.1, A.2 and A.3 are fulfilled,*

(iv) $\overline{X}(v)$ *are nonempty sets for* $v \in Z_F$,

then there exists a constant $C > 0$ *such that*

$$P\{\omega : |\inf_{\overline{X}(k_F(\alpha))} \hat{g}_0(x) - \inf_{\overline{X}(k_{F^N}(\alpha))} \hat{g}_0(x)|| > t\} \le 2s\exp\{-2N(\vartheta t/LCs))^2\}$$

for every $N \in \mathcal{N}$ *and* $t > 0$ *such that* $0 < t < 2\delta$.

Proof. The assertion of Corollary 1 follows from Proposition 1, Lemmas 1 and 5, relations (10), (11) and the properties of the Euclidean norm. □

Remark 4. *Assumptions under which the sets* $X(v)$ *are nonempty can be found, e.g., in Birge and Louveaux (1992) or in Kaňková (2007a).*

Proposition 3. (Houda and Kaňková 2012) *Let* $s = 1$, $t > 0$ *and Assumptions A.1 and A.2 be fulfilled. If there exists* $\beta > 0, R := R(N) > 0$ *defined on* $\mathcal{N}$ *such that* $R(N) \longrightarrow_{N \longrightarrow \infty} \infty$ *and, moreover,*

$$\begin{gathered}
N^{\beta} \int_{-\infty}^{-R(N)} F(z)dz \longrightarrow_{N \longrightarrow \infty} 0, \qquad N^{\beta} \int_{R(N)}^{\infty} [1 - F(z)] \longrightarrow_{N \longrightarrow \infty} 0, \\
2NF(-R(N)) \longrightarrow_{N \longrightarrow \infty} 0, \qquad 2N[1 - F(R(N))] \longrightarrow_{N \longrightarrow \infty} 0, \\
\left(\frac{12N^{\beta}R(N)}{t} + 1 \right) \exp\{-2N \left(\frac{t}{12R(N)N^{\beta}} \right)^2\} \longrightarrow_{N \longrightarrow \infty} 0,
\end{gathered} \tag{14}$$

then

$$P\{\omega : N^{\beta} \int_{-\infty}^{\infty} |F(z) - F^N(z)|dz > t\} \longrightarrow_{N \longrightarrow \infty} 0. \tag{15}$$

According to the results of Dvoretzky et al. (1956) we can see that (under the assumptions of Proposition 3) the validity of the relation (14) depends on the tails behaviour.

Proposition 4. (Houda and Kaňková 2012) *Let $s = 1$, $t > 0$, $r > 0$, Assumptions A.1 and A.2 be fulfilled. Let moreover, ξ be a random variable such that $\mathsf{E}_F|\xi|^r < \infty$. If constants $\beta, \gamma > 0$ fulfil the inequalities*

$$0 < \beta + \gamma < 1/2, \quad \gamma > 1/r, \quad \beta + (1-r)\gamma < 0, \tag{16}$$

then the relations (14) *are valid.*

Propositions 3 and 4 guarantee an existence of $\beta > 0$, $(\beta := \beta(r))$ fulfilling (14) and consequently (15) only in the case when there exists $r > 2$ such that $\mathsf{E}_F|\xi|^r < +\infty$. According to heavy tailed distributions it means that the tail parameter ν has to be greater than 2. Consequently existence of $\beta > 0$ is not guaranteed in the case of stable distributions with the shape parameter $\nu < 2$ (for the definition of the stable distribution and shape parameter see, e.g, Klebanov 2003). The case $\nu \in (1, 2)$ corresponds very often to random elements appearing in the financial applications. To include it in our investigation we recall the following assertion.

Proposition 5. *Let $s = 1$, $\{\xi^i\}_{i=1}^N$, $N = 1, 2, \ldots$ be a sequence of independent random values corresponding to a heavy tailed distribution F with the shape parameter $\nu \in (1, 2)$. Then the sequence*

$$\frac{N}{N^{1/\nu}} \int_{-\infty}^{\infty} |F^N(z) - F(z)| dz, N = 1, \ldots \tag{17}$$

is stochastically bounded if and only if

$$\sup_{t>0} t^{\nu} P\{\omega : |\xi| > t\} < \infty. \tag{18}$$

Proof. The assertion of Proposition 5 follows from Theorem 2.2 of Burrio et al. (1999). □

According to the definition of the stochastically bounded random sequences it means (under the relation (18)) that

$$\lim_{M \longrightarrow \infty} \sup_N P\{\omega : \frac{N}{N^{1/\nu}} \int_{-\infty}^{\infty} |F(z) - F^N(z)| > M\} = 0. \tag{19}$$

Remark 5. *The shape parameter ν is determined by tails of the distribution. Smaller shape parameter corresponds to more heavy tails.*

Lemma 6. *Let $\alpha_i \in (0, 1)$, $u_0^i > 0$, $i = 1, \ldots, s$. If*

(i) *X is a compact convex set,*

(ii) *$g_i(x)$, $i = 1, \ldots, s$ are convex continuous bounded function on X,*

(iii) *X_F defined by the relation* (10) *are nonempty for $v \in Z_F$,* *then X_F is a compact set and fulfils the relation* (11).

Proof. The assertion of Lemma 6 follows from the properties of convex functions, convex sets and the relation (10). □

4. Main Results

Applying the auxiliary assertions from the former parts we obtain the following results.

Theorem 1. *Let Assumptions B.1, B.2, A.1 be fulfilled, X be a compact set and $P_F \in \mathcal{M}_1^1(R^s)$, then*

$$P\{\omega : |\overline{\varphi}(F^N, X) - \overline{\varphi}(F, X)| \longrightarrow_{N \longrightarrow \infty} 0\} = 1.$$

Proof. The assertion of Theorem 1 follows from Proposition 2 and Lemma 4. □

Remark 6. *It follows from Theorem 1 and the properties of the stable distributions (see. e.g., Klebanov 2003) that (under general assumptions) $\overline{\varphi}(F^N, X)$ is a consistent estimate of $\overline{\varphi}(F, X)$ also for all stable distributions with the shape parameter $\nu \in (1, 2\rangle$.*

Furthermore, we shall deal with a convergence rate.

Theorem 2. *Let Assumptions B.1, B.2, A.1, A.2 be fulfilled, $P_F \in \mathcal{M}_1^1(R^s)$, $X_F = X$, $t > 0$. If*

(i) *for some $r > 2$ it holds that $\mathsf{E}_{F_i}|\xi_i|^r < +\infty$, $i = 1, \ldots, s$,*

(ii) *$\beta, \gamma > 0$ fulfil the inequalities $0 < \beta + \gamma < 1/2$, $\gamma > 1/r$, $\beta + (1 - r)\gamma < 0$,*

then

$$P\{\omega : \sup_{x \in X} N^{\beta} |\mathsf{E}_{F^N} \overline{g}_0(x, \xi, \mathsf{E}_{F^N} h(x, \xi)) - \mathsf{E}_F \overline{g}_0(x, \xi, \mathsf{E}_F h(x, \xi))| > t\} \longrightarrow_{N \longrightarrow \infty} 0. \tag{20}$$

If, moreover, X is a compact set, then also

$$P\{\omega : N^{\beta} |\overline{\varphi}(F, X) - \overline{\varphi}(F^N, X)| > t\} \longrightarrow_{N \longrightarrow \infty} 0. \tag{21}$$

Proof. The first assertion follows from Proposition 2 and Proposition 4. The second assertion follows from the first one and from the properties of integrals and compact sets. (See a similar proof for the problem (1) in Houda and Kaňková 2012). □

The next assertion deals with a special case in which the objective function does not depend on the probability measure.

Theorem 3. *Let $\varepsilon > 0$, $t > 0$, $\alpha_i \in (0, 1)$, $i = 1, \ldots, s$. Let moreover X be a nonempty convex compact set. If*

(i) *$\hat{g}_0(x)$, $x \in R^n$ is a Lipschitz function with the Lipschitz constant L,*

(ii) $\overline{g}_0(x, z, y) = \hat{g}_0(x)$, $x \in R^n$, $y \in R^s$, $y \in R^{m_1}$,

(iii) *Assumptions A.1, A.2 and A.3 are fulfilled,*

(iv) $\overline{X}(v)$*, defined by the relation* (10)*, is nonempty set for* $v \in Z_F$ *such that* $\|v - k_F(\alpha)\| < 2\delta$,

(v) $g_i(x)$*,* $i = 1, \dots, s$ *are convex continuous bounded functions on* $X(\varepsilon)$,

then

$$P\{\omega : N^{\beta}|\overline{\varphi}(F, X_F) - \overline{\varphi}(F^N, X_{F^N})| > t\} \longrightarrow_{N \longrightarrow \infty} 0 \text{ for } \beta \in (0, 1/2). \tag{22}$$

Proof. The assertion of Theorem 3 follows from the assertion of Corollary 1. □

Remark 7. *Evidently setting* $G = F^N$ *in Lemma 3 we can see that (under the corresponding assumptions)*

$$\mathsf{E}_F \overline{g}_0(x, \xi, \mathsf{E}_F h(x, \xi)), \, \mathsf{E}_{F^N} \overline{g}_0(x, \xi, \mathsf{E}_{F^N} h(x, \xi))$$

are Lipschitz functions on X with the same Lipschitz constant not depending on $\omega \in \Omega$. *Consequently, we can see that the assertion of Theorem 3 is valid if the function* $\overline{g}_0$ *fulfils the assumptions of Lemma 3 (instead of the assumptions (i), (ii) of Theorem 3).*

According to the last Remark applying the relation (9) we can present the following assertion.

Theorem 4. *Let* $t > 0$*, X be a nonempty convex compact set,* $P_F \in \mathcal{M}_1^1(R^s)$*. Let, moreover,* $\alpha_i \in (0, 1)$*,* $u_0^i > 0$*,* $\alpha = (\alpha_1, \dots, \alpha_s)$ *,* $u_0 = (u_0^1, \dots, u_0^s)$*. If*

(i) *Assumptions B.1, B.2, A.1, A.2 and A.3 are fulfilled,*

(ii) *for some* $r > 2$ *it holds that* $\mathsf{E}_{F_i}|\xi_i|^r < +\infty$*,* $i = 1, \dots, s$,

(iii) $\beta, \gamma > 0$ *fulfil the inequalities* $0 < \beta + \gamma < 1/2$*,* $\gamma > 1/r$*,* $\beta + (1 - r)\gamma < 0$,

(iv) $g_0(x, y, y)$*,* $h(x, z)$ *are Lipschitz functions on X with the Lipshitz constants* L_{g_0}*,* L_h *not depending on* $z \in Z_F$,

(v) $\overline{X}(v)$*, defined by the relation* (10)*, is nonempty set for* $v \in Z_F$ *such that* $\|v - k_F(\alpha)\| < 2\delta$,

(vi) $g_i(x)$*,* $i = 1, \dots, s$ *are convex continuous bounded functions on* $X(\varepsilon)$,

then

$$P\{\omega : N^{\beta}|\overline{\varphi}(F, X_F) - \overline{\varphi}(F^N, X_{F^N})| > t\} \longrightarrow_{N \longrightarrow \infty} 0. \tag{23}$$

Proof. Setting in Relation (9) $G = F^N$ and applying Remark under Theorem 3 we see that the assertion follows from Theorems 2, 3. □

Evidently, the convergence rate $\beta := \beta(r)$ introduced by Theorems 2, 4 depends on the absolute moments existence; it holds that $\beta(r) \longrightarrow_{r \longrightarrow \infty} 1/2$, $\beta(r) \longrightarrow_{r \longrightarrow 2^+} 0$.

Consequently, the best convergence rate is valid not only for exponential tails but also for every distribution with finite all absolute moments (e.g. Weibull); even in the case when finite moment generating function does not exist. Unfortunately we can not obtain (by this approach) any results in the case when there exist only finite $\mathsf{E}_F|\xi_i|^r$, $i = 1, \ldots, s$ for $r < 2$. This is the case of stable distributions (with exception of normal distribution) or the case of Pareto distribution with a shape parameter $\nu \le 2$.

Theorem 5. *Let the assumptions B.1, B.2, A.1 and A.2 be fulfilled, $P_F \in \mathcal{M}_1^1(R^s)$, $M > 0$, X be a compact set. If one dimensional components ξ_i, $i = 1, \ldots, s$ of the random vector ξ have the distribution functions F_i with tail parameters $\nu_i \in (1, 2)$ fulfilling the relations*

$$\sup_{t>0} t^{\nu_i} P\{\omega : |\xi_i| > t\} < \infty, \qquad i = 1, 2, \ldots, s,$$

then

$$\lim_{M \longrightarrow \infty} \sup_N P\{\omega : \frac{N}{N^{1/\nu}} |\overline{\varphi}(F^N, X) - \overline{\varphi}(F, X)| > M\} = 0 \quad \textit{with} \quad \nu = \min(\nu_1, \ldots, \nu_s). \tag{24}$$

Proof. Let $M > 0$, $\nu \in (1, 2)$. First, it follows from Proposition 2 successively that

$$\begin{aligned}
&\sup_N P\{\omega : \tfrac{N}{N^{1/\nu}} |\overline{\varphi}(F^N, X) - \overline{\varphi}(F, X)| > M\} &&\le \\
&\sup_N P\{\omega : \tfrac{N}{N^{1/\nu}} L \sum_{i=1}^{s} \int_{-\infty}^{\infty} |F^N(z) - F(z)| dz > M\} &&\le \\
&\sum_{i=1}^{s} \sup_N P\{\omega : \tfrac{N}{N^{1/\nu_i}} \int_{-\infty}^{\infty} |F^N(z) - F(z)| dz > M/Ls\}.
\end{aligned}$$

Consequently, according to Proposition 5 and (19) we can obtain

$$\begin{aligned}
&\lim_{M \longrightarrow \infty} \sup_N P\{\omega : \tfrac{N}{N^{1/\nu}} |\overline{\varphi}(F^N, X) - \overline{\varphi}(F, X)| > M\} \le \\
&\sum_{i=1}^{s} \lim_{M' \longrightarrow \infty} \sup_N P\{\omega : \tfrac{N}{N^{1/\nu_i}} \int_{-\infty}^{\infty} |F^N(z) - F(z)| dz > M'\}, \quad M' = M/Ls.
\end{aligned}$$

Now already we can see that the assertion of Theorem 5 holds. □

5. Conclusion

The paper deals with empirical estimates in the case of static stochastic optimization problem. In particular the paper deals with the rate of convergence of optimal value estimate in the case when the dependence on the probability measure is not linear. Evidently, the results of Houda and Kaňková (2012), Kaňková (2012) are generalized. Moreover, it was shown that the results corresponding to the case of nonlinear dependence are very similar to them in the case of linear dependence. Consequently, both results can be completed by simulation technique credited to M. Houda and published in Houda and Kaňková (2012) and to V. Omelchenko (for stable distribution) that can be found in Omelchenko (2012).

The paper deals only with the optimal value estimates. Employing some growth conditions (see, e.g. Römisch 2003) the introduced results can be transformed to the estimates of the optimal solution. However the investigation in this direction is beyond the scope of this paper.

Acknowledgment The research was supported by the Czech Science Foundation under Grants P402/10/0956, P402/11/0150 and P402/10/1610.

References

Barrio, E., Giné, E. and Matrán, C. (1999). Central Limit Theorems for the Wassersten Distance between the Empirical and the True Distributions. *Annals of Probability,* 27(2), 1009–1071.

Birge, J. R. and Louveaux, F. (1992). *Introduction in Stochastic Programming.* Berlin, Springer.

Dai, L., Chen, C. H. and Birge, J. N. (2000). Convergence Properties of Two-Stage Stochastic Programming. *Journal of Optimization Theory and Applications*, 106, 489–509.

Dvoretzky, A., Kiefer, J. and Wolfowitz, J. (1956). Asymptotic Minimax Character of the Sample Distribution Function and the Classical Multinomial Estimate. *The Annals of Mathematical Statistics,* 27, 642–669.

Dupačová, J. and Wets, R. J.-B. (1984). Asymptotic Behaviour of Statistical Estimates and Optimal Solutions of Stochastic Optimization Problems. *The Annals of Statistics,* 16, 1517–1549.

Dupačová, J., Hurt, J. and Štěpán, J. (2002). *Stochastic Modelling in Economics and Finance.* New York, Kluwer.

Dupačová, J. (2009). *Portfolio Optimization and Risk Management via Stochastic Programming.* Osaka, Osaka University Press.

Houda, M. and Kaňková, V. (2012). Empirical Estimates in Economic and Financial Optimization Problems. *Bulletin of the Czech Econometric Society,* 19(29), 50–69.

Kaniovski, Y. M., King, A. J. and Wets, R. J.-B. (1995). Probabilistic Bounds (via Large Deviations) for the Solutions of Stochastic Programming Problems. *Annals of Operations Research,* 56, 189–208.

Kaňková, V. (1978). An Approximative Solution of Stochastic Optimization Problem. In Kožešník, J. (ed.), *Transactions of the Eighth Prague Conference 1978.* Prague, Academia, 349–353.

Kaňková, V. (1994). A Note on Estimates in Stochastic Programming. *Journal of Computational and Applied Mathematics,* 56, 97–112.

Kaňková, V. (1997). On the Stability in Stochastic Programming: The Case of Individual Probability Constraints. *Kybernetika*, 33(5), 525–546.

Kaňková, V. and Houda, M. (2006). Empirical Estimates in Stochastic Programming. In Hušková, M. and Janžura, M. (eds.), *Proceedings of Prague Stochastics 2006.* Prague, Matfyzpress, 426–436.

Kaňková, V. (2007a). Stochastic Programming Problems with Recourse via Multiobjective Optimization Problems. In Czechárová, K., Halická, M., Barbelová, V. and Lacko, V. (eds.), *Proceedings 15th International Scientific Conference on Mathematical Methods in Economics and Industry*. Univerzita P. J. Šafárika v Košicích, Košice, 68–78.

Kaňková, V. (2010). Empirical Estimates in Stochastic Optimization via Distribution Tails. *Kybernetika*, 46(3), 459–471.

Kaňková, V. (2012). Empirical Estimates in Optimization Problems: Survey with Special Regard to Heavy Tails and Dependent Sample. *Bulletin of the Czech Econometric Society,* 19(30), 50–69.

Klebanov, L. B. (2003). *Heavy Tailed Distributions.* Prague, Matfyzpress.

Kozubowski, T. J., Panorska A. K. and Rachev, S. T. (2003). Statistical Issues in Modelling Stable Portfolios. In Rachev, S. T. (ed.), *Handbook of Heavy Tailed Distributions in Finance*. Amsterdam, Elsevier, 134–168.

Konno, H. and Yamazaki, H. (1991). Mean-Absolute Deviation Portfolio Optimization Model and Its Application to Tokyo Stock Market. *Management Science,* 37(5), 519–531.

Mandelbort, M. M. (2003). Heavy Tails in Finance for Independent or Multifractal Price Increments. In Rachev, S. T. (ed.), *Handbook of Heavy Tailed Distributions in Finance.* Amsterdam, Elsevier, 1–34.

Meerschaert, M. M. and Scheffler, H. P. (2003). Portfolio Modelling with Heavy Tailed Random Vectors. In Rachev, S. T. (ed.), *Handbook of Heavy Tailed Distribution in Finance.* Amsterdam, Elsevier, 595–640.

Omelchenko, V. (2012). Behaviour and Convergence of Wasserstein Metric in the Framework of Stable Distributions. *Bulletin of the Czech Econometric Society,* 19(30), 124–138.

Pflug, G. CH. (1999). Stochastic Programs and Statistical Data. *Annals of Operations Research*, 85, 59–78.

Pflug, G. CH. (2001). Scenario Tree Generation for Multiperiod Financial Optimization by Optimal Discretization. *Mathematical Programming Series B*, 89(2), 251–271.

Pflug, G. Ch. and Römisch, W. (2007). *Modeling, Measuring and Managing Risk.* Singapore, World Scientific Publishing Company.

Rachev, S. T. and Römisch, W. (2002). Quantitative Stability and Stochastic Programming: The Method of Probability Metrics. *Mathematics of Operations Research,* 27, 792–818.

Rockafellar, T. R. and Wets, R. J-B. (1998). *Variational Analysis.* Berlin, Springer.

Rockafellar, T. R. and Uryasev, S. P. (2002). Conditional Value-at-Risk for General Loss Distribution. *Journal of Banking and Finance,* 26, 1443–1471.

Römisch, W. (2003). Stability in Stochastic Programming Problems. In Ruszczynski, A. and Shapiro, A. (eds.), *Stochastic Programming. Handbooks in Operations Research and Management Science, Vol. 10.* Amsterdam, Elsevier, 483–554.

Shapiro, A. (2003). Monte-Carlo Sampling Methods. In Ruszczynski, A. and Shapiro, A. (eds.), *Stochastic Programming. Handbooks in Operations Research and Management Science, Vol. 10.* Amsterdam, Elsevier, 353–456.

Shiryaev, A. N. (1999). *Essentials in Stochastic Finance.* New Jersey, World Scientific.

Shorack, G. R. and Welner, J. A. (1986). *Empirical Processes with Applications in Statistics.* New York, John Wiley & Sons.

Šmíd, M. (2009). The Expected Loss in the Discretization of Multistage Stochastic Programming Problems – Estimation and Convergence rate. *Annals of Operations Research,* 165(1), 29–45.

Wets, R. J-B. (1974). A Statistical Approach to the Solution of Stochastic Programs with (Convex) Simple Recourse. University Kentucky, Working Paper.

Appendix

Proof of Lemma 5. First evidently, $k_F(\alpha)$ depends on $\omega \in \Omega$, and, moreover, if (for some $\omega \in \Omega$) it holds that $F^N(k_F(\alpha)+t) > 1-\alpha$, $F^N(k_F(\alpha)-t) < 1-\alpha$, then

$$k_{F^N}(\alpha) \in \langle k_F(\alpha)-t, k_F(\alpha)+t \rangle.$$

According to the results of Dvoretzky, Kiefer and Wolfowitz (1956),

$$\begin{aligned}
& P\{\omega : F^N(k_F(\alpha)+t) > 1-\alpha), F^N(k_F(\alpha)-t) < 1-\alpha\} = \\
&= P\{\omega : F^N(k_F(\alpha)+t) > F(k_F(\alpha)), F^N(k_F(\alpha)-t) < F(k_F(\alpha))\} = \\
&= P\{\omega : F(k_F(\alpha)+t) - F^N(k_F(\alpha)+t) < F(k_F(\alpha)+t) - F(k_F(\alpha)), \\
&\qquad F^N(k_F(\alpha)-t) - F(k_F(\alpha)-t) < F(k_F(\alpha)) - F(k_F(\alpha)-t)\} \geq \\
&\geq P\{\omega : F(k_F(\alpha)+t) - F^N(k_F(\alpha)+t) < \vartheta t, \\
&\qquad F^N(k_F(\alpha)-t) - F(k_F(\alpha)-t) < \vartheta t)\} \geq \\
&\geq 1 - 2\exp\{-2N(\vartheta t)^2\},
\end{aligned}$$

we can see that the assertion of Lemma 5 is valid. $\square$