Kybernetika

Vlasta Kaňková

Empirical estimates in stochastic optimization via distribution tails

Kybernetika, Vol. 46 (2010), No. 3, 459--471

Persistent URL: http://dml.cz/dmlcz/140761

KYBERNETIKA — VOLUME 46 (2010), NUMBER 3, PAGES 459–471

EMPIRICAL ESTIMATES IN STOCHASTIC OPTIMIZATION VIA DISTRIBUTION TAILS

Vlasta Kaňková

"Classical" optimization problems depending on a probability measure belong mostly to nonlinear deterministic optimization problems that are, from the numerical point of view, relatively complicated. On the other hand, these problems fulfil very often assumptions giving a possibility to replace the "underlying" probability measure by an empirical one to obtain "good" empirical estimates of the optimal value and the optimal solution. Convergence rate of these estimates have been studied mostly for "underlying" probability measures with suitable (thin) tails. However, it is known that probability distributions with heavy tails better correspond to many economic problems. The paper focuses on distributions with finite first moments and heavy tails. The introduced assertions are based on the stability results corresponding to the Wasserstein metric with an "underlying" $\mathcal{L}_1$ norm and empirical quantiles convergence.

Keywords: stochastic programming problems, stability, Wasserstein metric, $\mathcal{L}_1$ norm, Lipschitz property, empirical estimates, convergence rate, exponential tails, heavy tails, Pareto distribution, risk functionals, empirical quantiles

Classification: 90C15

1. INTRODUCTION

To introduce a "classical" one-stage stochastic optimization problem let $(\Omega, \mathcal{S}, P)$ be a probability space; $\xi\,(:= \xi(\omega) = [\xi_1(\omega), \ldots, \xi_s(\omega)])$ s-dimensional random vector defined on $(\Omega, \mathcal{S}, P)$; $F(:= F(z),\, z \in \mathbb{R}^s)$ the distribution function of ξ; $F_i(:= F_i(z_i),\, z_i \in \mathbb{R}^1)$, $i = 1, \ldots, s$ one-dimensional marginal distribution functions corresponding to F; P_F, Z_F the probability measure and the support corresponding to F. Let, moreover, $g_0(:= g_0(x, z))$ be a real-valued (say continuous) function defined on $\mathbb{R}^n \times \mathbb{R}^s$; $X_F \subset \mathbb{R}^n$ be a nonempty set depending (generally) on F; $X \subset \mathbb{R}^n$ be a nonempty "deterministic" set not depending on F.

If the symbol E_F denotes the operator of mathematical expectation corresponding to F, then a rather general "classical" one-stage stochastic programming problem can be introduced in the form:

Find

$$\varphi(F) := \varphi(F, X_F) = \inf\{\mathrm{E}_F g_0(x, \xi) | x \in X_F\}. \tag{1.1}$$

Since in applications very often the measure P_F has to be replaced by empirical one, the solution of (1.1) has to be (mostly) sought w.r.t. an "empirical problem":

Find

$$\varphi(F_\omega^N) := \varphi(F_\omega^N,\, X_{F_\omega^N}) = \inf\{\mathrm{E}_{F_\omega^N} g_0(x,\,\xi) | x \in X_{F_\omega^N}\}. \tag{1.2}$$

F_ω^N denotes an empirical distribution function determined by a random sample $\{\xi^i\}_{i=1}^N$ (not necessarily independent) corresponding F. It is known that under rather general assumptions $\varphi(F_\omega^N)$ is a "good" estimate of $\varphi(F)$.

The investigation of these estimates started in 1974 (see [34]) and was followed by a "statistical" approach and the stability investigation e. g. in [2, 7, 22, 26, 31]. The investigation of the convergence rate started in [8] and follows e. g. in [1, 6, 19, 23, 29]. Let us recall the first result about the convergence rate.

Theorem 1.1. (Kaňková [8]) Let $t > 0$, X be a nonempty compact, convex set. If

1. $g_0(x,\, z)$ is a uniformly continuous, bounded function on $X \times Z_F$,
2. $g_0(x,\, z)$ is a Lipschitz function on X with the Lipschitz constant $L^{'}$,
3. $\{\xi^i\}_{i=1}^N$, $N = 1,\, 2,\, \ldots$ is an independent random sample corresponding to P_F,

then there exist $K(t,\, X,\, L^{'}),\; k_1(M) > 0, \quad (|(g_0(x,\, z)| \le M,\; M > 0)$ such that

$$P\{\omega : |\varphi(F,\, X) - \varphi(F_\omega^N,\, X)| > t\} \le K(t,\, X,\, L^{'}) \exp\{-N k_1(M) t^2\}.$$

Remarks.

1. Under the assumptions of Theorem 1.1 it has been proven that

$$P\{\omega : N^\beta |\varphi(F,\, X) - \varphi(F_\omega^N,\, X)| > t\} \underset{(N\to\infty)}{\longrightarrow} 0 \quad \text{for} \quad \beta \in (0,\, 1/2).$$

2. The assertion of Theorem 1.1 is valid independently of the distribution function F; consequently also for the distribution functions with heavy tails. On the other hand $g_0(\cdot,\, \cdot)$ must be a bounded function. This condition substitutes, evidently, the assumption on a bounded support of the corresponding random element in the Hoeffding paper [5].

3. L. Dal, C. H. Chen and J. R. Birge (see [1]) have tried to generalize the last assumption (for $s = 1$) to the case when

$$\mathrm{E}_F \exp\{\theta\xi\} < \infty \quad \text{for} \quad 0 \le \theta \le \theta_0, \quad \theta_0 \quad \text{constant}. \tag{1.3}$$

Evidently, the relation (1.3) can be fulfilled only for F with thin tails. Tito Homen-de-Mello (see [19]) has continued in the last direction.

The assumption of "thin" tails is not fulfilled in many applications. A relatively detailed analysis of heavy tailed distributions is presented in [20]. A relationship between the stable distributions and heavy tailed distributions can be found e. g. in [15]; between the stable heavy tailed distributions and the Pareto tails is known and can be found e. g. in [15] and [18] (see also [21]). Furthermore, it follows from the relation (1.1) that the assertion of Theorem 1.1 is valid for problems in which the objective function is in the form of a linear functional of probability measure P_F. This assumption is not fulfilled in many cases in which risk measures appear (for more details see e. g. [24]). We are intend to deal also with new above mentioned situations. We start our investigation with problems in which $X_F = X$ or when

$$X_F \,:=\, X_F(\delta) \,=\, \bigcap_{i=1}^{s} \{x \in X : P_{F_i}\{\omega : g_i(x) \le \xi_i(\omega)\} \ge \delta_i\}, \tag{1.4}$$

with $g_i(:= g_i(x))$, $i = 1, \ldots, s$ real-valued (say continuous) functions defined on $\mathbb{R}^n$, $\delta = (\delta_1, \ldots, \delta_s)$, $\delta_i \in (0, 1)$, $i = 1, \ldots, s$. To get the new results we plan to employ the stability assertion corresponding to the Wasserstein metric with an "underlying" $\mathcal{L}_1$ norm [12]. According to an analysis presented in [20], a transformation to one dimensional random element can be (from the economic point of view) very suitable.

Remark. The problem (1.1) with X_F fulfilling (1.4) is known as the problem with individual probabilistic constraints (corresponding to random right-hand sides). The problems with joint probability constraints cover more applications. It is known that we can approximate the joint probabilistic constraints by individual case (see e. g. [10, 25]) and, consequently, to transform them also to one dimensional case.

2. SOME DEFINITIONS AND AUXILIARY ASSERTIONS

First, let $\mathcal{P}(\mathbb{R}^s)$ denote the set of Borel probability measures on $\mathbb{R}^s$, $s \ge 1$. We set

$$\mathcal{M}_1(\mathbb{R}^s) = \left\{ P \in \mathcal{P}(\mathbb{R}^s) : \int_{\mathbb{R}^s} \|z\|_s^1 P(\mathrm{d}z) < \infty \right\}, \quad \|\cdot\|_s^1 \text{ denotes } \mathcal{L}_1 \text{ norm in } \mathbb{R}^s.$$

Let, furthermore, $k_F(\delta) = (k_{F_1}(\delta_1), \ldots, k_{F_s}(\delta_s))$, $\delta = (\delta_1, \ldots, \delta_s)$ be defined by

$$k_{F_i}(\delta_i) \,=\, \sup_{z_i \in \mathbb{R}^1} P\{\omega : z_i \le \xi_i(\omega)\} \ge \delta_i, \quad i = 1, \ldots, s. \tag{2.5}$$

We introduce the system of the assumptions:

A.1
- $g_0(x, z)$ is a uniformly continuous function on $X \times \mathbb{R}^s$,
- $g_0(x, z)$ is for $x \in X$ a Lipschitz function of $z \in \mathbb{R}^s$ with the Lipschitz constant L (corresponding to the $\mathcal{L}_1$ norm) not depending on x,

A.2
- $\{\xi^i\}_{i=1}^{\infty}$ is a sequence of independent random vectors corresponding to F,
- F_ω^N is an empirical distribution function determined by $\{\xi^i\}_{i=1}^{N}$, $N = 1, 2, \ldots,$

A.3 P_{F_i}, $i = 1, \ldots, s$ are absolutely continuous w.r.t. the Lebesgue measure on $\mathbb{R}^1$ (we denote by f_i, $i = 1, \ldots, s$ the probability densities corresponding to F_i),

A.4 there exist constants $\vartheta_i > 0$, $i = 1, \ldots, s$ and neighborhoods $U_i(k_{F_i}(\delta_i))$ of $k_{F_i}(\delta_i)$ such that $f_i(z_i) > \vartheta_i$ for $z_i \in U_i(k_{F_i}(\delta_i))$,

A.5 $\mathrm{E}_F g_0(x, \xi)$ is a Lipschitz function on X.

2.1. Stability assertions

Proposition 2.1. (Kaňková and Houda [12]) Let $P_F, P_G \in \mathcal{M}_1(\mathbb{R}^s)$, X be a compact set. If the assumption A.1 is fulfilled, then

$$|\varphi(F, X) - \varphi(G, X)| \le L \sum_{i=1}^{s} \int_{-\infty}^{+\infty} |F_i(z_i) - G_i(z_i)| \, \mathrm{d}z_i.$$

Employing the triangular inequality and (1.4) we can obtain

$$|\varphi(G, X_G) - \varphi(F, X_F)| \le |\varphi(G, X_G) - \varphi(F, X_G)| + |\varphi(F, X_G) - \varphi(F, X_F)|. \quad (2.6)$$

According to the relations (1.4) and (2.5) we can write

$$X_F := \bar{X}_F(k_F(\delta)) = \bigcap_{i=1}^{s} \{x \in X : \ g_i(x) \le k_{F_i}(\delta_i)\}. \quad (2.7)$$

If X is a compact set, g_i, $i = 1, \ldots, s$ continuous functions on X, then X_F, X_G are compact sets. Employing Proposition 2.1 we obtain the upper bounds for $|\varphi(G, X_G) - \varphi(F, X_G)|$. If, furthermore, $\Delta[\cdot, \cdot] = \Delta_n[\cdot, \cdot]$ denotes the Hausdorff distance in the space of nonempty, closed subsets of $\mathbb{R}^n$ (for the definition see e. g. [28]), then in [9] are introduced assumptions under which it is possible to evaluate $\bar{C} > 0$ such that

$$\Delta[X_F(\delta), X_G(\delta)] = \Delta[\bar{X}(k_F(\delta)), \bar{X}(k_G(\delta))] \le \bar{C} \sum_{i=1}^{s} |k_{F_i}(\delta_i) - k_{G_i}(\delta_i)|,$$
$$k_{G_i}(\delta_i) \in U_i(k_{F_i}(\delta_i)), \ i = 1, \ldots, s; \quad U_i(k_{F_i}(\delta_i)) \quad \text{defined by A.4.} \quad (2.8)$$

Consequently, it follows from Proposition 2.1, the relations (2.6), (2.8) that the upper bound of $|\varphi(F, X_F) - \varphi(G, X_G)|$ can be numerically evaluated. Moreover, this bound can be employed for a construction of solutions approximate schemes with known deterministic approximation error bound (similar approach has been already employed in [11, 32]).

2.2. Empirical estimates

Replacing G by F^N_ω we can investigate properties of the empirical estimates $\varphi(F^N_\omega)$.

Lemma 2.2. (Shorack and Wellner [33]) Let $s = 1$, $P_F \in \mathcal{M}_1(\mathbb{R}^1)$. Let, moreover, the assumption A.2 be fulfilled, then

$$P\left\{\omega : \int_{-\infty}^{\infty} |F(z) - F^N_\omega(z)| \, \mathrm{d}z \underset{(N \to \infty)}{\longrightarrow} 0 \right\} = 1.$$

Proposition 2.3. (Kaňková [13]) Let $s = 1,\, t > 0,$ the assumption A.3 be fulfilled. If there exists $\psi(N,\, t,\, R)$ such that the empirical distribution function F_ω^N fulfils for $R > 0$ the relation

$$P\{\omega :\, |F(z) - F_\omega^N(z)| > t\} \le \psi(N,\, t,\, R) \quad \text{for every} \quad z \in (-R,\, R),$$

then for $\frac{t}{4R} < 1$ it holds that

$$\begin{aligned} &P\left\{\omega :\, \int_{-\infty}^{\infty} |F(z) - F_\omega^N(z)|\,\mathrm{d}z \,>\, t\right\} \\ &\le \left(\frac{12R}{t} + 1\right)\psi(N,\, \frac{t}{12R},\, R) \,+\, P\left\{\omega :\, \int_{-\infty}^{-R} F(z)\,\mathrm{d}z > \frac{t}{3}\right\} \\ &+P\left\{\omega :\, \int_{R}^{\infty} (1 - F(z))\,\mathrm{d}z > \frac{t}{3}\right\} \,+\, 2NF(-R) \,+\, 2N(1 - F(R)). \end{aligned}$$

Corollary 2.4. Let $s = 1,\, t > 0,$ the assumptions A.2, A.3 be fulfilled. If there exists $\beta > 0,\, R := R(N) > 0$ defined on $\mathcal{N}$ such that $R(N) \underset{(N\to\infty)}{\longrightarrow} \infty$ and, moreover,

$$\begin{aligned} &N^\beta \int_{-\infty}^{-R(N)} F(z)\,\mathrm{d}z \underset{(N\to\infty)}{\longrightarrow} 0, \quad N^\beta \int_{R(N)}^{\infty} [1 - F(z)]\,\mathrm{d}z \underset{(N\to\infty)}{\longrightarrow} 0, \\ &2NF(-R(N)) \underset{(N\to\infty)}{\longrightarrow} 0, \qquad 2N[1 - F(R(N))] \underset{(N\to\infty)}{\longrightarrow} 0, \\ &\left(\frac{12N^\beta R(N)}{t} + 1\right) \exp\left\{-2N\left(\frac{t}{12R(N)N^\beta}\right)^2\right\} \underset{(N\to\infty)}{\longrightarrow} 0, \end{aligned} \tag{2.9}$$

then

$$P\{\omega :\, N^\beta \int_{-\infty}^{\infty} |F(z) - F^N(z)_\omega| > t\} \underset{(N\to\infty)}{\longrightarrow} 0.$$

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

P r o o f. Since it has been proven in [3] that for independent random sample

$$P\{\omega :\, |F(z) - F_\omega^N(z)| > t\} \le 2\exp\{-2Nt^2\} \quad \text{indepedently on} \quad z \in \mathbb{R}^1,$$

the assertion follows from the assertion of Proposition 2.3. □

Remark. According to the about mentioned inequality we can write (in the case of independent random sample) $\psi(N,\, t)$ instead of $\psi(N,\, t,\, R)$. Furthermore, employing the last inequality, the assertions of Proposition 2.1, Proposition 2.3, we can numerically (for every given F) evaluate $P\{\omega :\, |\varphi(F,\, X) - \varphi(F_\omega^N\, X)| > t\}$ for $N \in \mathcal{N},\, t > 0$.

Corollary 2.5. (Kaňková [13]) Let $s = 1$, $t > 0$, the assumptions A.2, A.3 be fulfilled. If there exists constants C_1, C_2 and $T > 0$ such that

$$f(z) \le C_1 \exp\{-C_2|z|\} \quad \text{for} \quad z \in (-\infty, -T) \cup (T, \infty),$$

then

$$P\{\omega : N^{\beta} \int_{-\infty}^{\infty} |F(z) - F_{\omega}^{N}(z)| > t\} \underset{(N \to \infty)}{\longrightarrow} 0 \quad \text{for} \quad \beta \in (0, 1/2).$$

To apply Corollary 2.4 to "heavy" tails, we recall the Pareto distribution.

Definition 2.6. Meerschaert [20]. A random variable $\xi(:= \xi(\omega))$ has a Pareto distribution if

$$P\{\omega : \xi > z\} = Cz^{-\alpha}, \qquad f(z) = \begin{array}{lll} C\alpha z^{-\alpha-1} & \text{for} & z > C^{\frac{1}{\alpha}}, \\ 0 & & z \le C^{\frac{1}{\alpha}}, \end{array} \tag{2.10}$$

where $C > 0$, $\alpha > 0$ are constants and $f(:= f(z))$ is a probability density.

The Pareto distribution has only one tail and for $\alpha > 1$ we obtain $P_F \in \mathcal{M}_1(\mathbb{R}^1)$.

Corollary 2.7. Let $s = 1$, $t > 0$, $\alpha > 1$ and $\beta, \gamma > 0$ fulfil the inequalities $\gamma > \frac{1}{\alpha}$, $\frac{\gamma}{\beta} > \frac{1}{\alpha - 1}$, $\gamma + \beta < \frac{1}{2}$. Let, moreover, the assumptions A.2, A.3 be fulfilled. If there exist constants $C > 0$, $T > 0$ such that

$$f(z) \le C\alpha|z|^{-\alpha-1} \quad \text{for} \quad z \in (-\infty, -T) \cup (T, \infty),$$

then

$$P\{\omega : N^{\beta} \int_{-\infty}^{\infty} |F(z) - F_{\omega}^{N}(z)| > t\} \underset{(N \to \infty)}{\longrightarrow} 0.$$

P r o o f. First, it follows from the assumptions for $z > T$, $R(N) = N^{\gamma}$, $R(N) > T$ that

$$N^{\beta} \int_{R(N)}^{\infty} [1 - F(z)] \, \mathrm{d}z \le N^{\beta} [C(-\alpha + 1) z^{-\alpha+1}]_{R(N)}^{\infty} = -C(-\alpha + 1) N^{\beta} N^{\gamma(-\alpha+1)},$$
$$N[1 - F(R(N))] \le NCN^{-\alpha\gamma} = CN^{1-\alpha\gamma}.$$

Setting $R(N) = N^{\gamma}$ we can see that the assertion follows from the assertion of Corollary 1, the last system of inequalities and the properties of the distribution functions. □

Examples. The following two cases of combinations of α, γ, β fulfil the assumptions of Corollary 3.

1. $\alpha = 3 + \varepsilon$, $\gamma = \frac{1}{3}$, $\beta = \frac{1}{6}$, $\varepsilon > 0$ arbitrary small,
2. $\alpha = 4 + \varepsilon$, $\gamma = \frac{1}{4}$, $\beta = \frac{1}{4}$, $\varepsilon > 0$ arbitrary small.

Lemma 2.8. Let $s = 1$, $t > 0$, $\beta \in (0, \frac{1}{2})$, $\delta \in (0, 1)$, A.2, A.3, A.4 be fulfilled, then

1. $P\{\omega : k_{F^N}(\delta) \underset{(N\to\infty)}{\longrightarrow} k_F(\delta)\} = 1,$

2. $P\{\omega : N^{\beta}|k_F(\delta) - k_{F^N}(\delta)| > t\} \underset{(N\to\infty)}{\longrightarrow} 0.$

Proof. The assertion of Lemma 2.8 follows from [30] (see also [14]). □

2.3. Bivariate Pareto distributions

A few definitions of slightly different univariate Pareto distributions exist in the literature. We recall the *Pareto(I)*(σ, α) distribution (introduced in [17]) that is very similar to the definition corresponding to the relation (2.10) ($C := \sigma^{\alpha}$).

Definition 2.9. (Kotz, Balakrishnan and Johnson [17]) The random value ξ is said to have a univariate *Pareto(I)*(σ, α) distribution if $P_{\alpha}\{\omega : \xi > z\} = \left(\frac{z}{\sigma}\right)^{-\alpha}$ for $z \ge \sigma$, $\sigma > 0$, $\alpha > 0$.

Mostly (in applications), a random element is represented by an s-dimensional random vector ($s > 1$). A bivariate and multivariate Pareto distributions corresponding to $P(I)(\sigma, \alpha)$ are introduced in [17]. We recall the bivariate case only.

Definition 2.10. (Kotz, Balakrishnan and Johnson [17]) The random two dimensional vector $\xi = (\xi_1, \xi_2)$ is said to have a bivariate Pareto distribution of the first kind if the joint probability density function $f_{\xi_1,\xi_2}(z_1, z_2)$ fulfil the relation

$$\begin{aligned} f_{\xi_1,\xi_2}(z_1, z_2) = \quad & (\alpha+1)\alpha(\theta_1\theta_2)^{\alpha+1}(\theta_2 z_1 + \theta_1 z_2 - \theta_1\theta_2)^{-(a+1)}, \\ & z_1 \ge \theta_1 > 0,\ z_2 \ge \theta_2 > 0,\ \alpha > 0. \end{aligned}$$

Evidently, the marginal densities are

$$f_{\xi_i}(z_i) = \alpha\theta_i^{\alpha} z_i^{-(\alpha+1)}, \quad z_i \ge \theta_i > 0, \quad i = 1, 2; \quad \text{consequently} \quad \xi_i =_d PI\left(\frac{1}{\theta_i}, \alpha\right).$$

Remark. A survey of Pareto distributions applications can be found in [20]. There exists also an analysis about an approach that α_i, $i = 1, \ldots, s$ are not necessary the same for all components.

3. MAIN RESULTS

3.1. Consistency

Theorem 3.1. Let X be a compact set, the assumptions A.1 and A.2 be fulfilled. If $P_F \in \mathcal{M}_1(\mathbb{R}^s)$, then

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

If, moreover, the assumptions A.3, A.4, A.5 and the relation (2.8) are fulfilled, then also

$$P\{\omega : |\varphi(F, X_F) - \varphi(F_\omega^N, X_{F_\omega^N})| \underset{(N\to\infty)}{\longrightarrow} 0\} = 1.$$

Proof. The assertion follows from (2.6), Proposition 2.1, Lemma 2.2 and Lemma 2.8. □

Remark. According to the fact that $P_F \in \mathcal{M}_1(\mathbb{R}^s)$ for many stable (for definition see e. g. [15]) and Pareto distributions we can see that $\varphi(F_\omega^N)$ is a consistent estimate of $\varphi(F)$ for many heavy tails distributions.

3.2. Convergence rate

Theorem 3.2. (Kaňková [13]) Let $t > 0$, X be a compact set, the assumptions A.1, A.2 and A.3 be fulfilled. If there exist constants $C_1, C_2 > 0$ and $T > 0$ such that

$$f_i(z_i) \le C_1 \exp\{-C_2|z_i|\} \quad \text{for} \quad z_i \in (-\infty, -T) \cup (T, \infty), \quad i = 1, \dots, s, \tag{3.11}$$

then

$$P\{\omega :\, N^\beta|\varphi(F, X) - \varphi(F_\omega^N, X)| > t\} \underset{(N\to\infty)}{\longrightarrow} 0 \quad \text{for } \beta \in (0, 1/2).$$

If, moreover, the assumptions A.4, A.5 and (2.8) are fulfilled, then also

$$P\{\omega :\, N^\beta|\varphi(F, X_F) - \varphi(F_\omega^N, X_{F_\omega^N})| > t\} \underset{(N\to\infty)}{\longrightarrow} 0 \quad \text{for } \beta \in (0, 1/2).$$

Theorem 3.3. Let $t > 0$, X be a compact set, $C > 0$, $\alpha_i > 1$, $i = 1, \dots, s$, the assumptions A.1, A.2 and A.3 be fulfilled. If

1. there exists a constant $T > 0$ such that

$$f_i(z) \le C\alpha_i |z_i|^{-\alpha_i - 1} \quad \text{for} \quad z_i \in (-\infty, -T) \cup (T, \infty), \quad i = 1, \dots, s, \tag{3.12}$$

2. $\alpha_i > 1$, $\gamma_i > 0$, $i = 1, \dots, s$, $\beta > 0$ fulfil the inequalities

$$\gamma_i > \frac{1}{\alpha_i}, \quad \frac{\gamma_i}{\beta} > \frac{1}{\alpha_i - 1}, \quad \gamma_i + \beta < \frac{1}{2},$$

then

$$P\{\omega : N^{\beta}|\varphi(F_{\omega}^{N}, X) - \varphi(F, X)| > t\} \underset{(N\to\infty)}{\longrightarrow} 0.$$

If, moreover, the assumptions A.4, A.5 and (2.8) are fulfilled, then also

$$P\{\omega : N^{\beta}|\varphi(F, X_F) - \varphi(F_{\omega}^{N}, X_{F_{\omega}^{N}})| > t\} \underset{(N\to\infty)}{\longrightarrow} 0.$$

Proof. Evidently, under the assumptions $P_F \in \mathcal{M}_1(\mathbb{R}^s)$. The first assertion of Theorem 3.3 follows from Proposition 2.1 and Corollary 2.7. The second assertion follows from the first one and the relation (2.6), (2.8). □

4. APPLICATION TO PORTFOLIO SELECTION

Heavy tails distributions are applied also to assets theory (see e.g. [20]). Moreover, it follows e.g. from [16, 23, 24] that risk measures are not necessary a linear "functional" of the probability measure. Consequently, new types of optimization problems arise. To explain this fact, we start with a classical portfolio problem:

Find

$$\max_{x\in X}\sum_{k=1}^{n}\xi_k x_k, \quad X = \left\{x \in \mathbb{R}^n : \sum_{k=1}^{n} x_k \le 1, \quad x_k \ge 0, \quad k = 1, \ldots, n\right\}, \quad s = n,$$

where x_k is a fraction of the unit wealth invested in the asset k, $\xi_k x_k$ denotes the return of the value x_k invested in the asset $k \in \{1, 2, \ldots n\}$. If $\xi_k, k = 1, \ldots, n$ are known, then the last problem is a linear programming problem. Since $\xi_k, k = 1, \ldots, n$ are mostly random variables with unknown realizations in a time decision, it is reasonable to set to the portfolio selection two-objective optimization problem:

Find

$$\max\sum_{k=1}^{n}\mu_k x_k, \quad \min\sum_{k=1}^{n}\sum_{j=1}^{n} x_k c_{k,j} x_j \quad \text{subject to} \quad x = (x_1, \ldots, x_n) \in X, \tag{4.13}$$

where $\mu_k = \mathrm{E}_F\xi_k$, $c_{k,j} = \mathrm{E}_F(\xi_k - \mu_k)(\xi_j - \mu_j)$, $k, j = 1, \ldots n$. Markowitz sets to the problem (4.13) the following one-objective problem:

Find

$$\max\left[\sum_{k=1}^{n}\mu_k x_k - K\sum_{k=1}^{n}\sum_{j=1}^{n} x_k c_{k,j} x_j\right] \quad \text{subject to} \quad x \in X; \quad K > 0 \quad \text{is a constant.} \tag{4.14}$$

Evidently, $\sigma^2(x) = \sum_{k=1}^{n}\sum_{j=1}^{n} x_k c_{k,j} x_j = \mathrm{E}_F\{\sum_{j=1}^{n}\xi_j x_j - \mathrm{E}_F[\sum_{j=1}^{n}\xi_j x_j]\}^2$ can be considered as a risk measure, that can be (see [16]) replaced by

$$\begin{aligned} w(x) &= \mathrm{E}_F\left|\sum_{k=1}^{n}\xi_k x_k - \mathrm{E}_F\left[\sum_{k=1}^{n}\xi_k x_k\right]\right|, \\ w^{+}(x) &= \mathrm{E}_F\left|\sum_{k=1}^{n}\xi_k x_k - \mathrm{E}_F\left[\sum_{k=1}^{n}\xi_k x_k\right]\right|^{+}, \\ w^{-}(x) &= \mathrm{E}_F\left|\sum_{k=1}^{n}\xi_k x_k - \mathrm{E}_F\left[\sum_{k=1}^{n}\xi_k x_k\right]\right|^{-}. \end{aligned} \tag{4.15}$$

Replacing in (4.14) $\sigma^2(x)$ by $w(x)$, $w^+(x)$ and $w^-(x)$ we obtain the problems:
Find

$$\begin{aligned}
\varphi^1(F) &:= \max_{x\in X} \Big[\textstyle\sum_{k=1}^n \mu_k x_k - K \mathrm{E}_F \Big| \sum_{k=1}^n \xi_k x_k - \mathrm{E}_F \Big[\sum_{k=1}^n \xi_k x_k \Big]\Big|\Big],\\
\varphi^2(F) &:= \max_{x\in X} \Big[\textstyle\sum_{k=1}^n \mu_k x_k - K \mathrm{E}_F \Big| \sum_{k=1}^n \xi_k x_k - \mathrm{E}_F \Big[\sum_{k=1}^n \xi_k x_k \Big]\Big|^+\Big],\\
\varphi^3(F) &:= \max_{x\in X} \Big[\textstyle\sum_{k=1}^n \mu_k x_k - K \mathrm{E}_F \Big| \sum_{k=1}^n \xi_k x_k - \mathrm{E}_F \Big[\sum_{k=1}^n \xi_k x_k \Big]\Big|^-\Big].
\end{aligned} \tag{4.16}$$

Evidently the problems (4.16) are covered by a more general problem:
Find

$$\varphi(F) := \bar{\varphi}(F,\, X) = \inf \left\{ \mathrm{E}_F g_0^1(x,\, \xi, \mathrm{E}_F h(x,\, \xi)) | x \in X \right\}, \tag{4.17}$$

where $h(x,\, z) = (h_1(x,\, z),\, \ldots,\, h_{m_1}(x,\, z))$ generally can be an m_1-dimensional vector function defined on $X \times \mathbb{R}^s$, $g_0^1(x,\, z,\, y)$ is a real valued function defined on $X \times \mathbb{R}^s \times Y$, $Y \subset \mathbb{R}^{m_1}$ nonempty set. $\mathrm{E}_F[\sum_{k=1}^n \xi_k x_k]$ corresponds in (4.16) to $\mathrm{E}_F h(x,\, \xi)$. Furthermore employing the approach of [23] we can this general case transform to the case of Theorem 3.2 or Theorem 3.3 (for more details see [13]).

Proposition 4.1. (Kaňková [13]) Let X be a compact set, G be an arbitrary s-dimensional distribution function. Let, moreover, P_F, $P_G \in \mathcal{M}_1(\mathbb{R}^s)$. If

1. $g_0^1(x,\, z,\, y)$ is for $x \in X$, $z \in \mathbb{R}^s$ a Lipschitz function of $y \in Y$ with a Lipschitz constant L^y; $Y = \{y \in \mathbb{R}^{m_1} : y = h(x,\, z) \quad \text{for some} \quad x \in X,\, z \in \mathbb{R}^s\}$,
2. for every $x \in X$, $y \in Y$ there exist finite mathematical expectations

   $\mathrm{E}_F g_0^1(x,\, \xi,\, \mathrm{E}_F h(x,\, \xi))$, $\quad \mathrm{E}_F g_0^1(x,\, \xi,\, \mathrm{E}_G h(x,\, \xi))$, $\quad \mathrm{E}_G g_0^1(x,\, \xi,\, \mathsf{G}_F h(x,\, \xi))$,
3. $h_i(x,\, z)$, $i = 1,\, \ldots,\, m_1$ are for every $x \in X$ Lipschitz functions of z with the Lipschitz constants L_h^i (corresponding to $\mathcal{L}_1$ norm),
4. $g_0^1(x,\, z,\, y)$ is for every $x \in X$, $y \in \mathbb{R}^{m_1}$ a Lipschitz function of $z \in \mathbb{R}^s$ with the Lipschitz constant L^z (corresponding to $\mathcal{L}_1$ norm),

then there exists $\hat{C}$ such that

$$|\bar{\varphi}(F,\, X) - \bar{\varphi}(G,\, X)| \le \hat{C} \sum_{i=1}^s \int_{-\infty}^{\infty} |F_i(z_i) - G_i(z_i)| \, \mathrm{d}z_i. \tag{4.18}$$

Replacing G by empirical one F_ω^N we obtain "empirical problems":
Find

$$\begin{aligned}
\varphi^2(F_\omega^N) &= \max_{x\in X} \left[\mathrm{E}_{F_\omega^N} \left[\textstyle\sum_{k=1}^n \xi_k x_k \right] - K \mathrm{E}_{F_\omega^N} \left| \sum_{k=1}^n \xi_k x_k - \mathrm{E}_{F_\omega^N} \left[\sum_{k=1}^n \xi_k x_k \right] \right| \right],\\
\varphi^3(F_\omega^N) &= \max_{x\in X} \left[\mathrm{E}_{F_\omega^N} \left[\textstyle\sum_{k=1}^n \xi_k x_k \right] - K \mathrm{E}_{F_\omega^N} \left| \sum_{k=1}^n \xi_k x_k - \mathrm{E}_{F_\omega^N} \left[\sum_{k=1}^n \xi_k x_k \right] \right|^+ \right],\\
\varphi^4(F_\omega^N) &= \max_{x\in X} \left[\mathrm{E}_{F_\omega^N} \left[\textstyle\sum_{k=1}^n \mu_k x_k \right] - K \mathrm{E}_{F_\omega^N} \left| \sum_{k=1}^n \xi_k x_k - \mathrm{E}_{F_\omega^N} \left[\sum_{k=1}^n \xi_k x_k \right] \right|^- \right].
\end{aligned} \tag{4.19}$$

Employing furthermore the technique of the Theorem 3.2, Theorem 3.3 and Proposition 4.1 proofs we can see that

$$P\left\{\omega : N^{\beta}|\varphi^i(F) - \varphi^i(F^N_\omega)| > t\right\} \underset{(N\to\infty)}{\longrightarrow} 0, \quad i = 1, 2, 3, \tag{4.20}$$

where the value of coefficient β is determined by the relations (3.11) or (3.12).

5. DISCUSSION

The paper deals with stability and empirical estimates of the optimal value in stochastic programming problems. In particular, the aim of the paper is to focus on heavy tails and Pareto distributions. The presented results are based on the stability assertions based on the Wasserstein metric corresponding to $\mathcal{L}_1$ norm. These stability results are obtained under the assumptions of compact feasible set, continuity of the "underlying" objective functions, existence of probability density and so on. Consequently, the assumptions are rather strong, however on the other hand this approach enables to evaluate numerically approximations of upper bounds in the case of deterministic approximative solutions schemes (for details see e. g. [11] or [32]), as well as the probability of the Monte Carlo error in the empirical approximations. At the end of the paper, the reported results are applied to some nonlinear (w.r.t. probability measure) functionals. To obtain the results for optimal solution some growth assumptions can be assumed (see e. g. [27]). However, this investigation is beyond the scope of this paper.

More stronger theoretical assertions (under weaker assumptions, i. e. without the assumptions of compactness of the feasible set, continuity of objective function, only individual probability constraints, absolutely continuous probability measure w.r.t. Lebesque measure and so on) are known from the stochastic programming literature (see e. g. [4] or [27]). These papers also present results concerning the optimal solutions. However, all these results are based on essentially more general "complicated" theoretical probability metrics. These abstract assertions are of great value from the theoretical point of view, and special cases can be sometimes obtained from them. Our results only try to complete them for possibilities of the numerical employment and approximation in new arising economic problems.

ACKNOWLEDGEMENT

The author would like to thank an anonymous referee for many helpful comments.

This work was partially supported by the Czech Science Foundation under Grants 402/07/1113, 402/08/0107 and the Grant Agency of Ministry of Youth, Education and Sport of the Czech Republic under grant LC 06075.

(Received April 14, 2010)

REFERENCES

[1] L. Dai, C. H. Chen, and J. R. Birge: Convergence properties of two-stage stochastic programming. J. Optim. Theory Appl. *106* (2000), 489–509.

[2] J. Dupačová and R. J.-B. Wets: Asymptotic behaviour of statistical estimates and optimal solutions of stochastic optimization problems. Ann. Statist. *16* (1984), 1517–1549.

[3] A. Dvoretzky, J. Kiefer, and J. Wolfowitz: Asymptotic minimax character of the sample distribution function and the classical multinomial estimate. Ann. Math. Statist. *56* (1956). 642–669.

[4] R. Henrion and W. Römisch: Metric regularity and quantitative stability in stochastic programs with probability constraints. Math. Programming *84* (1999), 55–88.

[5] W. Hoeffding: Probability inequalities for sums of bounded random variables. J. Amer. Statist. Assoc. *58* (1963), 301, 13–30.

[6] Y. M. Kaniovski, A. J. King, and R. J.-B. Wets: Probabilistic bounds (via large deviations) for the solutions of stochastic programming problems. Ann. Oper. Res. *56* (1995), 189–208.

[7] V. Kaňková: Optimum solution of a stochastic optimization problem with unknown parameters. In: Trans. Seventh Prague Conference, Academia, Prague 1977, pp. 239–244.

[8] V. Kaňková: An approximative solution of stochastic optimization problem. In: Trans. Eighth Prague Conference, Academia, Prague 1978, pp. 349–353.

[9] V. Kaňková: On the stability in stochastic programming: the case of individual probability constraints. Kybernetika *33* (1997), 5, 525–546.

[10] V. Kaňková: Unemployment problem, restructuralization and stochastic programming. In: Proc. Mathematical Methods in Economics 1999 (J. Plešingr, ed.), Czech Society for Operations Research and University of Economics Prague, Jindřichův Hradec, pp. 151–158.

[11] V. Kaňková and M. Šmíd: On approximation in multistage stochastic programs: Markov dependence. Kybernetika *40* (2004), 5, 625–638.

[12] V. Kaňková and M. Houda: Empirical estimates in stochastic programming. In: Proc. Prague Stochastics 2006 (M. Hušková and M. Janžura, eds.), Matfyzpress, Prague 2006, pp. 426–436.

[13] V. Kaňková: Empirical Estimates via Stability in Stochastic Programming. Research Report ÚTIA AV ČR No. 2192, Prague 2007.

[14] V. Kaňková: Multistage stochastic programs via autoregressive sequences and individual probability constraints. Kybernetika *44* (2008), 2, 151–170.

[15] L. B. Klebanov: Heavy Tailed Distributions. Matfyzpress, Prague 2003.

[16] H. Konno and H. Yamazaki: Mean-absolute deviation portfolio optimization model and its application to Tokyo stock markt. Management Sci. *37* (1991), 5, 519–531.

[17] S. Kotz, N. Balakrishnan, and N. L. Johnson: Continuous Multiviariate Distributions. Volume 1: Models and Applications. Wiley, New York 2000.

[18] T. J. Kozubowski, A. K. Panorska, and S. T. Rachev: Statistical issues in modeling stable portfolios. In: Handbook of Heavy Tailed Distributions in Finance (S. T. Rachev, ed.), Elsevier, Amsterdam 2003, pp. 131–168.

[19] T. Homen de Mello: On rates of convergence for stochastic optimization problems under non-i.i.d. sampling. SIAM J. Optim. *19* (2009), 2, 524–551.

[20] M. M. Meerschaert and H.-P. Scheffler: Portfolio modeling with heavy tailed random vectors. In: Handbook of Heavy Tailed Distributions in Finance (S. T. Rachev, ed.), Elsevier, Amsterdam 2003, pp. 595–640.

[21] V. Omelchenko: Stable Distributions and Application to Finance. Diploma Thesis (supervisor L. Klebanov), Faculty of Mathematics and Physics, Charles University Prague, Prague 2007.

[22] G. Ch. Pflug: Scenario tree generation for multiperiod financial optimization by optimal discretization. Math. Program. Ser. B *89* (2001), 251–271.

[23] G. Ch. Pflug: Stochastic optimization and statistical inference. In: Stochastic Programming (Handbooks in Operations Research and Management Science, Vol. 10, A. Ruszczynski and A. A. Shapiro, eds.), Elsevier, Amsterdam 2003, pp. 427–480.

[24] G. Ch. Pflug and W. Römisch: Modeling Measuring and Managing Risk. World Scientific Publishing Co. Pte. Ltd, New Jersey, 2007.

[25] A. Prékopa: Probabilistic programming. In: Stochastic Programming, (Handbooks in Operations Research and Managemennt Science, Vol. 10, (A. Ruszczynski and A. A. Shapiro, eds.), Elsevier, Amsterdam 2003, pp. 267–352.

[26] W. Römisch and R. Schulz: Stability of solutions for stochastic programs with complete recourse. Math. Oper. Res. *18* (1993), 590–609.

[27] W. Römisch: Stability of stochastic programming problems. In: Stochastic Programming, Handbooks in Operations Research and Managemennt Science, Vol 10 (A. Ruszczynski and A. A. Shapiro, eds.), Elsevier, Amsterdam 2003, pp. 483–554.

[28] G. Salinetti and R. J. B. Wets: On the convergence of closed-valued measurable multifunctions. Trans. Amer. Math. Soc. *266* (1981), 1, 275–289.

[29] R. Schulz: Rates of convergence in stochastic programs with complete integer recourse. SIAM J. Optim. *6* (1996), 4, 1138–1152.

[30] J. R. Serfling: Approximation Theorems of Mathematical Statistics. Wiley, New York 1980.

[31] A. Shapiro: Quantitative stability in stochastic programming. Math. Program. *67* (1994), 99–108.

[32] M. Šmíd: The expected loss in the discretization of multistage stochastic programming problems-estimation and convergence rate. Ann. Oper. Res. *165* (2009), 29–45.

[33] G. R. Shorack and J. A. Wellner: Empirical Processes and Applications to Statistics. Wiley, New York 1986.

[34] R. J.-B. Wets: A Statistical Approach to the Solution of Stochastic Programs with (Convex) Simple Recourse. Research Report, University Kentucky, USA 1974.

Vlasta Kaňková, Institute of Information Theory and Automation – Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 4, 182 08 Praha 8. Czech Republic.
e-mail: kankova@utia.cas.cz