# Empirical Estimates in Optimization Problems: Survey with Special Regard to Heavy Tails and Dependent Sample *

Vlasta Kaňková

*Academy of Sciences of the Czech Republic, Institute of Information Theory and Automation, Pod Vodárenskou věží 4, Praha 8, CZ-182 08, Czech Republic.*

*e-mail: kankova@utia.cas.cz*

**Abstract.** Economic processes are usually influenced simultaneously by a random factor and a decision parameter. Since the decision parameter has to be mostly determined before realization of the random element, deterministic optimization problems which depend on a probability measure often correspond to the above mentioned situations. A complete knowledge of the "underlying" measure would be a necessary assumption to determine both an exact optimal solution and an exact optimal value. Since this condition is not usually fulfilled, the solution is often determined on an empirical data base. Corresponding estimates can only be obtained using this approach.

Many efforts have been made to investigate the above mentioned estimates. The consistency, convergence rate and an asymptotic distribution have been examined. This was mostly done under assumptions of linear dependence on the probability measure, distributions with "thin" tails and an assumption of independent data. The aim of this paper is to consider the cases in which these assumptions are rather relaxed. To this end we employ stability results based on the Wasserstein metric corresponding to $\mathcal{L}_1$ norm and some results on mixing sequences.

**Keywords:** Stochastic optimization, empirical estimates, Wasserstein metric, $\mathcal{L}_1$ norm, Lipschitz property, consistency, convergence rate, thin and heavy tails, stable distributions, shape parameter, independent and weak dependent random samples

**JEL classification:** C44
**AMS classification:** 90C15

## 1. Introduction

To introduce a "classical" one–stage stochastic programming 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 $(\Omega, \mathcal{S}, P)$; $F := F(z)$ where $z \in \mathbb{R}^s$, $P_F$ and $Z_F$ denote the distribution function, the probability measure and the support corresponding to $\xi$. Let, moreover, $g_0 := g_0(x, z)$ be a real–valued (say, continuous) function defined on $\mathbb{R}^n \times \mathbb{R}^s$;

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

$X_F \subset X \subset \mathbb{R}^n$ be a nonempty set generally depending on $F$, and $X \subset \mathbb{R}^n$ be a nonempty deterministic set. If the symbol $\mathbb{E}_F$ denotes the operator of mathematical expectation corresponding to $F$ and if for every $x \in X_F$ there exists a finite $\mathbb{E}_F g_0(x, \xi)$, then a rather general "classical" one–stage stochastic programming problem can be introduced in the form:

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

In real–life models it is often necessary to replace the measure $P_F$ by its stochastic estimate to obtain at least an approximate optimal value and an optimal solution. An empirical probability measure is a very suitable candidate for this estimate. Consequently, the solution to Problem (1) often has to be sought with respect to an empirical problem:

$$\text{Find} \quad \varphi(F^N,\, X_{F^N}) = \inf\{\mathbb{E}_{F^N} g_0(x, \xi) \,|\, x \in X_{F^N}\}, \tag{2}$$

where $F^N$ denotes an empirical distribution function determined by a random sample $\{\xi^i\}_{i=1}^N$ (not necessarily independent) corresponding to the distribution function $F$. If we denote the optimal solutions sets of (1) and (2) by $\mathcal{X}(F,\, X_F), \mathcal{X}(F^N,\, X_{F^N})$, then (under rather general assumptions) $\varphi(F^N,\, X_{F^N}), \mathcal{X}(F^N,\, X_{F^N})$ are "good" stochastic estimates of $\varphi(F,\, X_F)$, $\mathcal{X}(F,\, X_F)$. The properties of the above–mentioned estimates have been investigated many times in stochastic programming literature. It was shown that these estimates are consistent under rather general assumptions. The convergence rate and asymptotic distributions have been studied as well. It has been proven that the corresponding estimates usually have very "pleasant" properties. However, these results have mostly been obtained for "underlying" "classical" situations. This means, under the assumptions of the distribution with "thin" tails, random data corresponding to independent samples and the problems depending linearly on the probability measure. More recently it has been recognized that random elements corresponding to economic and financial situations do not fulfill these conditions everywhere. Consequently, a question has arisen if the above mentioned estimates then also have "acceptable" properties. The aim of this paper is to recall some results obtained also under the assumptions of the distributions with "heavy" tails (including the stable distributions) and random data which fulfill mixing conditions.

The paper is organized as follows. First, we try to give a brief survey of corresponding historical results and approaches (section 2). Section 3 is devoted to necessary definitions and some auxiliary assertions. The essential results are presented in Section 4 (starting with a survey of heavy tails in economic and financial applications, and followed by mathematical results). Finally, Section 5 deals with dependent random samples in empirical estimates.

## 2. Brief Survey

The investigation of the empirical estimates started in [50] in 1974; followed by many works (see, e. g., [6], [13], [16], [35], [37], [41], [47]). Let us recall some of these well–known results. We emphasize that the first results were obtained mostly under the assumption $X_F = X$.

**Theorem 1.** *[13] If*

1. *$X$ is a compact set, $g_0(x, z)$ is a uniformly continuous bounded function on $\mathbb{R}^s \times X$,*

2. *$\{\xi^i\}_{i=-\infty}^{\infty}$ is an ergodic sequence,*

*then*

$$P\{\omega : |\varphi(F^N, X) - \varphi(F, X)| \xrightarrow[N\to\infty]{} 0\} = 1.$$

*Remark 2.* Theorem 1 was proven under the assumption that $\{\xi^i\}_{i=1}^{\infty}$ is an ergodic sequence (for the definition of the ergodic random sequence see, e.g. [2]). Of course the ergodic property covers an independent random sample.

The results on consistency have many times been generalized. Later on we shall see that the consistency results are guaranteed (under some additional assumptions) by the finite first moments of the random elements. But now let us recall one of the well–known result concerning this problem.

**Theorem 3.** *Let $X$ be a nonempty compact set. If*

1. *for every $x \in X$ the function $g_0(x, z)$ is a continuous function of $x$ for almost every $z \in Z_F$ (w.r.t. $P_F$),*

2. *$g_0(x, z), x \in X$ is dominated by an integrable (w.r.t. $F$) function,*

3. *$\{\xi^i\}_{i=1}^{N}, N = 1, 2, \ldots$ is an independent random sample,*

*then*

$$P\{\omega : |\varphi(F^N, X) - \varphi(F, X)| \xrightarrow[N\to\infty]{} 0\} = 1.$$

*Proof.* The assertion of Theorem 3 follows immediately from Proposition 5.2 and Theorem 7.48 proven in [39].

Investigation of the convergence rate began back in 1978 [14] by this assertion:

**Theorem 4.** *[14] Let $t > 0$, $X$ be a nonempty compact, convex set. If*

1. *$g_0(x, z)$ is a uniformly continuous function on $X \times Z_F$, bounded by $M' > 0$ (i. e., $|g_0(x, z)| \leq M'$),*

2. *$g_0(x, z)$ is a Lipschitz function on $X$ with the Lipschitz constant $L'$ not depending on $z$,*

3. *$\{\xi^i\}_{i=1}^{N},\ N = 1, 2, \ldots$ is an independent random sample,*

*then there exist constants $K(t, X, L')$, $k_1(M') > 0$ such that*

$$P\{\omega : |\varphi(F, X) - \varphi(F^N, X)| > t\} \leq K(t, X, L') \exp\{-Nk_1(M')t^2\}.$$

*Remark 5.* $K(t, X, L')$ depends on $t, X, L^{'}$ and $k_1(M^{'})$ on $M^{'}$. Employing their estimates presented in [14] it has been proven in [18] that

$$P\{\omega : N^{\beta}|\varphi(F,\, X) - \varphi(F^N,\, X)| > t\} \xrightarrow[N \longrightarrow \infty]{} 0 \qquad \text{for} \qquad \beta \in (0, \tfrac{1}{2}).$$

Moreover, if $g_0(x,\, z)$ is a uniformly strongly convex function of $x \in X$ with a parameter $\rho > 0$, then $\mathcal{X}(F,\, X)$ and $\mathcal{X}(F^N,\, X)$ are singletons and

$$P\{\omega :\, N^{\beta}\|\mathcal{X}(F,\, X) - \mathcal{X}(F^N,\, X)\|^2 > t\} \xrightarrow[N \longrightarrow \infty]{} 0 \qquad \text{for} \qquad \beta \in (0, \tfrac{1}{2}).$$

Recall that $g_0(x,\, z)$ is a (uniformly) strongly convex function on convex set $X$ if there exists a constant $\rho > 0$ such that the relation

$$g_0(x,\, z) \le \lambda g_0(x^1,\, z) + (1-\lambda) g_0(x^2,\, z) - \lambda(1-\lambda)\rho\|x^1 - x^2\|^2$$

is valid for every $\lambda \in \langle 0,\, 1\rangle$, $x = \lambda x^1 + (1-\lambda)x^2$, $x^1,\, x^2 \in X$, $z \in \mathbb{R}^s$ (for more details see e.g. [34]; $\|\cdot\| = \|\cdot\|_n^2$ denotes the Euclidean norm in $\mathbb{R}^n$).

The assertion of Theorem 4 is valid independently of distribution function $F$, and consequently it is true for distribution functions with heavy tails as well. On the other hand, $g_0$ must be a bounded function. This condition substitutes, evidently, the assumption on a bounded support of the random element in the Hoeffding paper [8]. Today the assumptions of Theorem 1 and Theorem 3 seems very strong. However we can mention that results very similar to Theorem 4 have been proven by A. Tsybakov in 1981 (see [45]). More recently, the theory of large deviations has been employed to investigate the convergence rate (see, e.g., [3], [12]). If the moment–generating function $M_{g_0}(u)$, corresponding to $g_0(x, \xi)$ is defined by the relation

$$M_{g_0}(u) := \mathbb{E}_F\{e^{u[g_0(x,\xi) - \mathbb{E}_F g_0(x,\xi)]}\}, \quad u \in \mathbb{R}^1,$$

then the following assertion was presented in [38].

**Theorem 6.** *[38] Let $X \subset \mathbb{R}^n$ be a nonempty compact set. If*

1. *for every $x \in X$ the moment generating function $M_{g_0}(u)$ is finite valued for all $u$ in a neighbourhood of zero,*
2. *there exists a measurable function $\kappa : Z_F \to \mathbb{R}_+$ and a constant $\gamma^{'} > 0$ such that*
$$|g_0(x', z) - g_0(x, z)| \le \kappa(z)\|x' - x\|^{\gamma^{'}}$$
*for all $z \in Z_F$ and all $x, x' \in X$,*
3. *the moment generating function $M_{\kappa}(u)$ of $\kappa(\xi)$ is finite valued for all $u$ in a neighbourhood of zero,*
4. *$\{\xi^i\}_{i=1}^N$, $N = 1, 2, \ldots$ is an independent random sample,*

*then for any $t > 0$ there exist positive constants $\bar{C} := \bar{C}(t)$ and $\bar{\beta} := \bar{\beta}(t)$, independent of $N$, such that*

$$P\{\sup_{x \in X} |\mathbb{E}_{F^N} g_0(x, \xi) - \mathbb{E}_F g_0(x, \xi)| \ge t\} \le \bar{C}(t) e^{-N\bar{\beta}(t)}.$$

*Remark 7.* In Theorem 6 it is assumed that the function $g_0$ is Lipschitz with respect to the decision parameter and, moreover, the Lipschitz constant can depend on the random element. The upper bound is exponential; however since the types of the functions $\bar{C}$, $\bar{\beta}$ are not specified, the results of Theorem 6 are not suitable for the determination of the convergence rate.

The stability results and the large deviations technique have been employed to investigate convergence rate of empirical estimates. First, the Kolmogorov metric was employed (see, e.g., [15]). Later, e.g., the Lipschitz metric, Wasserstein metric based on the Euclidean norm, and Fortet–Mourier metric were employed (for the definitions see e.g., [31]). The special metric depending on the problem (with mathematical expectation generally also constrainted) has been constructed; see e.g., [36]. A relationship between convergence rate and the "underlying" probability measure was investigated therein.

The recalled approaches have been employed mostly for independent data. However, results admitting some types of dependence have appeared in the literature. The investigation of estimates corresponding to the dependent data also began in 1990s (see, e.g., [17], [48]). Since then others approaches have appeared (see, e.g. [26], [49]).

## 3. Some Definitions and Auxiliary Assertions

We introduce the essential system of the assumptions:

A.1
- $g_0(x, z)$ is either a uniformly continuous function on $X \times \mathbb{R}^s$ or there exists $\varepsilon > 0$ such that $g_o(x, z)$ is a convex bounded function on $X(\varepsilon)$ ($X(\varepsilon)$ denotes the $\varepsilon$–neighborhood of the set$X$),
- $g_0(x, z)$ 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 an independent random sequence corresponding to $F$ (we denote by symbol $F^N$ an empirical distribution function determined by $\{\xi^i\}_{i=1}^{N}, N = 1, 2, \dots$),

A.3 $P_{F_i}, i = 1, \dots, s$ are absolutely continuous w. r. t. the Lebesgue measure on $\mathbb{R}^1$ ($F_i$, $P_{F_i}$, $i = 1, 2, \dots, s$ denote one–dimensional marginal distribution functions and probability measures corresponding to $F$),

A.4 For every $i \in \{1, \dots, s\}$ there exist $\delta > 0$ and $\vartheta > 0$ such that $f_i(z_i) > \vartheta$ for $z_i \in Z_{F_i}$, $|z_i - k_{F_i}(\alpha_i)| < 2\delta$ ( $k_{F_i}(\alpha_i) = \sup\{z_i : P_{F_i}\{\omega : z_i \le \xi_i(\omega)\} \ge \alpha_i\}, \alpha_i \in (0, 1).$)

First, employing the triangular inequality for two $s$–dimensional distribution functions $F$, $G$ for which the Problem (1) is well defined, we can obtain

$$|\varphi(F, X_F) - \varphi(G, X_G)| \le |\varphi(F, X_F) - \varphi(G, X_F)| + |\varphi(G, X_F) - \varphi(G, X_G)|. \quad (3)$$

Furthermore, if there exist real–valued continuous functions $g_i := g_i(x)$, $x \in \mathbb{R}^n$ and $\alpha_i$, $i = 1, 2, \ldots, s$ such that

$$\begin{aligned} X_F (:= X_F(\alpha)) &= \bigcap_{i=1}^{s} \{x \in X : P[\omega : g_i(x) \le \xi_i] \ge \alpha_i\}, \\ & \quad \alpha_i \in (0, 1),\, i = 1, \ldots, s, \quad \alpha = (\alpha_1, \ldots, \alpha_s), \end{aligned} \tag{4}$$

and if Assumption A.3 is fulfilled, then

$$X_F = \bigcap_{i=1}^{s} \{x \in X : g_i(x) \le k_{F_i}(\alpha_i)\}. \tag{5}$$

Consequently, setting

$$\bar{X}(v) = \bigcap_{i=1}^{s} \{x \in X : \, g_i(x) \le v_i\},\, v = (v_1, \ldots, v_s) \tag{6}$$

we obtain

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

Let us recall the definition of the Hausdorff distance in the space of nonempty subsets of $\mathbb{R}^n$.

**Definition 8.** *[34] If $X^{'}$, $X^{''} \subset \mathbb{R}^n$ are two nonempty 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 9.** *Let $X$ be nonempty set. If*

1. *$\hat{g}_0 := \hat{g}_0(x)$, $x \in \mathbb{R}^n$ is a Lipschitz function on $X$ with the Lipschitz constant $L$,*
2. *$\bar{X}(v)$ are nonempty sets for every $v \in Z_F$ and, moreover, there exists a constant $\hat{C} > 0$ such that*

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

*then*

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

*Proof.* Proposition 9 is a slightly modified version of Proposition 1 in [20].

Employing the technique of [20] or [22] we can see that the following assertion is valid.

**Lemma 10.** *[20] Let $\varepsilon > 0$, $X$ be a convex, compact and nonempty set. If*

1. *$g_i(x)$, $i = 1, \dots s$ are convex, continuous, bounded functions on $X(\varepsilon)$,*
2. *$\bar{X}(v)$ is a nonempty set for every $v \in \mathbb{R}^s$, $v \in Z_F$,*

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

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

**Lemma 11.** *Let $s = 1$, $\alpha \in (0, 1)$. If Assumptions A.2, A.3 and A.4 are fulfilled, $0 < t^{'} < \delta$, then*

$$P\{\omega : |k_{F^N}(\alpha) - k_F(\alpha)| > t^{'}\} \le 2\exp\{-2N(\vartheta t^{'})^2\}, \quad N \in \mathbb{N}.$$

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

*Proof.* First, $k_{F^N}(\alpha)$ evidently depends on $\omega \in \Omega$, and, moreover, if (for some $\omega \in \Omega$) it holds that $F^N(k_F(\alpha) + t^{'}) > 1 - \alpha$, $\quad F^N(k_F(\alpha) - t^{'}) < 1 - \alpha$, then

$$k_{F^N}(\alpha) \in \langle k_F(\alpha) - t^{'}, k_F(\alpha) + t^{'}\rangle.$$

According to the assumptions and to the classical results of [7] we can obtain that

$$\begin{aligned}
&P\{\omega :\ F^N(k_F(\alpha) + t^{'}) > 1 - \alpha, \quad F^N(k_F(\alpha) - t^{'}) < 1 - \alpha\} = \\
&= P\{\omega :\ F^N(k_F(\alpha) + t^{'}) > F(k_F(\alpha)), \quad F^N(k_F(\alpha) - t^{'}) < F(k_F(\alpha))\} = \\
&P\{\omega :\ F(k_F(\alpha) + t^{'}) - F^N(k_F(\alpha) + t^{'}) < F(k_F(\alpha) + t^{'}) - F(k_F(\alpha)), \\
&F^N k_F(\alpha) - t^{'}) - F(k_F(\alpha) - t^{'}) < F(k_F(\alpha)) - F(k_F(\alpha) - t^{'})\} \ge \\
&P\{\omega : F(k_F(\alpha) + t^{'}) - F^N(k_F(\alpha) + t^{'}) < \vartheta t^{'}, \\
&F^N(k_F(\alpha) - t^{'}) - F(k_F(\alpha) - t^{'}) < \vartheta t^{'}\} \ge \\
&1 - 2\exp\{-2N(\vartheta t^{'})^2\}.
\end{aligned}$$

Now we can already see that the assertion of Lemma 11 is valid. □

(For the exponential convergence rate of empirical quantiles see also [42].)

**Corollary 12.** *Let $X$ be a convex, compact and nonempty set, $\alpha_i \in (0, 1)$, $i = 1, \dots, s$, $\alpha = (\alpha_1, \dots, \alpha_s)$. If*

1. *$\hat{g}_0(x)$, $x \in \mathbb{R}^n$ is a Lipschitz function on $X$ with the Lipschitz constant $L$,*
2. *A.2, A.3, A.4 are fulfilled,*

3. $\bar{X}(v)$ *are nonempty sets for* $v \in Z_F$,

4. *there exists* $\varepsilon > 0$ *such that* $g_i(x)$, $i = 1, \ldots, s$ *are convex continuous functions on* $X(\varepsilon)$,

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

$$P\{\omega : |\inf_{\bar{X}(k_F(\alpha))} \hat{g}_0(x) - \inf_{\bar{X}(k_{F^N}(\alpha))} \hat{g}_0(x)|| > t\} \leq 2s \exp\{-2N(\vartheta t/LCs)^2\}$$
$$\text{for } N \in \mathbb{N} \quad \text{and} \quad t > 0 \quad \text{such that} \quad 0 < t/LCs < \delta.$$

*Proof.* The assertion of Corollary 12 follows from the assertions of Lemmas 10 and, 11 and the properties of the Euclidean norm and the probability measure.

To recall the definition of the Wasserstein metric $d_{W_1^1}(F, G) := d_{W_1^1}(P_F, P_G)$, let $\mathcal{P}(\mathbb{R}^s)$ denote the set of all (Borel) probability measures on $\mathbb{R}^s$. If $\mathcal{M}_1^1(\mathbb{R}^s) = \{\nu \in \mathcal{P}(\mathbb{R}^s) : \int\limits_{\mathbb{R}^s} \|z\|_s^1 \nu(dz) < \infty\}$ and $\mathcal{D}(P_F, P_G)$ denotes the set of those measures on $\mathcal{P}(\mathbb{R}^s \times \mathbb{R}^s)$ whose marginal measures are $P_F$ and $P_G$, $\| \cdot \|_s^1$ corresponds to the $\mathcal{L}_1$ norm in $\mathbb{R}^s$, then

$$d_{W_1^1}(F, G) := d_{W_1^1}(P_F, P_G) = \inf\{ \int\limits_{\mathbb{R}^s \times \mathbb{R}^s} \|z - \bar{z}\|_s^1 \kappa(dz \times d\bar{z}) : \kappa \in \mathcal{D}(P_F, P_G)\}$$
$$\text{for } P_F, P_G \in \mathcal{M}_1^p(\mathbb{R}^s), \quad p = 1, 2.$$

(For more general types of the Wasserstein metric see, e.g., [31].)

Employing the Wasserstein metric corresponding to $\mathcal{L}_1$ norm and the results of [46], the following stability assertion has been proven.

**Proposition 13.** *[21] Let* $P_F, P_G \in \mathcal{M}_1^1(\mathbb{R}^s)$ *and let* $X$ *be a compact set. If Assumption A.1 is fulfilled, then*

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

Proposition 13 reduces (from the mathematical point of view) an $s$– dimensional case to one–dimensional. Of course, a stochastic dependence between components of the random vector is neglected there. The idea to reduce an $s$–dimensional case, $s > 1$ to a one dimensional case is credited to G. Pflug [30] (see also [44]).

*Remark 14.* Consider two very simple optimization problems:

Find $\min\limits_{x \in \langle 1, 2\rangle} \mathbb{E}_{F_\xi} \xi x$ ; Find $\min\limits_{x \in \langle 1, 2\rangle} \mathbb{E}_{F_\eta} \eta x$,

where $\xi$ is a random value with uniform distribution $F_\xi$ on $\langle 0, 1\rangle$ and $\eta = \xi + a$, $a \in (0, 1)$. Evidently, then

$$|\min_{x \in \langle 1, 2\rangle} \mathbb{E}_{F_\xi} \xi x - \min_{x \in \langle 1, 2\rangle} \mathbb{E}_{F_\eta} \eta x| = a$$

and, simultaneously,

$$\int\limits_{-\infty}^{\infty} |F_\xi(u) - F_\eta(u)|\,du = \int\limits_{-\infty}^{\infty} (F_\xi(u) - F_\eta(u))du =$$
$$\int\limits_{0}^{a} udu + \int\limits_{1}^{1+a} (1 - u + a)du + \int\limits_{a}^{1} adu = \tfrac{a^2}{2} + \tfrac{a^2}{2} + a - a^2 = a.$$

($F_\eta$ denotes the distribution function of $\eta$.)

Consequently setting $F := F_\xi$, $G := F_\eta$ and $X := \langle 1,\, 2\rangle$ in Proposition 13 we have a case in which the inequality introduced in Proposition 13 turns to an equality.

Replacing $G$ by $F^N$ in Proposition 13, we can investigate properties of the empirical estimate $\varphi(F^N,\, X)$. It follows from Proposition 13 that properties of $|\varphi(F,\, X) - \varphi(F^N,\, X)|$ follows from the properties of $\int\limits_{-\infty}^{+\infty} |F_i(z_i) - F_i^N(z_i)|\mathrm{d}z_i,\ i = 1,\, 2,\, \ldots,\, s$. We recall the following assertions:

**Lemma 15.** *[43] Let $s = 1$ and $P_F \in \mathcal{M}_1^1(\mathbb{R}^1)$. Let, moreover, Assumption A.2 be fulfilled. Then*

$$P\{\omega : \int\limits_{-\infty}^{\infty} |F(z) - F^N(z)|\mathrm{d}z \xrightarrow[N\to\infty]{} 0\} = 1.$$

**Proposition 16.** *[9] Let $s = 1, t > 0$ and Assumptions A.2, A.3 be fulfilled. If there exists $\beta > 0$, $R := R(N) > 0$ defined on $\mathbb{N}$ such that $R(N) \xrightarrow[N\to\infty]{} \infty$ and, moreover,*

$$\begin{aligned} &N^\beta \int\limits_{-\infty}^{-R(N)} F(z)\mathrm{d}z \xrightarrow[N\to\infty]{} 0, \qquad N^\beta \int\limits_{R(N)}^{\infty} [1 - F(z)]\mathrm{d}z \xrightarrow[N\to\infty]{} 0,\\ &2NF(-R(N)) \xrightarrow[N\to\infty]{} 0, \qquad 2N[1 - F(R(N))] \xrightarrow[N\to\infty]{} 0,\\ &(\frac{12N^\beta R(N)}{t} + 1)\exp\{-2N(\frac{t}{12R(N)N^\beta})^2\} \xrightarrow[N\longrightarrow\infty]{} 0, \end{aligned} \tag{7}$$

*then*

$$P\{\omega :\, N^\beta \int\limits_{-\infty}^{\infty} |F(z) - F^N(z)|\mathrm{d}z > t\} \xrightarrow[N\to\infty]{} 0. \tag{8}$$

$\mathbb{N}$ *denotes the set of natural numbers.*

According to the assertion of Proposition 16 and to the old results of Dvoretzky–Kiefer–Wolfowitz (see, e.g., [7]) we recognize that the convergence rate can depend on the distribution tails. Evidently, if $F_i$, $i = 1,\, \ldots,\, s$ have exponential tails, then (8) is valid for $\beta \in (0,\, 1/2)$ (compare with the Kolmogorov limit theorem; [33]). Consequently, the case of heavy tails is more interesting.

# 4. Heavy Tails in Stochastic Optimization

## 4.1. Heavy Tails in Economic and Financial Applications

Previously, it was assumed (in stochastic optimization theory) that the "underlying" probability measures belong to the class of distribution functions with thin tails. Consequently it was supposed that distributions can mostly be approximated by the normal distributions. However, recently it has been recognized that many data correspond to distributions for which the above mentioned assumptions are not fulfilled. In particular, empirical distributions determined by the data are "near" to the heavy–tailed distributions

A relatively good analysis of the "heavy"–tailed distributions in economy and finance is presented, e.g., in [28]. In particular it was discovered that some data on river flow, cotton, exchange rate and returns correspond to different random parameters with heavy–tailed distributions. The Weibull distribution often corresponds to lifetime value as well as to problems about wind speed and power, and rainfall intensity. (It seems useful to comment on the position of the Weibull distribution that can belong to the thin and also the heavy tails distributions depending on the parameter value.) Furthermore, it was mentioned in [11] that some gold prices, telecommunication, quality control data, well as problems about incomes correspond to the lognormal distribution. A relationship between heavy–tailed distributions and the stable distributions can be found, e.g. in [25]. The relationship between the tails of the stable distributions and the Pareto tails is shown in [40].

According to the aforementioned facts, it is easy to see that the distributions with "heavy" tails correspond well to many economic and financial data. Consequently, a question arises: how good are empirical estimates corresponding to them? Are these estimates consistent and what can be said about the convergence rate and an asymptotic distribution? Some results about consistency as well as well convergence rate are already known. In this paper we try to also cover the results corresponding to heavy–tailed distributions including the stable distributions with a shape parameter, corresponding to the tail, $\nu \in (1, 2)$ (for the definition of the shape parameter and stable distributions see, e.g., [25]). An alternative approach to the optimization problems with the"underlying" heavy–tailed distribution can be found, e.g., in [32].

## 4.2. Consistency

First, according to Proposition 13 and Lemma 15 the following assertion is valid.

**Theorem 17.** *Let Assumptions A.1, A.2, and A.3 be fulfilled, $P_F \in \mathcal{M}_1^1(\mathbb{R}^s)$, $X$ be a compact set. Then*

$$P\{\omega : |\varphi(F^N, X) - \varphi(F, X)| \xrightarrow[N\to\infty]{} 0\} = 1.$$

*Remark 18.* According to Theorem 17 we can see that $\varphi(F^N, X)$ is (under rather general assumptions) a consistent estimate of $\varphi(F, X)$. Evidently this assertion is also valid for the stable distributions with the shape parameter $\nu > 1$.

### 4.3. Convergence Rate

The next assertion follows from Proposition 13 and Relations (7), (8).

**Theorem 19.** *[9] Let Assumptions A.1, A.2, and A.3 be fulfilled, $P_F \in \mathcal{M}_1^1(\mathbb{R}^s)$, $t > 0$, $X$ be a compact set. If*

1. *for some $r > 2$ it holds that $\mathbb{E}_{F_i}|\xi_i|^r < +\infty, \quad i = 1, \ldots, s,$*
2. *constants $\beta, \gamma > 0$ fulfill the inequalities $0 < \beta + \gamma < 1/2, \quad \gamma > 1/r,$ $\beta + (1 - r)\gamma < 0,$*

*then*

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

Evidently, the convergence rate $\beta := \beta(r)$ (introduced by Theorem 19) generally depends on existence of the finite absolute moments and, moreover, it is easy to see that

$$\beta(r) \xrightarrow[r \to +\infty]{} 1/2, \qquad \beta(r) \xrightarrow[r \to 2^+]{} 0.$$

Unfortunately, we cannot obtain (using this theoretical approach) results in the case when there exist only $\mathbb{E}_F|\xi_i|^r$, $i = 1, \ldots s$ for $r < 2$. It seems that a relationship exists between the convergence rate and a domain of attraction for a normal law (for the definition of domain of attraction see, e.g. [27]). To obtain at least a weaker result for the case when the finite moment exists only for $r < 2$, we recall the results of [1]. To this end we mention that stable distributions with $\nu < 2$ belong to this class. The stable distributions, with exception of a normal case, belong to the heavy tails distributions. The shape parameter expresses how "heavy" the tails of the distribution are. The case $\nu = 2$ correspond to normal distribution; of course where the finite second moment exists. However, the finite second moment does not exist for $\nu < 2$.

**Proposition 20.** *[1] 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*

$$\frac{N}{N^{1/\nu}} \int_{-\infty}^{\infty} |F^N(z) - F(z)| dz, \; N = 1, \ldots, \tag{9}$$

*is stochastically bounded if and only if*

$$\sup_{t > 0} t^{\nu}\, P\{\omega : |\xi| > t\} < \infty. \tag{10}$$

The assertion of Proposition 20 follows from Theorem 2.2 [1]. According to the definition of stochastically bounded random sequences, it holds (under the validity of the relation (10)) that

$$\lim_{M \longrightarrow \infty} \sup_{N} P\{\omega : \frac{N}{N^{1/\nu}} \int_{-\infty}^{\infty} |F(z) - F^N(z)| > M\} = 0. \tag{11}$$

Applying the assertion of Proposition 20, the following weaker assertion can be proven.

**Theorem 21.** *[24] Let Assumptions A.1, A.2 and A.3 be fulfilled, $P_F \in \mathcal{M}_1^1(R^s)$, $M > 0$, $X$ be a compact set. If one–dimensional components $\xi_i$, $i = 1, \ldots, s$ of the random vector $\xi$ have distribution functions $F_i$ with tails 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*

$$\lim_{M \longrightarrow \infty} \sup_N P\{\omega : \frac{N}{N^{1/\nu}} |\varphi(F^N, X) - \varphi(F, X)| > M\} = 0 \text{ with } \nu = \min(\nu_1, \ldots, \nu_s). \tag{12}$$

*Remark 22.* Let us assume that Assumptions of Theorem 21 are fulfilled and $\beta := \beta(\nu) = 1 - 1/\nu$. Then $\beta(\nu)$ is an increasing function of $\nu$ and holds up to

$$\lim_{\nu \longrightarrow 1^+} \beta(\alpha) = 0, \quad \lim_{\nu \longrightarrow 2^-} \beta(\nu) = \frac{1}{2}.$$

We have considered the special case $X_F = X$. However, a generalized assertion can be proven in the case of $X_F$ fulfilling Relation (4).

**Theorem 23.** *Let $X$ be a convex compact nonempty set, $P_F \in \mathcal{M}_1^1(R^s)$, $t, \varepsilon > 0, \alpha_i \in (0, 1)$, $i = 1, \ldots, s$, $\alpha = (\alpha_1, \ldots, \alpha_s)$. If*

1. *A.1, A.2, A.3 and A.4 are fulfilled,*
2. *for some $r > 2$ it holds that $\mathsf{E}_{F_i} |\xi_i|^r < +\infty$, $\quad i = 1, \ldots, s$,*
3. *$\beta, \gamma > 0$ fulfill inequalities $0 < \beta + \gamma < 1/2$, $\quad \gamma > 1/r$, $\quad \beta + (1 - r)\gamma < 0$,*
4. *$g_0(x, z)$ is a Lipschitz function on $X(\varepsilon)$ with the Lipschitz constant $L'$ not depending on $z \in Z_F$,*
5. *$X_F := X_F(\alpha)$ is defined by Relation (4) with the continuous convex and bounded functions $g_i(x)$, $i = 1, \ldots, s$ on $X(\varepsilon)$ and, moreover, $\bar{X}(v)$ (defined by Relation (6)) are nonempty sets for $v \in Z_F$,*

*then*

$$P\{\omega : N^\beta |\varphi(F, X_F) - \varphi(F^N, X_{F^N})| > t\} \xrightarrow[N \to \infty]{} 0. \tag{13}$$

*Proof.* First, according to the assumptions and to the relations (4), (5) and (6) we can see that $X_F(\alpha)$ is, for a fixed $\alpha$, a convex, compact and nonempty set; consequently it follows from Theorem 19 that

$$P\{\omega : N^\beta |\varphi(F, X_F) - \varphi(F^N, X_F)| > t\} \xrightarrow[N \to \infty]{} 0. \tag{14}$$

It follows from Lemma 10 and Assumption 5 that a constant $C^{'}$ exists such that for $\omega \in \Omega$ the following inequality is valid

$$\Delta_n[\bar{X}(k_F(\alpha)),\, \bar{X}(k_{F^N}(\alpha))] \,\leq\, C^{'}\|k_F(\alpha) - k_{F^N}(\alpha)\|.$$

According to Assumption 4 we can see that $\mathbb{E}_{F^N} g_0(x,\, \xi)$ on $X(\varepsilon)$ is a Lipschitz function of $x$ with the Lipschitz constant $L^{'}$ not depending on $\omega \in \Omega$. Consequently, according to Proposition 9 we can obtain

$$|\inf_{\bar{X}(k_F(\alpha))} \mathbb{E}_{F^N}\, g_0(x,\, \xi) - \inf_{\bar{X}(k_{F^N}(\alpha))} \mathbb{E}_{F^N} g_0(x,\, \xi)| \,\leq\, L^{'}C^{'}\|k_F(\alpha) - k_{F^N}(\alpha)\| \quad \text{for} \quad \omega \in \Omega;$$

consequently also for $t > 0$

$$\begin{aligned}&P\{\omega : |\inf_{\bar{X}(k_F(\alpha))} \mathbb{E}_{F^N}\, g_0(x,\, \xi) - \inf_{\bar{X}(k_{F^N}(\alpha))} \mathbb{E}_{F^N} g_0(x,\, \xi)| \,> t\} \leq \\ &P\{\omega : L^{'}C^{'}\|k_{F^N}(\alpha) - k_F(\alpha)\| \geq t\}.\end{aligned}$$

Now, employing the assertion of Lemma 11 and the properties of the Euclidean norm, we obtain for every $t > 0$ such that $t/L^{'}C^{'}s < \delta$ that

$$P\{\omega : |\inf_{\bar{X}(k_F(\alpha))} \mathbb{E}_{F^N}\, g_0(x,\, \xi) - \inf_{\bar{X}(k_{F^N}(\alpha))} \mathbb{E}_{F^N} g_0(x,\, \xi)| \,> t\} \leq 2s \exp\{-2N(\vartheta t/sL^{'}C^{'})^2\}.$$

However, it evidently follows from the last inequality that for $\beta > 0$ and enough large $N \in \mathbb{N}$ that

$$\begin{aligned}&P\{\omega : N^{\beta}|\inf_{\bar{X}(k_F(\alpha))} \mathbb{E}_{F^N}\, g_0(x,\, \xi) - \inf_{\bar{X}(k_{F^N}(\alpha))} \mathbb{E}_{F^N} g_0(x,\, \xi)| \,> t\} \,\leq \\ &2s \exp\{-2N(\vartheta t/N^{\beta}L^{'}C^{'}s)^2\}\end{aligned}.$$

and furthermore, employing the properties of the exponential functions we obtain

$$\begin{aligned}&P\{\omega : N^{\beta}|\inf_{\bar{X}(k_F(\alpha))} \mathbb{E}_{F^N}\, g_0(x,\, \xi) - \inf_{\bar{X}(k_{F^N}(\alpha))} \mathbb{E}_{F^N} g_0(x,\, \xi)| \,> t\} \xrightarrow[N \to \infty]{} 0, \\ &\quad \text{for} \quad \beta \in (0,\, 1/2).\end{aligned} \tag{15}$$

The assertion of Theorem 23 now follows immediately from Relations (3), (14) and (15).

□

The following assertion follows from the arguments used in the proof of Theorem 23 (of Relation (15)).

**Corollary 24.** *Let $X$ be a convex, compact nonempty set, $P_F \in \mathcal{M}_1^1(R^s)$, $t,\, \varepsilon > 0, \alpha_i \in (0,\, 1)$, $i = 1,\, \ldots,\, s$, $\alpha = (\alpha_1,\, \ldots,\, \alpha_s)$. If*

1. *A.1, A.2, A.3 and A.4 are fulfilled,*
2. *there exists a function $\hat{g}_0$ defined on $\mathbb{R}^n$ such that $g_0(x,\, z) = \hat{g}_0(x)$ for $x \in X(\varepsilon)$, $z \in Z_F$,*

3. $X_F := X_F(\alpha)$ *is defined by Relation (4) with the continuous convex and bounded functions* $g_i(x)$, $i = 1, \ldots, s$ *on* $X(\varepsilon)$ *and, moreover,* $\bar{X}(v)$ *(defined by Relation (6)) are nonempty sets for every* $v = (v_1, \ldots, v_s)$, $v \in Z_F$,

*then*

$$P\{\omega : N^{\beta}|\varphi(F, X_F) - \varphi(F^N, X_{F^N})| > t\} \longrightarrow_{N \longrightarrow \infty} 0 \quad \text{for} \quad \beta \in (0, 1/2). \tag{16}$$

*Proof.* Since it follows Assumption 2 of Corollary 24 that

$$\mathbb{E}_F g_0(x, \xi) = \hat{g}_0(x), \quad \mathbb{E}_{F^N} g_0(x, \xi) = \hat{g}_0(x), \quad \text{for every} \quad x \in X(\varepsilon),$$

we can see that

$$\begin{aligned} \varphi(F, X_F) &= \inf_{X_F(\alpha)} \mathbb{E}_F g_0(x, , \xi) &= \inf_{X_F(\alpha)} \hat{g}_0(x) \\ \varphi(F^N, X_F) &= \inf_{X_F(\alpha)} \mathbb{E}_{F^N} g_0(x, , \xi) &= \inf_{X_F(\alpha)} \hat{g}_0(x) \end{aligned}$$

and so

$$N^{\beta}|\varphi(F, X_F) - \varphi(F^N, X_F)| = 0 \quad \text{for every} \quad \omega \in \Omega, \, \beta > 0, \, N \in \mathbb{N}.$$

Consequently,

$$\begin{aligned} &P\{\omega : N^{\beta}|\varphi(F, X_F) - \varphi(F^N, X_{F^N})| > t\} = \\ &P\{\omega : N^{\beta}| \inf_{\bar{X}(k_F(\alpha))} \mathbb{E}_{F^N} g_0(x, \xi) - \inf_{\bar{X}(k_{F^N}(\alpha))} \mathbb{E}_{F^N} g_0(x, \xi)| > t\}. \end{aligned}$$

The assertion of Corollary 24 follows from the last relation and Relations (3), (15).

□

*Remark 25.* In the case when the function under the operator of mathematical expectation in (1) does not depend on the random factor, we obtain the "best" convergence rate under very general conditions (including the class of the stable distributions with the shape parameter from the interval (1, 2).)

**Theorem 26.** *Let Assumptions A.1, A.2, A.3 and A.4 be fulfilled,* $P_F \in \mathcal{M}_1^1(R^s)$, $M > 0$, $\varepsilon > 0$, $\alpha_i \in (0, 1)$, $i = 1, \ldots, s$, $\alpha = (\alpha_1, \ldots, \alpha_s)$, $X$ *be a compact set. If*

1. *one–dimensional components* $\xi_i$, $i = 1, \ldots, s$ *of the random vector* $\xi$ *have the distribution functions* $F_i$ *with the shape parameters* $\nu_i \in (1, 2)$ *fulfilling the relations*
$$\sup_{t > 0} t^{\nu_i} P\{\omega : |\xi_i| > t\} < \infty, \qquad i = 1, 2, \ldots, s,$$

2. $g_0(x, z)$ *is a Lipschitz function on* $X(\varepsilon)$ *with the Lipschitz constant not depending on* $z \in Z_F$,

3. $X_F := X_F(\alpha)$ *is defined by Relation (4) with the continuous convex and bounded functions* $g_i(x)$, $i = 1, \ldots, s$ *on* $X(\varepsilon)$ *and, moreover,* $\bar{X}(v)$ *(defined by Relation (6)) are nonempty sets for every* $v = (v_1, \ldots, v_s)$, $v \in Z_F$,

*then*

$$\lim_{M \longrightarrow \infty} \sup_{N} P\{\omega : \tfrac{N}{N^{1/\nu}} |\varphi(F^N, X_{F^N}) - \varphi(F, X_F)| > M\} = 0$$
with $\nu = \min(\nu_1, \ldots, \nu_s)$.

*Proof.* We can successively obtain

$$\begin{aligned}
&\lim_{M \longrightarrow \infty} \sup_{N} P\{\omega : \tfrac{N}{N^{1/\nu}} |\varphi(F^N, X_{F^N}) - \varphi(F, X_F)| > M\} \leq \\
&\lim_{M \longrightarrow \infty} \sup_{N} P\{\omega : \tfrac{N}{N^{1/\nu}} |\varphi(F^N, X_F) - \varphi(F, X_F)| > M/2\} + \\
&\lim_{M \longrightarrow \infty} \sup_{N} P\{\omega : \tfrac{N}{N^{1/\nu}} |\varphi(F^N, X_{F^N}) - \varphi(F^N, X_F)| > M/2\}.
\end{aligned}$$

Employing the main idea of the proof of Theorem 24 we can first see that the following assertion is brought forth from Theorem 21

$$\lim_{M \longrightarrow \infty} \sup_{N} P\{\omega : \frac{N}{N^{1/\nu}} |\varphi(F^N, X_F) - \varphi(F, X_F)| > M/2\} = 0;$$

furthermore,

$$\begin{aligned}
&\lim_{M \longrightarrow \infty} \sup_{N} P\{\omega : \tfrac{N}{N^{1/\nu}} |\varphi(F^N, X_{F^N}) - \varphi(F^N, X_F)| > M/2\} \leq \\
&\lim_{M \longrightarrow \infty} \sup_{N} P\{\omega : N^{(1-1/\nu)} |\varphi(F^N, X_{F^N}) - \varphi(F^N, X_F)| > M/2\} \leq \\
&\lim_{M \longrightarrow \infty} \sup_{N} P\{\omega : N^{(1-1/\nu)} \|k_F(\alpha) - k_{F^N}(\alpha)\|\} > M/2\} \leq \\
&\lim_{M \longrightarrow \infty} \sup_{N} 2s \exp\{-2N(M/2N^{(1-1/\nu)})^2\} \leq \lim_{M \longrightarrow \infty} 2s \exp\{-2(M))^2\}.
\end{aligned}$$

We can see that, the assertion of Theorem 26 follows from the last systems of the relations.

□

# 5. Dependent Samples in Stochastic Optimization

## 5.1. Historical Survey

It has often been recognized that the data do not correspond to an independent random sample. On the other hand, it can frequently be assumed that a stochastic dependence between elements of the corresponding random sequence is going to zero with growing time. It means that such data can often be modeled by mixing random sequences. Some applications leading to such situations can be found, e.g. in [4], [5], [29] or [51].

The investigation of empirical estimates with an "underlying" dependent random sample started in the 1990s (see, e.g., [17], [19] or [48]). These results have been followed by [3], [10], [23], [26], [49]. In this paper we focus on the convergence rate. To this end, we first, let $\{\xi^i\}_{i=-\infty}^{+\infty}$ be a random sequence defined on $(\Omega, \mathcal{S}, P)$. Moreover, let $\mathcal{B}(-\infty,\, a)$ be a $\sigma$–field generated by $\ldots,\, \xi^{a-1},\, \xi^a$, and $\mathcal{B}(b,\, +\infty)$ be the $\sigma$–field generated by $\xi^b,\, \xi^{b+1},\, \ldots$. We recall some definitions:

**Definition 27.** *[5] The sequence $\{\xi^i\}_{i=-\infty}^{+\infty}$ is said to be $m$–dependent if $\mathcal{B}(-\infty,\, a)$, $\mathcal{B}(b,\, +\infty)$ are mutually independent for $b - a > m$.*

**Definition 28.** *[51] The sequence $\{\xi^i\}_{i=-\infty}^{+\infty}$ is called $\Phi$–mixing (uniformly mixing) whenever there exists $\Phi_N$ such that $\Phi_N \longrightarrow 0$ as $N \longrightarrow \infty$ fulfilling the relation*

$$|P(A \bigcap B) - P(A)P(B)| \le \Phi_N P(A), \quad A \in \mathcal{B}(-\infty,\, k),\, \mathcal{B}(k+N,\, +\infty),$$
$$-\infty < k < \infty,\, N \ge 1.$$

**Definition 29.** *[51] Let $\{\xi^i\}_{i=-\infty}^{+\infty}$ be a strongly stationary random sequence. We say that $\{\xi^i\}_{i=-\infty}^{+\infty}$ is absolutely regular with $\bar{\beta}(N)$ if*

$$\bar{\beta}(N) = \sup_k \mathbb{E} \sup_{A \in \mathcal{B}(N+k,\, +\infty)} |P(A|\mathcal{B}(-\infty,\, k)) - P(A)| \downarrow 0 \quad (N \longrightarrow \infty).$$

In [19] certain convergence rates for Problems (1) with $X_F = X$ were proven under the assumptions of $\{\xi^i\}_{i=-\infty}^{+\infty}$ fulfilling the assumptions of one of the above mentioned type of dependence. Moreover, some results were also introduced in the case of $X_F$ corresponding to the joint probability constraints (for the definition of the problems with joint probability constraints see, e.g., [39]). Some assertions on convergence rate can also be found in [23]. In this paper we focus mostly on the case of $m$–dependent random samples and generalization by the assertions recalled above.

### 5.2. $m$–dependent Random Sample

Consider the cases when $s = 1$ and when $\{\xi^i\}_{i=-\infty}^{+\infty}$ is an $m$–sequence. Employing the technique of [21] it is easy to see that for $N \in \mathbb{N}$, $n > m$ there exist $k \in \{0,\, 1,\, \ldots,\}$ and $r \in \{0,\, 1\, \ldots,\, m\}$ such that $N = mk + r$ and

$$|F^N(z) - F(z)| \le \sum_{j=1}^{m} \frac{N_j}{N} |F^{N_j}(z) - F(z)|, \tag{17}$$

where $F^{N_j}$ are empirical distributions functions already determined by $N_j$ independent random variables with the distribution function $F$; moreover $N_j = k+1$ for $mk + 1 \le N \le mk + r$ and $N_j = k$ for $mk + r < N \le m(k+1)$.

Accordingly, it is easy to see that results very similar to Theorem 19, and Theorem 21 are also valid in the case of $m$–random samples. Some results valid under the $\Phi$ conditions can be found in [23].

## 6. Conclusion

This paper deals with the investigation of empirical estimates of the optimal value in the case of one–stage "classical" (rather general) stochastic programming problems. The aim of the paper is to recall some known results under non–standard assumptions. The majority has been devoted to the case when the "underlying" distribution function has heavy tails. The presented results also cover stable distributions with a shape parameter from the interval (1, 2). Employing the idea used in [22] we can see that practically all results for heavy–tailed distributions can be generalized to some subclass of optimization problems in which dependence of the objective function on the probability measure is not linear (see e.g. [22]). Theoretical results corresponding to the part of distributions with heavy tails has been completed by simulations. The corresponding models are credited to M. Houda [9] and V. Omelchenko (published in the same issue). These simulation results appear to confirm the theoretical hypothesis.

The second part has been devoted to the case of weakly dependent random data where attention has mainly been focused on the $m$–sequences. Great effort will need to be paid to dependent sample in the future. However, in both parts the results have only been presented for estimates of the optimal value. By employing some of the growth conditions, the introduced results can be transformed to the corresponding estimates of the optimal solution (see, e.g., [36]).

Both of these groups of results appear to be suitable for the investigation of empirical estimates in the case of stochastic multistage problems with an "underlying" random sequence corresponding to an autoregressive random dependence (for the corresponding definition of the multistage stochastic programs see, e.g., [39]). However the investigation in all of these directions goes beyond the scope of this paper.

The author would like to thank anonymous referees for their many helpful comments.

## References

[1] E. Barrio, E. Giné and E. Matrán: Central limit theorems for a Wasserstein distance between empirical and the true distributions. Ann. Prob. **27** (1999), 2, 1009-1071.

[2] P. Billingsley: Ergodic Theory and Information. John Wiley & Sons, New York, 1965.

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

[4] J. Dedecker, P. Doukham, G. Lang, J. R. León, S Louhich and C. Prieur: Weak Dependence (With Examples and Applications). Springer, Berlin, 2007.

[5] P. Doukham: Mixing Properties and Examples. Lectures Notes and Statistics 85. Springer, Berlin, 1994.

[6] 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.

[7] A. Dvoretzky, J. Kiefer, and J. Wolfowitz: Asymptotic minimax character of the sample distribution function and the classical multinomial estimate. Ann. Math. Statist. **27** (1956), 642–669.

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

[9] M. Houda and V. Kaňková: Empirical estimates in economic and financial optimization problems. Bulletin of the Czech Econometric Society **19** (2012), 29, 50–69.

[10] M. Houda: Convexity in Stochastic Programming Model with Indicators of Ecological Stability. In: *Proceedings of the 30th International Conference on Mathematical Methods in Economics 2012, Part I* (Jaroslav Ramík and Daniel Stavárek, eds.), Slesian University in Opava, School of Business in Karviná, Karviná (Czech Republic), 2012, 314–319.

[11] N. L. Johnson, S. Kotz and N. Balakrisnan: Continuous Univariate Distributions (Volume 1). John Wiley & Sons, New York, 1994.

[12] Y. M. Kaniovski, A. King and R. J.-B. Wets: Probabilistic bounds for the solution of stochastic programming problems. Annals of Operations Research **56** (1995), 189–208.

[13] V. Kaňková: Optimum Solution of a Stochastic Optimization Problem. In: *Trans. 7th Prague Conf. 1974*, Academia, Prague, 1977, 239–244.

[14] V. Kaňková: An Approximative Solution of Stochastic Optimization Problem. In: *Trans. 8th Prague Conf. 1978*, Academia, Prague, 1978, 349–353.

[15] V. Kaňková: Uncertainty in Stochastic Programming. In: *Proceedings of International Conference on Stochastic Optimization, Kiev 1984* (V. I. Arkin, A. Shyriaev and R. Wets, eds.), Lectures Notes on Optimization 81, Springer, Berlin, 1986, 327–332.

[16] V. Kaňková: Estimates in stochastic programming–chance constrained case. Problems of Control and Information Theory **18** (1989), 4, 251–260.

[17] V. Kaňková: On the convergence rate of empirical estimates in chance stochastic programming. Kybernetika **26** (1990), 6, 449–451.

[18] V. Kaňková and P. Lachout: Convergence rate of empirical estimates in stochastic programming. Informatica **3** (1992), 4, 497–523.

[19] V. Kaňková: A note on estimates in stochastic programming. J. Comput. Appl. Math. **56** (1994), 97–112.

[20] V.Kaňková: On the stability in stochastic programming: the case of individual probability constraints. Kybernetika **33** (1997), 5, 525–544.

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

[22] V. Kaňková: Empirical estimates in stochastic programming via distribution tails. Kybernetika **46** (2010), 3, 459–471.

[23] V. Kaňková: Dependent Data in Economic and Financial Problems. In: *Proceedings of the 29th International Conference on Mathematical Methods in Economics 2011, Part I* (M. Dlouhý and V. Skočdopolová, eds.), University of Economics in Prague, Prague, 2011, 327–332.

[24] V. Kaňková: Empirical Estimates in Economic and Financial Problems via Heavy Tails. In: *Proceedings of the 30th International Conference on Mathematical Methods in Economics 2012, Part I* (Jaroslav Ramík and Daniel Stavárek, eds.), Slesian University in Opava, School of Business in Karviná, Karviná (Czech Republic), 2012, 396–401.

[25] L. B. Klebanov: Heavy Tailed Distributions. MATFYZPRESS, Prague, 2003.

[26] J. Klicnerová: On Limit Theorems for Weakly Dependent Samples. In: *Proceedings of the 29th International Conference on Mathematical Methods in Economics 2011, Part I* (Hartin Dlouhý and Veronika Skočdopolová, eds.), University of Economics in Prague, Prague, 2011, 333–338.

[27] M. M. Meerschaert and H.-P.Scheffler: Limit Distributions for Sums of Independent Random Vectors (Heavy Tails in Theory and Practice). John Wiley & Sons, New York, 2001.

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

[29] P. A. Nze and P. Doukham: Weak dependence: models and applications to econometrics. Econometric Theory **20** (2004), 905–1045.

[30] G. Ch. Pflug: Scenario tree generation for multiperiod financial optimization by optimal discretization. Mat. Program. Series B **89** (2001), 2, 251–271.

[31] S. Rachev: Probability Metrics and the Stability of Stochastic Models. John Wiley & Sons, Chichester, 1991.

[32] S. T. Rachev and W. Römisch: Quantitative stability and stochastic programming: the method of probabilistic metrics. Math. Oper. Res. **27** (2002), 792–818.

[33] A. Rényi: Probability Theory. North–Holland, Amsterdam, 1970.

[34] R. T. Rockafellar, and R. J. B. Wets: Variational Analysis. Springer, Berlin, 1983.

[35] W. Römisch and A. Wakolbinger: Obtaining Convergence Rate for Approximation in Stochastic Programming. In: *Parametric Optimization and Related Topics* (J. Guddat, H. Th. Jongen, B. Kummer and F. Nožička, eds.), Akademie–Verlag, Berlin, 1987, 327–343.

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

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

[38] A. Shapiro and H. Xu: Stochastic mathematical programs with equilibrium constraints, modelling and sample average approximation. Optimization **57** (2008), 395–418.

[39] A. Shapiro, D. Dentcheva and A. Ruszczynski: *Lectures on Stochastic Programming (Modeling and Theory)*. Published by Society for Industrial and Applied Mathematics and Mathematical Programming Society, Philadelphia, 2009.

[40] N. Shiryaev: Essential in Stochastic Finance. World Scientific, New Jersey, 1999.

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

[42] J. R. R.Serfling: Approximation Theorems of Mathematical Statistics. John Wiley & Sons, New York, 1980.

[43] G. R. Shorack and J. A. Wellner: Empirical Processes with Applications to Statistics. John Wiley & Sons, New York, 1986.

[44] M. Šmíd: The expected loss in the discretization of multistage stochastic programming problems – estimation and convergence rate. Ann. Oper. Res. **165** (2009), 1, 29–45.

[45] A. B. Tsybakov: Error bounds for the methods of minimization of empirical risk (in Russian). Problemy Peredachi. Inform. **17** (1981), 50–61.

[46] S. S. Valander: Calculation of the Wasserstein distance between probability distribution on the line. Theor. Probab. Appl. **18** (1973), 784–786.

[47] S. Vogel: Stability results for stochastic programming problems, Optimization **19** (1988), 269–288.

[48] L. Wang and J. Wang: Limit distribution of statistical estimators for stochastic programs with dependent samples. Z. Angew. Math. Mech. **79** (1999), 4, 257–266.

[49] L. Wang: Asymptotics of statistical estimates in stochastic programming problems with long–range dependent samples. Math. Meth. Oper. Res. **55** (2002), 37–54.

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

[51] K. Yoshihara: Weakly Dependent Sequences and Their Applications. Vol. 1: Summation Theory for Weakly Dependent Sequences. Sanseido, Tokyo, 1992.