Empirical Estimates in Economic and Financial Optimization Problems*

Michal Houda
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: houda@utia.cas.cz

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. Many applications from economic and financial practice lead to optimization problems depending on a probability measure. A complete knowledge of the "underlying" measure is a necessary assumption to determine an exact optimal solution and an exact optimal value. Since this condition is not usually fulfilled, the solution is often determined using empirical data. Estimates of the optimal value and the optimal solution sets can be obtained by this approach only.
Many efforts has been paid to the investigation of the above mentioned estimates. Especially the consistency and the convergence rate have been investigated. However, it was mostly done for "classical" problems and "underlying" distributions with "thin" tails. The aim of this paper is to analyze these estimates from the point of the distribution tails. To this end, first, we recall some known results. Furthermore, we recall stability results based on the Wasserstein metric corresponding to $\mathcal{L}_1$ norm (see e. g. [16], [17]) and employ them to the case of "heavy" tails. Results based on simulation techniques complete our investigation.

Keywords: Stochastic programming, empirical estimates, moment generating function, stability, Wasserstein metric, $\mathcal{L}_1$ norm, Lipschitz property, consistence, convergence rate, normal distribution, Pareto distribution, Weibull distribution, distribution tails, simulation.

JEL classification: C44
AMS classification: 90C15

1. Introduction

Economic processes are usually influenced simultaneously by a random factor and a decision parameter. Consequently, constructing their mathematical model we of-

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

ten obtain an optimization problem depending on a probability measure. These problems can be static (one–stage) or dynamic. Multistage stochastic programming programs belong to a dynamic type. Employing a recursive definition (see e. g. [5], [19], [27]) we obtain a system of one–stage (mostly) parametric problems. Consequently, under some assumptions (corresponding to many practical situations), the results obtained for one–stage problems can be employed to investigate the multistage problems. So the investigation of one–stage problems is crucially important also for the multistage case.

To introduce "classical" one–stage stochastic programming problem let $(\Omega, \mathcal{S}, P)$ be a probability space; $\xi := \xi(\omega) = (\xi_1(\omega), \dots, \xi_s(\omega))$ s–dimensional random vector defined on $(\Omega, \mathcal{S}, P)$; $F := F(z)$ with $z \in \mathbb{R}^s$, P_F and Z_F denote the distribution function, the probability measure and the support corresponding to ξ. Let, moreover, $g_0 := g_0(x, z)$ be a real–valued (say continuous) function defined on $\mathbb{R}^n \times \mathbb{R}^s$; $X \subset \mathbb{R}^n$ be a nonempty set. If the symbol $\mathbb{E}_F$ denotes the operator of mathematical expectation corresponding to F and if for every $x \in X$ 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

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

In applications, we very often have to replace the measure P_F by its stochastic estimate to obtain at least an approximate optimal value and optimal solution. An empirical probability measure is a very suitable candidate for this measure estimate. Consequently, the solution of the problem (1) has to be often sought with respect to an empirical problem

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

where F^N denotes an empirical distribution function determined by a random sample $\{\xi^i\}_{i=1}^N$ (not necessary independent) corresponding to the distribution function F. If we denote the optimal solutions sets of (1) and (2) by $\mathcal{X}(F), \mathcal{X}(F^N)$, then (under rather general assumptions) $\varphi(F^N), \mathcal{X}(F^N)$ are "good" stochastic estimates of $\varphi(F), \mathcal{X}(F)$.

The properties of the above mentioned estimates have been investigated many times in the stochastic programming literature. It was shown there that these estimates are consistent under rather general assumptions. Furthermore, the convergence rate has been studied. However, there mostly results for "underlying" distribution with "weak" tails have been obtained. This assumption corresponds to "classical" situations and, moreover, it has been proven that the corresponding results have very "pleasant" properties. However, it has been recognized later that random elements corresponding to economic and financial situations do not fulfil these conditions. Consequently, a question has arisen if the above mentioned estimates have then also "acceptable" properties. The aim of this paper is to deal with the case of "heavy" tails. In particular, we try to show a relationship between finite moments existence and a rate of convergence of the corresponding estimates.

2. Historical Survey

The investigation of the empirical estimates started in 1974 by [43]; followed by many works (see e. g. [4], [13], [35], [36], [37], [39]). This topic appears also e. g. in [2], [20] or [22]). Let us recall some of the known results. First, we recall consistent results.

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. *$\{\xi^i\}_{i=1}^{\infty}$ is an ergodic sequence,*

then

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

Remark. Theorem 1 has been proven under the assumption that $\{\xi^i\}_{i=1}^{\infty}$ is an ergodic sequence. We recall that ergodic property corresponds to an invariant transformation. Of course ergodic property covers independent case. For more details see e. g. [1].

Theorem 2. *Let X be a nonempty compact set. If*

1. *in 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 function (w. r. t. F),*

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

then

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

Proof. The assertion of Theorem 2 follows immediately from Proposition 5.2 and Theorem 7.48 proven in [39].

Furthermore, we recall convergence rate results.

Theorem 3 ([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., $|g_0(x, z)| \le M$),*

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

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

Remarks.

1. $K(t, X, L')$ depends on t, X, L' and $k_1(M)$ on M. Employing their estimations presented in [14] it has been proven ([15]) that
$$P\{\omega : N^{\beta}|\varphi(F) - \varphi(F^N)| > t\} \xrightarrow[N\to\infty]{} 0 \qquad \text{for } \beta \in (0, \tfrac{1}{2}).$$
Moreover if $g_0(x,\, z)$ is uniformly strongly convex function of $x \in X$ with a parameter $\rho > 0$, then $\mathcal{X}(F)$ and $\mathcal{X}(F^N)$ are singletons and
$$P\{\omega :\, N^{\beta}\|\mathcal{X}(F) - \mathcal{X}(F^N)\|^2 > t\} \xrightarrow[N\to\infty]{} 0 \qquad \text{for } \beta \in (0, \tfrac{1}{2}).$$
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 $\lambda \in \langle 0, 1\rangle$, $x = \lambda x^1 + (1-\lambda)x^2$, $x^1, x^2 \in X$, $z \in R^s$; $\|\cdot\| = \|\cdot\|_n^2$ denotes the Euclidean norm in $\mathbb{R}^n$. For more details see e. g. [34].

1. The assertion of Theorem 3 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 [9].

If the moment generating function $M_{g_0}(t)$, corresponding to $g_0(x, \xi)$, is defined by the relation
$$M_{g_0}(t) := \mathbb{E}_F\left\{e^{t[g_0(x,\xi) - \mathbb{E}_F g_0(x,\xi)]}\right\},$$
then the following assertion has been proven in [38].

Theorem 4 ([38]). *Let $X \subset \mathbb{R}^n$ be a nonempty closed set. If*

1. *for every $x \in X$ the moment generating function $M_{g_0}(t)$ is finite valued for all t in a neighbourhood of zero,*
2. *there exists a measurable function $\kappa\colon Z_F \to \mathbb{R}_+$ 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 $M_\kappa(t)$ of $\kappa(\xi)$ is finite valued for all t in a neighbourhood of zero,*
4. *$\{\xi^i\}_{i=1}^N$, $N = 1, 2, \ldots$ is an independent random sample,*

then for any $\varepsilon > 0$ there exist positive constants $C = C(\varepsilon)$ and $\beta = \beta(\varepsilon)$, independent of N, such that
$$P\left\{\sup_{x\in X}|\mathbb{E}_{F^N} g_0(x, \xi) - \mathbb{E}_F g_0(x, \xi)| \ge \varepsilon\right\} \le C(\varepsilon)e^{-N\beta(\varepsilon)},$$
$\|\cdot\| = \|\cdot\|_n^2$ denotes the Euclidean norm in $\mathbb{R}^n$.

Remark. In Theorem 4 no assumptions are made on the function $g_0(x, z)$. The corresponding assumptions are concerning only the moment generating function M_{g_0}. This fact contrasts the assumptions of Theorems 1, 2, 3. (Moreover, M_{g_0} is in fact the moment generating function corresponding to the central moment.)

Evidently, according to the above mentioned results (as well as to the results published e. g. in [3], [10], [12], [39] and [44]), the existence of the corresponding finite moment generating functions is a "rather simple" sufficient condition to guarantee exponential convergence rate in many cases (see e. g. linear objective functions). Of course, the normal distribution fulfils this assumption and, moreover, the normal distribution (or at least normal approximation) has been former times employed usually in economic and financial applications. We can recall the well known Markowitz model of a portfolio selection where the normal distribution is usually considered and where the variance corresponds to a risk measure. It has been shown in [24] that there exists a relationship between this risk measure and absolute deviation just in the case of the normal distribution. However, relatively soon it has been recognized that many data correspond to the distributions (with "heavy" tails) for which the finite moment generating function does not exist.

A relatively "good" analysis of the "heavy" tailed distributions in economy and finance is presented e. g. in [30]. There is mentioned e. g. the fact that some data about river flow, cotton, exchange rate, returns and so on correspond just to different random parameters with heavy tails distributions. The Weibull distribution corresponds often to lifetime value as well as to problems about wind speed and power, rainfall intensity and so on. Furthermore, it was mentioned in [7] that some date about gold prices, telecommunication, quality control, but also problems about incomes correspond to the lognormal distribution. A relationship between heavy tailed distributions and the stable distributions can be found e. g. in [23]; between the stable heavy tailed distributions and the Pareto tails is known and can be found e. g. in [23] and [26] (see also [31]).

According to the above recalled facts, it is easy to see that the distributions with "heavy" tails correspond really to many economic and financial data. Consequently, a question arises: how "good" are empirical estimates corresponding to them. Are these estimates consistent and what is it valid about a convergence rate and an asymptotic distribution? Some results about consistency are known (see also Theorem 2). A minor result (about the convergence rate) has been proven in [21]. In this paper we try to extend the last theoretical results, especially we focus our investigation on the value of the convergence rate. Moreover, we complete theoretical results by simulation techniques.

3. Some Definitions and Auxiliary Assertions

Let F, G be two s–dimensional distribution functions. To recall the definition of the Wasserstein metric $d_{W_1^p}(F, G) = d_{W_1^p}(P_F, P_G)$ with $p = 1, 2$, let $\mathcal{P}(\mathbb{R}^s)$ denote the set of all (Borel) probability measure on $\mathbb{R}^s$. If $\mathcal{M}_1^p(\mathbb{R}^s) = \{\nu \in \mathcal{P}(\mathbb{R}^s) : \int_{\mathbb{R}^s} \|z\|_s^p \mathrm{d}z < \infty\}$ 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, $\|\cdot\|_s^2$ corresponds to the Euclidean norm, $\|\cdot\|_s^1$ to the $\mathcal{L}_1$

norm in $\mathbb{R}^s$, then

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

We introduce the following 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 an independent random sequence corresponding to F,
- F^N is 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 $\mathbb{R}^1$ ($P_{F_i}, i=1,2,\dots,s$ denote one-dimensional marginal probability measures corresponding to F).

Employing the Wasserstein metric corresponding to $\mathcal{L}_1$ norm and the results of [42], the following stability assertion has been proven.

Proposition 5 ([17]). *Let $P_F, P_G \in \mathcal{M}_1^1(\mathbb{R}^s)$, and let X be a compact set. If the assumption A.1 is fulfilled, then*

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

Proposition 5 reduces (from the mathematical point of view) s–dimensional case to one–dimensional. Of course, stochastic dependence between components of the random vector is there neglected.

Replacing G by F^N in Proposition 5 we can investigate properties of the empirical estimates $\varphi(F^N)$. It follows from Proposition 5 that the investigation (of the problem (2)) cannot be only reduced to the case $s=1$ however that properties of $|\varphi(F)-\varphi(F^N)|$ follows from the properties of $\int_{-\infty}^{+\infty}|F_i(z_i)-F_i^N(z_i)|\mathrm{d}z_i,\ i=1,2,\dots,s$; ($F_i^N, i=1,\dots,s$ denotes one–dimensional marginal empirical distribution functions corresponding to F^N).

Remark. The Wasserstein metric based on $\mathcal{L}_1$ norm has appeared already in [41] (see also [32]).

We recall the following assertions.

Lemma 6 ([40]). *Let $s=1$ and $P_F\in\mathcal{M}_1^1(\mathbb{R}^1)$. Let, moreover, the assumption A.2 be fulfilled. Then*

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

Proposition 7 ([18], [21]). *Let $s = 1, t > 0$ and the 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,*

$$
\begin{aligned}
N^{\beta} \int_{-\infty}^{-R(N)} F(z)\mathrm{d}z \xrightarrow[N\to\infty]{} 0, \qquad & N^{\beta} \int_{R(N)}^{\infty} [1 - F(z)]\mathrm{d}z \xrightarrow[N\to\infty]{} 0,\\
2NF(-R(N)) \xrightarrow[N\to\infty]{} 0, \qquad & 2N[1 - F(R(N))] \xrightarrow[N\to\infty]{} 0,\\
\left(\frac{12N^{\beta}R(N)}{t} + 1\right) \exp\left\{-2N\left(\frac{t}{12R(N)N^{\beta}}\right)^{2}\right\} & \xrightarrow[N\to\infty]{} 0,
\end{aligned}
\tag{3}
$$

then

$$
P\left\{\omega :\ N^{\beta}\int_{-\infty}^{\infty} |F(z) - F^{N}(z)|\mathrm{d}z > t\right\} \xrightarrow[N\to\infty]{} 0. \tag{4}
$$

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

Proof. First replacing in Proposition 5 G by F^N and employing the properties of the probability measure we can obtain for $t > 0$, $R > 0$ and $s = 1$ the following relation

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

If furthermore we set for $N = 1, 2, \ldots$

$$
\begin{aligned}
\Omega(N, R) &= \{\omega : \xi^{i}(\omega) \in (-R, R),\ i = 1, \ldots, N\},\\
\Omega^{c}(N, R) &= \Omega - \Omega(N, R),
\end{aligned}
$$

then evidently

$$
\begin{aligned}
&P\Big\{\omega : \int_{-\infty}^{\infty} |F(z) - F^{N}(z)|\mathrm{d}z > t\Big\}\\
&\quad \le P\Big\{\omega : \int_{-R}^{R} |F(z) - F^{N}(z)|\mathrm{d}z > \frac{t}{3}\Big\}\\
&\quad + P\Big\{\omega \in \Omega(N, R) : \int_{-\infty}^{-R} |F(z) - F^{N}(z)|\mathrm{d}z > \frac{t}{3}\Big\}\\
&\quad + P\Big\{\omega \in \Omega(N, R) : \int_{R}^{\infty} |F(z) - F^{N}(z)|\mathrm{d}z > \frac{t}{3}\Big\}\\
&\quad + P\Big\{\omega \in \Omega^{c}(N, R) : \int_{-\infty}^{-R} |F(z) - F^{N}(z)|\mathrm{d}z > \frac{t}{3}\Big\}\\
&\quad + P\Big\{\omega \in \Omega^{c}(N, R) : \int_{R}^{\infty} |F(z) - F^{N}(z)|\mathrm{d}z > \frac{t}{3}\Big\}
\end{aligned}
$$

and, consequently,

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

Since

1. $P\{\omega : \omega \in \Omega^c(N,\, R)\} \le NF(-R) + N(1 - F(R))$,
2. it has been proven in [6] (for independent random sample) that
$$P\left\{\omega : |F(z) - F^N(z)| > t\right\} \le 2\exp\{-2Nt^2\} \quad \text{independently on } z \in \mathbb{R}^1,$$
3. employing the results of [6] it has been proven in [18] (see also [17]) that
$$P\left\{\omega : \int_{-R}^{R} |F(z) - F^N(z)| \mathrm{d}z > t\right\} \le \left(\frac{1}{t} + 1\right) \exp\big\{-2Nt^2\big\} \quad \text{for } \frac{t}{4R} < 1,$$

we can see that the assertion of Proposition 7 is valid.

To complete this part, we recall the assertion corresponding to the case $\beta = \frac{1}{2}$.

Proposition 8 ([11], [16]). *Let the assumptions A.2 and A.3 be fulfilled,* $s = 1$, $P_F \in \mathcal{M}_1^1(\mathbb{R}^1)$. *If*

$$\int_{-\infty}^{+\infty} \sqrt{F(z)(1 - F(z)} \mathrm{d}z < +\infty,$$

then

$$\int_{-\infty}^{+\infty} \sqrt{N} |F^N(z) - F(z)| \mathrm{d}z \longrightarrow_d \int_{-\infty}^{+\infty} |\mathbb{U}(F(z))| \mathrm{d}z,$$

$\mathbb{U}$ *denotes the Brownian bridge.*

4. Convergence Rate – Analysis

In this section we try to introduce assumptions under which the relations (3) are valid and consequently (4) holds. To this end we restrict to the case $s = 1$. Let f denote the probability density corresponding to F. First, we recall the assertion corresponding to the exponential tails.

4.1. Exponential Tails

Proposition 9 ([21]). *Let $s = 1$, $t > 0$, $\beta \in (0, \frac{1}{2})$ and the assumptions A.2, A.3 be fulfilled. If there exist constants C_1, C_2 and $T > 0$ such that*

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

then

$$P\left\{\omega : N^{\beta} \int_{-\infty}^{\infty} |F(z) - F^N(z)| \mathrm{d}z > t\right\} \xrightarrow[N\to\infty]{} 0.$$

Of course, the finite moment generating function (corresponding to the random element ξ) exists in this case. Furthermore, we try to relax the assumption of "thin" (exponential) tails. In particular, we try to consider examples of the distributions with not everywhere existing finite moment generating function, however for which the assertion of Proposition 9 is valid. We consider the Weibull distribution.

4.2. Weibull Distribution

Definition 10 ([25]). *Let $s = 1$. A random variable ξ has a Weibull probability density f if there exist constants $c > 0, \nu > 0, z_0 \in \mathbb{R}^1$ such that*

$$f(z) = \begin{cases} \frac{c}{\nu}\left(\frac{z-z_0}{\nu}\right)^{c-1} \exp\{-((z-z_0)/\nu)^c\} & \textit{for } z > z_0, \\ 0 & \textit{for } z \le z_0. \end{cases}$$

Evidently, without loss of generality, we can set $z_0 = 0$. The distribution function $F(z)$ then fulfils the relation

$$F(z) = 1 - \exp\{-(z/\nu)^c\}.$$

Immediately, we can obtain for $R(N) = N^{\gamma}$, $\gamma \in (0, \frac{1}{2})$ that

$$\left(\frac{12N^{\beta}R(N)}{t} + 1\right) \exp\left\{-2N\left(\frac{t}{12R(N)N^{\beta}}\right)^2\right\}$$
$$= \left(\frac{12N^{\beta}N^{\gamma}}{t} + 1\right) \exp\left\{-2N\left(\frac{t}{12N^{\gamma}N^{\beta}}\right)^2\right\} \xrightarrow[N\to\infty]{} 0. \quad (5)$$

Furthermore, it is easy to see that the moment generating function exists in the case $c > 1$; while it does not exist when $0 < c < 1$ (for more details see [8]). To deal with the convergence rate let $\gamma, \beta > 0$, $\gamma + \beta \in (0, \frac{1}{2})$ and $R(N) = N^{\gamma}$. We distinguish two cases: $c \ge 1$, and $0 < c < 1$. If $c \ge 1$, then we can obtain (for enough large N) successively

$$N[1 - F(R(N))] = N \exp\{-(N^{\gamma}/\nu)^c\} \xrightarrow[N\to\infty]{} 0,$$
$$N^{\beta} \int_{N^{\gamma}}^{\infty} [1 - F(z)] \mathrm{d}z = N^{\beta} \int_{N^{\gamma}}^{\infty} \exp\{-(z/\nu)^c\} \mathrm{d}z \le N^{\beta} \int_{N^{\gamma}}^{\infty} \exp\{-(z/\nu)\} \mathrm{d}z$$
$$= N^{\beta}[-\nu \exp\{-(z/\nu)\}]_{N^{\gamma}}^{\infty} = N^{\beta} \nu \exp\{-N^{\gamma}/\nu\} \xrightarrow[N\to\infty]{} 0. \quad (6)$$

Consequently, it follows from the relations (3), (5) and (6) that under the assumptions A.2, A.3 the assertion of Proposition 7 is valid. It remains to consider the case $c \in (0, 1)$. First, we obtain in this case

$$N[1 - F(R(N))] = N \exp\{-(N^{\gamma}/\nu)^{c}\} \xrightarrow[N \to \infty]{} 0. \tag{7}$$

Furthermore, if we set $a = \frac{N^{c\gamma}}{\nu^{c}}$ and define $k \in \mathbb{N}$ such that $k \geq \frac{1}{c} - 1$, then (for enough large N) employing the per partes formula we obtain

$$\begin{aligned} N^{\beta} \int_{N^{\gamma}}^{\infty} [1 - F(z)]\mathrm{d}z = N^{\beta} \int_{N^{\gamma}}^{\infty} \exp\{-(z/\nu)^{c}\}\mathrm{d}z = N^{\beta} \int_{a}^{\infty} \frac{\nu}{c} u^{\frac{1}{c}-1} \exp\{-u\}\mathrm{d}u \\ \leq N^{\beta} \int_{a}^{\infty} \frac{\nu}{c} u^{k} \exp\{-u\}\mathrm{d}u \\ \leq N^{\beta} \left[-\frac{\nu}{c} u^{k} e^{-u} + \cdots + (-1)^{k} \left(\frac{\nu}{c}\right)^{k} k(k-1)\cdots 1 e^{-u}\right]_{a}^{\infty} + N^{\beta} \int_{a}^{\infty} e^{-u}\mathrm{d}u. \end{aligned} \tag{8}$$

Consequently, it follows from the properties of the exponential function that

$$N^{\beta} \int_{N^{\gamma}}^{\infty} [1 - F(z)]\mathrm{d}z \xrightarrow[N \to \infty]{} 0.$$

It follows from the relations, (5), (7) and (8) that under the assumptions A.2, A.3 the assertion of Proposition 7 is valid also in the case $c \in (0, 1)$. Evidently, we have proven the following assertion.

Proposition 11. *Let $s = 1$, $t > 0$, and the assumptions A.2, A.3 be fulfilled. If there exists constants $c > 0, \nu > 0, T > 0$ such that a probability density f fulfils the relation*

$$f(z) \leq \frac{c}{\nu} \left(\frac{|z|}{\nu}\right)^{c-1} \exp\{-(|z|/\nu)^{c}\} \quad \textit{for } z \in (-\infty, T) \cup (T, \infty),$$

then

$$P\left\{\omega : N^{\beta} \int_{-\infty}^{\infty} |F(z) - F^{N}(z)|\mathrm{d}z > t\right\} \xrightarrow[N \to \infty]{} 0 \quad \textit{for } \beta \in (0, \tfrac{1}{2}).$$

Consequently, it follows from Proposition 9 and Proposition 11 that the assertion of Proposition 7 is valid for $\beta \in (0,\ 1/2)$ in the case of the exponential tails as well as in the case tails corresponding to the Weibull distribution. It means that there can exist the distributions with “heavy” tails (for which the finite moment generating function does not exist) and simultaneously with the best convergence rate of optimal value empirical estimates (for more details see Kolmogorov limit theorem, see also [6]). This result looks very promising. However, recalling the results concerning Pareto distribution (see also simulations), we can see that only weaker results can be there guaranteed.

4.3. Pareto Distribution

Definition 12 ([30]). *Let $s = 1$. A random variable ξ has a Pareto distribution if*

$$P\{\omega : \xi > z\} = \left(\frac{C}{z}\right)^{\alpha}, \qquad f(z) = \alpha C^{\alpha} z^{-\alpha-1} \qquad \text{for } z \geq C$$

$$f(z) = 0 \qquad \text{for } z < C,$$

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

The Pareto distribution has only one tail and for $\alpha > 1$ we obtain $P_F \in \mathcal{M}_1^1(\mathbb{R}^1)$. The finite moments $\mathbb{E}_F \xi^r$ exist there for $r < \alpha$. We recall the following assertion.

Proposition 13 ([21]). *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*

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

then

$$P\Big\{\omega :\ N^{\beta} \int_{-\infty}^{\infty} |F(z) - F^N(z)| \mathrm{d}z > t\Big\} \xrightarrow[N \to \infty]{} 0.$$

Corollary 14. *Let the assumptions of Proposition 13 be fulfilled. If moreover*

1. $\alpha > 4$, $\gamma > \frac{1}{4}$, *then necessary* $\beta < \frac{1}{4}$;
2. $\alpha > 3$, $\gamma > \frac{1}{3}$, *then necessary* $\beta < \frac{1}{6}$;
3. $\alpha > 2$, $\gamma > \frac{1}{\alpha}$, *then we can obtain* $\beta := \beta(\alpha) \xrightarrow[\alpha \to 2^+]{} 0$.

If $\alpha \in (1, 2)$, then we can not obtain (in Pareto case) similar result to the results of Proposition 13.

Analyzing the cases of the Weibull and the Pareto distributions, we can see that all moments $\mathbb{E}_F \xi^r$ exist in the case of the Weibull distribution (for more details see e. g. [8]) while in the case of the Pareto distribution the moments exist only for $r < \alpha$. The convergence rate β, in the case of the Weibull distribution, is the same as in the case of exponential tails, while in the case of Pareto distribution is worse; evidently it depends on the value of the parameter α (consequently on the existence of finite moments). Moreover, it seems that for $\alpha, \alpha' > 2, \alpha < \alpha'$ the corresponding β, β' fulfil an inequality $\beta < \beta'$. Consequently, according to the assertions of Propositions 9, Proposition 11 and Proposition 13 it seems that the convergence rate depends not only on the existence of the finite moment generating function, but it depends on the existence of finite moments.

4.4. Relationship between Finite Moments Existence and Convergence Rate

Thinking about the previous results, it seems that more general results can be obtained. To this end, first, we recall a well known result of the mathematical statistics.

Proposition 15 ([8]). *Let $r > 0$, $s = 1$. Suppose that ξ is a non–negative random variable. Then*

$$\mathbb{E}_F \xi^r < \infty \quad \Longrightarrow \quad z^r P\{\omega : \xi > z\} \longrightarrow 0 \ \ \textit{as } z \longrightarrow \infty.$$

Remark. Proposition 15 is formulated for a non–negative random value ξ. However, if we assume $s = 1$, and $r > 0$ then successively

$$\mathbb{E}_F |\xi|^r < +\infty \quad \Longrightarrow \quad |z|^r P\{\omega : |\xi| > z\} \longrightarrow 0 \ \text{ as } |z| \longrightarrow \infty,$$

and simultaneously

$$P\{\omega : \xi < -z\} \le P\{\omega : |\xi| > z\}, \qquad P\{\omega : \xi > z\} \le P\{\omega : |\xi| > z\}, \ z > 0. \quad (9)$$

Evidently, we can employ the results of Proposition 7 and Proposition 15 to obtain the convergence rate for a relatively large class of the distribution functions with "heavy" tails.

Proposition 16. *Let $s = 1$, $t > 0$, $r > 0$, and the assumptions A.2, A.3 be fulfilled. Let, moreover, ξ be a random variable such that $\mathbb{E}_F|\xi|^r < \infty$. If constants $\beta, \gamma > 0$ fulfil the inequalities*

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

then

$$P\left\{\omega : N^\beta \int_{-\infty}^{\infty} |F(z) - F^N(z)| \mathrm{d}z > t\right\} \xrightarrow[N \to \infty]{} 0.$$

Proof. First setting $R(N) = N^\gamma$ we can see that for $\gamma + \beta < 1/2$ it holds

$$\left(\frac{12N^\beta R(N)}{t} + 1\right) \exp\left\{-2N\left(\frac{t}{12R(N)N^\beta}\right)^2\right\}$$
$$= \left(\frac{12N^\beta N^\gamma}{t} + 1\right) \exp\left\{-2N\left(\frac{t}{12N^\gamma N^\beta}\right)^2\right\} \xrightarrow[N \longrightarrow \infty]{} 0. \quad (10)$$

Furthermore it follows from the assumptions and from Proposition 15 for $\gamma > 1/r$ that

$$NF(-R(N)) \xrightarrow[N \to \infty]{} 0, \qquad N[1 - F(R(N))] \xrightarrow[N \to \infty]{} 0. \quad (11)$$

and, simultaneously, for $\beta + (1 - r)\gamma < 0$ that

$$N^\beta \int_{-\infty}^{-R(N)} F(z) \mathrm{d}z \xrightarrow[N \to \infty]{} 0, \qquad N^\beta \int_{R(N)}^{\infty} [1 - F(z)] \mathrm{d}z \xrightarrow[N \to \infty]{} 0. \quad (12)$$

The assertion of Proposition 16 follows now already from the assertion of Proposition 7 and the relations (10), (11) and (12).

5. Main Results

It follows from the former section that in the case $s = 1$ the convergence rate in the relation

$$P\left\{\omega : N^{\beta} \int_{-\infty}^{\infty} |F(z) - F^N(z)| \mathrm{d}z > t\right\} \xrightarrow[N \to \infty]{} 0 \quad \text{for } t > 0$$

depends on the range of finite moments existence. Consequently, employing the assertion of Proposition 5 we can obtain a relationship between the range of finite moments existence for one dimensional marginal distributions and the optimal value estimates. Consequently we obtain the results concerning the rate of convergence without the assumption that $g_0(x, z)$ is a bounded function as well as the finite moment generating function M_{g_0} exists. However, first we recall one result on consistency.

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

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

Proof. The assertion of Theorem 17 follows from Proposition 5 and Lemma 6.

The next assertion follows from Proposition 5. Evidently, it generalizes well–known results on convergence rate of empirical estimates in stochastic programming problem. Consequently, this result generalizes the former result and it can be useful in applications on economic problems. Especially according to the fact, that to verify the corresponding assumptions seems to be simple.

Theorem 18. *Let the assumptions A.1, A.2, and A.3 be fulfilled,* $P_F \in \mathcal{M}_1^1(\mathbb{R}^s)$, $t > 0$. *If*

1. for some $r > 2$ *it holds that*

$$\mathbb{E}_{F_i} |\xi_i|^r < +\infty, \quad i = 1, \ldots, s,$$

2. constants $\beta, \gamma > 0$ *fulfil the inequalities*

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

then

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

Proof. The assertion of Theorem 18 follows immediately from Proposition 5 and Proposition 16.

According to Theorem 17 we can see that under rather general assumption $\varphi(F^N)$ is a consistent estimate of $\varphi(F)$. Evidently, the assumptions of Theorem 17 can be very simply verified (however maybe they are a little stronger then the assumptions of Theorem 2). Furthermore Theorem 18 generalizes the assertions of Theorem 3 and Theorem 4. According to this result the "best" convergence rate is "practically" valid (under some additional assumptions) for all "underlying" distributions with $\mathbb{E}_{F_i}|\xi_i|^r < \infty$, $i = 1, \ldots, s$ for all $r \in \mathbb{N}$. Namely, generally, the convergence rate $\beta := \beta(r)$ depends on the absolute moments existence 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 can not obtain (by this theoretical approach) results in the case when there exist only $\mathbb{E}_F|\xi_i|^r$, $i = 1, \ldots s$ for $r \leq 2$. It seems that there exists a relationship between the convergence rate and a domain of attractions for a normal law (for the definition of domain of attraction see e. g. [29]). To obtain some further information we employ the simulation technique.

6. Simulation

In order to illustrate the asymptotic behaviour of the Wasserstein metric we compute values of the integrated empirical process

$$N^\beta \int_{-\infty}^{\infty} |F(z) - F^N(z_i)| \mathrm{d}z. \tag{13}$$

for different distributions – namely standard normal $N(0;1)$ distribution (as a benchmark), Pareto I distribution with the scale parameter $C = 1$ and different values of shape parameter α (from 1.8 to 3), and standard Weibull distribution with the shape parameters $c = 0.6$ and 0.8 and scale parameter $\nu = 1$ – and for different values of β (up to $\frac{1}{2}$). For each situation we generate random sequences of $N = 100, 1000$, and 2000 samples, and do 200 simulations in total in order to estimate the limiting distribution of the process (13).

The limiting distribution of Pareto integrated empirical process is seen to be of very large variance (compared to standard normal distribution) for $\alpha = 2$ and smaller α's; this also generates a negative impact for estimates (we have only a poor information about convergence rate). Also note the curve for $N = 100$ in Figure 1: it differs from curves for larger sample sizes – thus one concludes that using small sample sizes for estimating could be seen problematic.

We observe the convergence to zero for normal distribution with $\beta = 2/5$ in Figure 2 (the solid curve is the leftmost one). This is not evident for any of the used Pareto distributions: for $\alpha = 3$ we theoretically assure the convergence to zero only for small value of $\beta = 1/6$ (even though it is only a conservative estimate). For smaller αs, we have no theoretical results; we observe again a great value of limiting variance and we are not sure at all if the sample size of 2000 realization is enough to conclude about limiting distribution, at least empirically.

We got very similar behaviour simulating the process with $\beta = 1/3$ and with $\alpha = 1.6$ and 1.5: even greater variance and doubts about the sufficient sample size. These simulation are not presented in this paper.

For Weibull distribution we have chosen shape parameters c corresponding to situation where the moment generating function does not exists; however the assertion of the Proposition 7 is valid as shown in Section 4.2.. We observe moderate

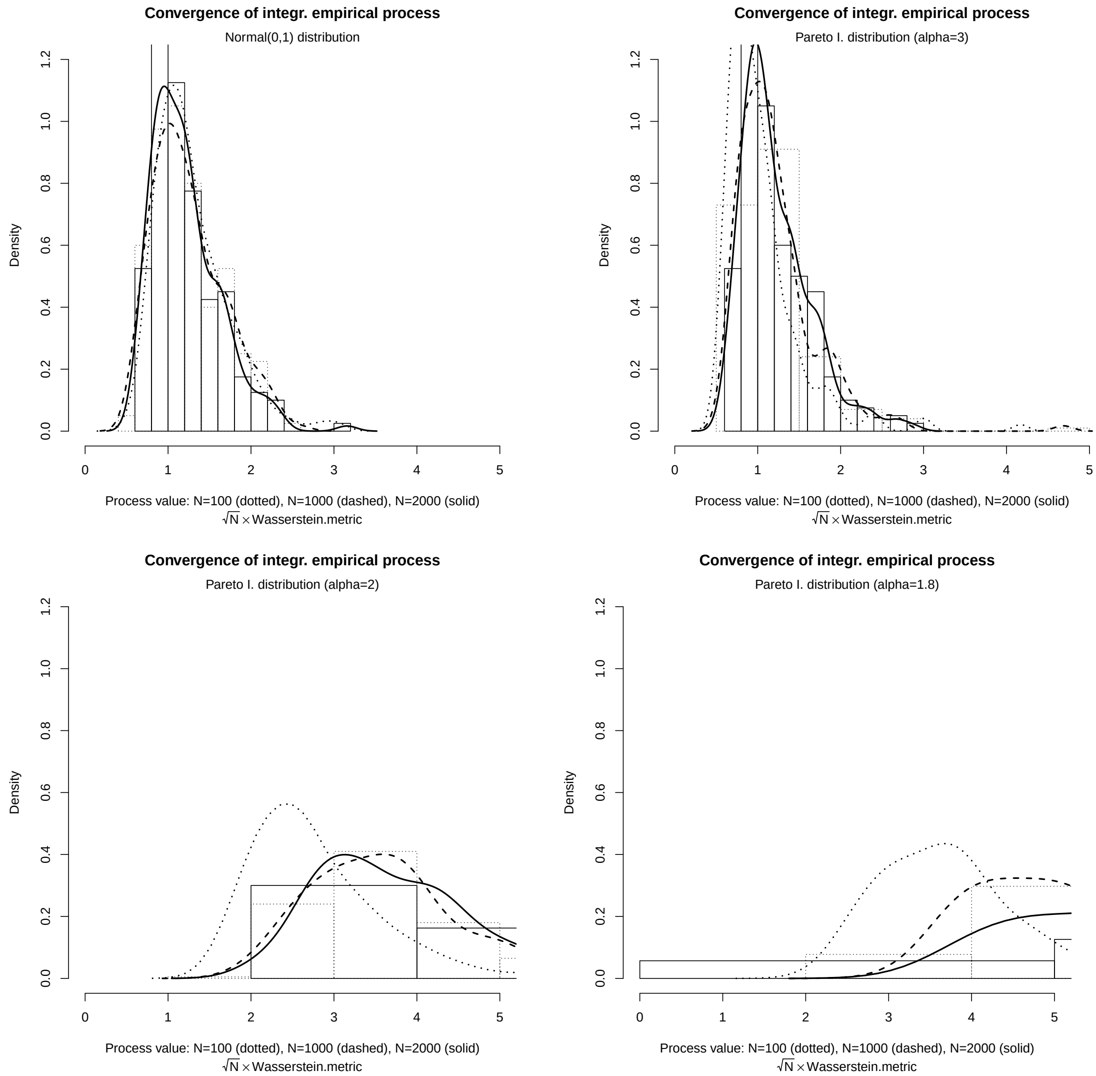

Figure 1: Normal and Pareto distributions ($\alpha = 3.0, 2.0, 1.8$) with $\beta = 1/2$

asymptotic properties for moderate value of $c = 0.8$ (see the top part of Figure 3); the convergence for $\beta < 1/2$ is theoretically assured but no visual difference between distribution of (13) for $N = 1000$ and 2000 samples is observed on the right-hand side graph. The situation is getting much more worse for the shape of $c = 0.6$ (bottom part of Figure 3): the variance of the limiting distribution for $\beta = 1/2$ (left-hand side) increases (left-hand side graph) and the convergence to zero for $\beta = 2/5$ is not visually observed even for 2000 sample size. So is for smaller values 0.5 and 0.3 which we have also computed but not finally included in this paper.

7. Conclusion

The paper deals with the investigation of empirical estimates of the optimal value in the case of one–stage (rather general) stochastic programming problems. First, some consistent results are introduced, however the main effort has been paid to the

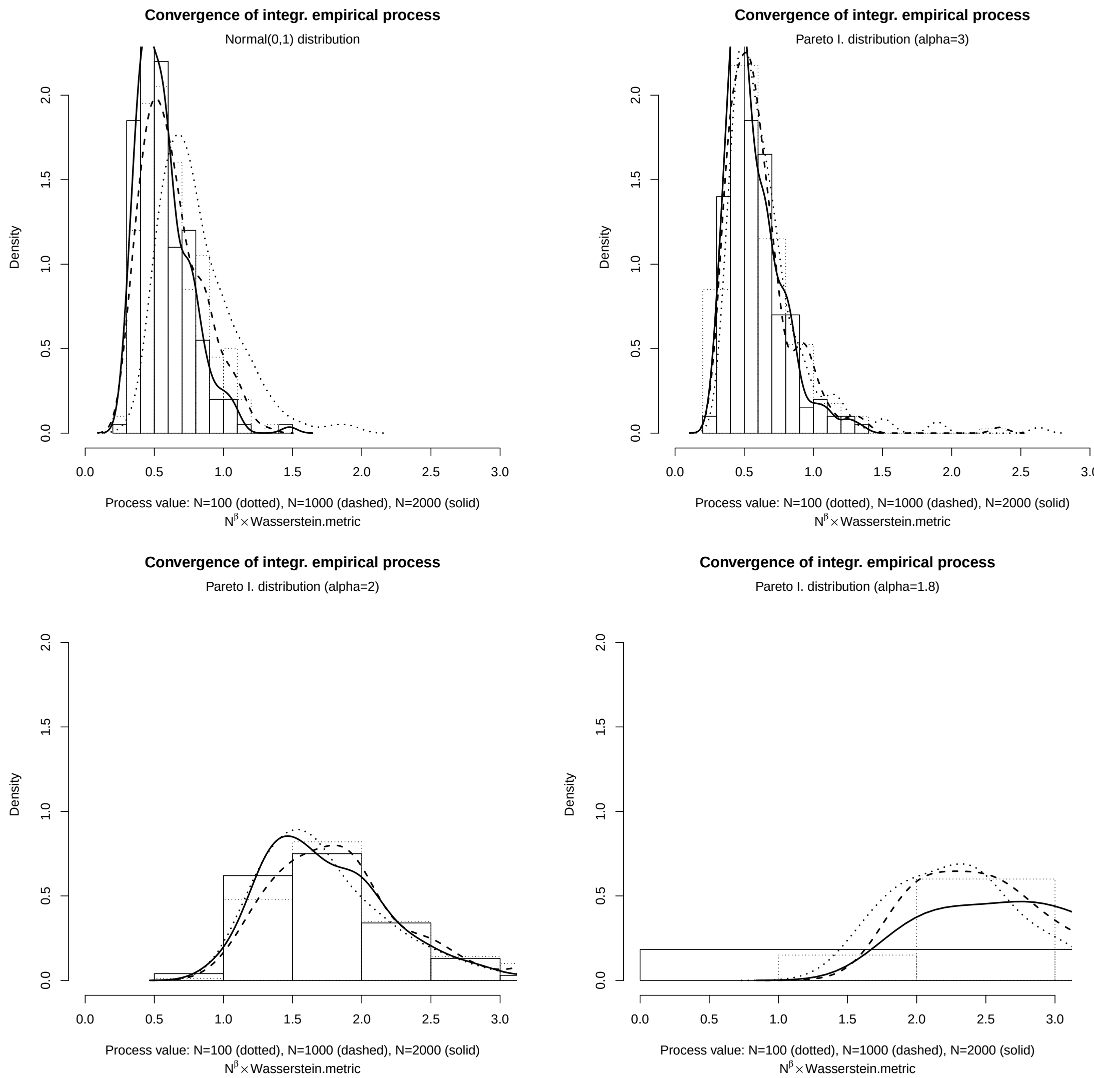

Figure 2: Normal and Pareto distributions ($\alpha = 3.0, 2.0, 1.8$) with $\beta = 2/5$

convergence rate of the optimal value with the "underlying" probability measure for which the finite moment generating function does not exist. Evidently, the classical results are generalized. Employing some growth conditions the introduced results can be transformed to the estimates of the optimal solution (for more details see e. g. [11], [34] and [35]).

Furthermore, it follows from [21] that the achieved results can be applied to some nonlinear functionals, covering some risk measures and to the multistage stochastic programming problems with "underlying" autoregressive (or at least Markov) sequences (see e. g. [19]). However the investigation in these directions is beyond the scope of this paper.

References

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

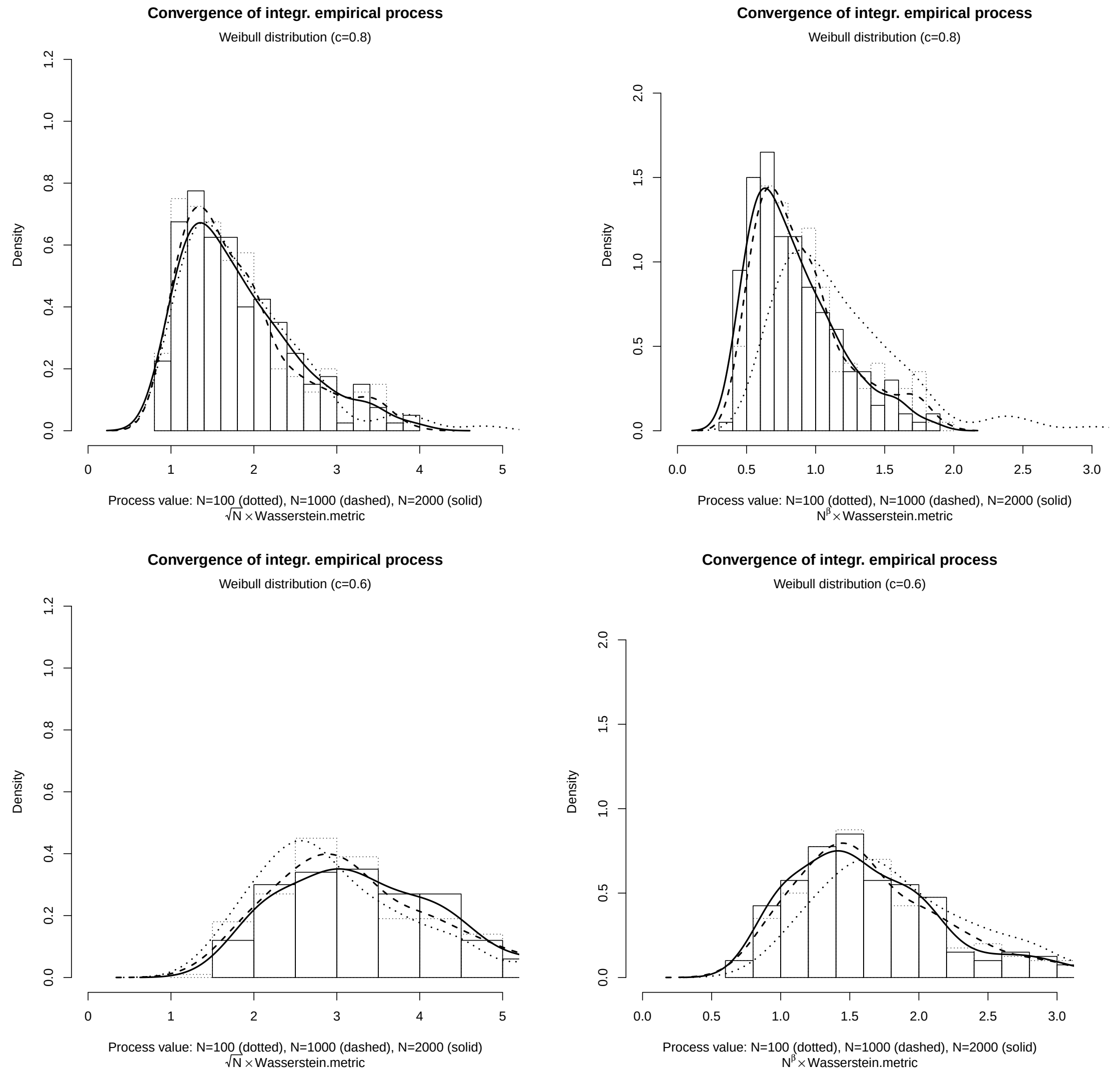

Figure 3: Weibull distribution ($c = 0.8, 0.6$) with $\beta = 1/2$ (left) and $2/5$ (right)

[2] M. Branda: Solving Real–Life Portfolio Problem using Stochastic Programming and Monte–Carlo Techniques. In: *Proceedings of the 28th International Conference on Mathematical Methods in Economics 2010* (M. Houda and J. Friebelová, eds.), University of South Bohemia, České Budějovice (Czech Republic) 2010, 24–28.

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

[5] J. Dupačová: Multistage stochastic programs: the state–of–the–art and selected bibliography. *Kybernetika* **31** (1995), 2, 151–174.

[6] 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.

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

[8] A. Gut: *Probability: A Graduate Course.* Springer, New York 2005.

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

[10] T. Homen–de–Mello: On rates of convergence for stochastic optimization problems under non–independent and identically distributed samples. *SIAM J. Optimization* **19** (2008), 524–521.

[11] M. Houda: *Stability and Approximations for Stochastic Programs.* Doctoral Thesis, Faculty of Mathematics and Physics, Charles University Prague, Prague 2009.

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

[16] V. Kaňková and M. Houda: Dependent samples in empirical estimation of stochastic programming problem. *Austrian Journal of Statistics* **35** (2006), 2&3, 271–279.

[17] 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.

[18] V. Kaňková: Empirical Estimates via Stability in Stochastic Programming. Research Report ÚTIA AS CR No. 2192, August 2007.

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

[20] V. Kaňková: A Remark on Empirical Estimates via Economic Problems. In: *Proceedings of the 27th International Conference on Mathematical Methods in Economics 2010* (H. Brožová and R. Kvasnička, eds.), Czech University of Life Sciences, Prague 2009, 169–173.

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

[22] 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.

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

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

[25] S. Kotz, N. Balakrishnan, and N. L. Johnson: *Continuous Multivariate Distributions (Volume 1: Models and Applications).* John Wiley & Sons, New York 2000.

[26] T. J. Kozubowski, A. K. Panorska and S. T. Rachev: Statistical Issues in Modelling Stable Portfolios. In: *Handbook of Heavy Tailed Distributions in Finance* (S. T. Rachev, ed.), Elsevier, Amsterdam 2003, 131–168.

[27] D. Kuhn: *Generalized Bounds for Convex Multistage Programs.* Springer, Berlin 2005.

[28] B. B. Mandelbort: Heavy Tails in Finance for Independent or Multifractal Price Increments. In: *Handbook of Heavy Tailed Distributions in Finance* (S. T. Rachev, ed.), Elsevier, Amstedam, 2003.

[29] 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.

[30] 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, 595–640 .

[31] V. Omelchenko: *Stable Distributions and Application to Finance.* Diploma Thesis, Faculty of Mathematics and Physics, Charles University Prague, Prague 2007.

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

[33] S. Rachev and S. Mitting: *Stable Paretian Models in Finance (Series in Financial Economics and Quantitative Analysis).* John Wiley & Sons, Chichester 2000.

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

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

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

[37] A. Shapiro: Monte Carlo Sampling Methods. In: *Stochastic Programming* (A. Ruszczynski and A. A. Shapiro, eds.) Handbooks in Operations Research and Managemennt Science, Volume 10, Elsevier, Amsterodam 2003, 353–426.

[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)*. Society for Industrial and Applied Mathematics and Mathematical Programming Society, Philadelphia 2009.

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

[41] 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.

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

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

[44] H. Xu: Uniform exponential analysis of sample average random functions under general sampling with applications in stochastic programming. *J. Math. Anal. Appl.* **368** (2010), 692–710.