Empirical Estimates in Stochastic Optimization: Special Cases

Vlasta Kaňková
Department of Econometrics
Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic

Abstract

"Classical" optimization problems depending on a probability measure (and corresponding to many applications) belong mostly to a class of nonlinear deterministic optimization problems that are (from the numerical point of view) relatively complicated. On the other hand, these problems fulfill very often "suitable" mathematical properties guaranteing the stability (w.r.t probability measure) and moreover giving a possibility to replace the "underlying" probability measure by an empirical one to obtain "good" stochastic estimates of the optimal value and the optimal solution. Properties of these (empirical) estimates have been studied mostly for standard types of the "underlying" probability measures with suitable (thin) tails and independent random samples. However, it is known that probability distributions with heavy tails correspond to many economic problems better (see e.g. [18]) and, moreover, many applications do not correspond to the "classical" above mentioned problems.

The aim of the paper is, first, to recall stability results in the case of heavy tails, furthermore, to introduce more general problems for which above mentioned results can be employed too.

Key words

Stochastic programming problems, stability, $\mathcal{L}_1$ norm, Lipschitz property, empirical estimates, convergence rate, exponential tails, heavy tails, Pareto distribution, risk functionals, multistage stochastic programming problems, multiobjective stochastic programming problems, stochastic linear programming problems with recourse, multistage multiobjective stochastic programming problems

1 Introduction

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

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

Find

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

Since in applications very often the probability measure P_F has to be replaced by empirical one, the solution of the problem (1) has to be (very often) sought w.r.t. an "empirical problem":

Find

$$\varphi(F^N) = \inf\{\mathsf{E}_{F^N} g_0(x, \xi) | x \in X\}, \tag{2}$$

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 it is known, that under rather general assumptions $\varphi(F^N)$, $\mathcal{X}(F^N)$ are "good" stochastic estimates of $\varphi(F)$, $\mathcal{X}(F)$.

The investigation of these empirical estimates started already in 1974 by R. Wets (see [23]). In the same time the consistency (under only an ergodic assumption) has been investigated in [6]. The investigation of the convergence rate started in [7]. Of course, the first results have been achieved for independent random samples and they have been based on the Hoeffding's paper [5] (Chernoff inequality). Let us recall the first result about the convergence rate.

Theorem 1. [7] Let $t > 0$, X be a nonempty compact set. If

1. $g_0(x, z)$ is
   a. a uniformly continuous, bounded function on $X \times Z_F$,
   b. a Lipschitz function on X with the Lipschitz constant L not depending on $z \in Z_F$,
2. $\{\xi^i\}_{i=1}^N$ for $N = 1, 2, \ldots$ is an independent random sample corresponding to P_F,

then

$$P\{\omega : |\varphi(F) - \varphi(F^N)| > t\} \leq K(t, X) \exp\{-Nk_1 t^2\}, \quad K(t, X), k_1 > 0 \quad \text{constants.}$$

L. Dal, C. H. Chen and J. R. Birge [2] have tried to generalize the assertion of Theorem 1 (for $s = 1$) to the case when

$$\mathsf{E}_F \exp\{\theta\xi\} < \infty \quad \text{for} \quad 0 \leq \theta \leq \theta_0, \quad \theta_0 \quad \text{constant.} \tag{3}$$

Evidently, the relation (3) can be fulfilled only for F with thin tails. However, the assumption of "thin" tails is not fulfilled in many applications. Relatively a detailed analysis about heavy tailed distributions (in economics and finance) is presented in [18]. Moreover, a relationship between the stable distribution (for definition see e.g. [14]) and heavy tailed distributions can be there found. A relationship between the stable distribution functions and the Pareto tails is mentioned e.g. in [14] and [17].

Furthermore, it follows from the relation (1) that the objective function is assumed there to be a linear functional of the probability measure P_F; the objective is a real valued function and, moreover, the problem is one–stage. Evidently, many problems (corresponding to the applications) do not fulfil these assumptions. In this contribution we introduce some types of problems not fulfilling these assumptions, however, for which a little modified corresponding results can be employed.

2 Some Definitions and Auxiliary Assertions

2.1 Stability Assertions

First, we recall some stability results achieved for the problem (1). To this end let $\mathcal{P}(R^s)$ denote the set of all Borel probability measures on R^s, $s \geq 1$. We set

$$\mathcal{M}_1(R^s) = \{P \in \mathcal{P}(R^s) : \int\limits_{R^s} \|z\|_s^1 P(dz) < \infty\}, \quad \|\cdot\|_s^1 \quad \text{denotes the} \quad \mathcal{L}_1 \quad \text{norm in} \quad R^s.$$

Proposition 1.[9] Let $X \subset R^n$ be nonempty compact set, $P_F, P_G \in \mathcal{M}_1(R^s)$. If

1. $g_0(x, z)$ is a Lipschitz function of $z \in R^s$ with the Lipschitz constant L (corresponding to $\mathcal{L}_1$ norm) not depending on $x \in X$,
2. finite $\mathsf{E}_F g_0(x, \xi)$, $\quad \mathsf{E}_G g_0(x, \xi)$ exist for every $x \in X$,
3. $g_0(x,\, z)$ is a uniformly continuous function on $X \times R^s$,

then

$$|\inf_{x \in X} \mathsf{E}_F g_0(x, \xi) - \inf_{x \in X} \mathsf{E}_G g_0(x, \xi)| \,\leq\, L \sum_{i=1}^{s} \int_{-\infty}^{+\infty} |F_i(z_i) - G_i(z_i)| dz_i.$$

Evidently, replacing the distribution function G by an empirical F^N we can investigate the convergence rate of the empirical estimates $\varphi(F^N)$, $\mathcal{X}(F^N)$.

2.2 Empirical Estimates

The results of this subsection will be introduced for a special case $s = 1$. First, we recall the assertion published in [22] (pp 66).

Lemma 1. [22]. Let $s = 1$, $P_F \in \mathcal{M}_1(R^1)$. Let, moreover, $\{\xi^i\}_{i=1}^{\infty}$ be a sequence of independent random values with a common distribution function F. If F^N is the empirical distribution function determined by $\{\xi^i\}_{i=1}^{N}$, $N = 1, 2, \ldots,$ then

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

The following assertion follows from the results that have been proven in [10].

Proposition 2. Let $s = 1$, $t > 0$. If

1. P_F is absolutely continuous with respect to the Lebesgue measure on R^1,
2. there exists $\psi(N,\, t) := \psi(N,\, t,\, R)$ such that the empirical distribution function F^N fulfils for $R > 0$ the relation

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

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

$$\begin{aligned} &P\{\omega : \int_{-\infty}^{\infty} |F(z) - F^N(z)| dz \,>\, t\} \,\leq \\ &(\tfrac{12R}{t} + 1)\psi(N,\, \tfrac{t}{12R},\, R) \,+ P\{\omega : \int_{-\infty}^{-R} F(z) dz > \tfrac{t}{3}\} + \\ &P\{\omega : \int_{R}^{\infty} (1 - F(z)) dz > \tfrac{t}{3}\} \,+\, 2NF(-R) \,+\, 2N(1 - F(R)). \end{aligned}$$

Proposition 3. [10]. Let $s = 1$, $t > 0$. Let, moreover $\{\xi^i\}_{i=1}^{\infty}$ be a sequence of independent random values with a common distribution function F. If

1. P_F is absolutely continuous with respect to the Lebesgue measure on R^1,
(we denote by f the probability density corresponding to F),

2. there exist constants $C_1, C_2 > 0$ and $T > 0$ such that

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

then

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

Applying the assertion of Proposition 2 to the Pareto distribution we can obtain also useful (from economic point of view) assertions. Of course, they are weaker and they depend on the "stability" coefficient α of the Pareto distribution. According to [16] there exist a few slightly different definitions of a univariate Pareto distribution; there exists relationship between them. To recall this type of distributions we employ the definition from [18]. According to it the random value ξ has a Pareto distribution if

$$P\{\omega : \xi > z\} = Cz^{-\alpha}, \qquad f(z) = \begin{array}{ll} C\alpha z^{-\alpha-1} & \text{for} \quad z > C^{\frac{1}{\alpha}}, \\ 0 & \quad\quad\;\; z \le C^{\frac{1}{\alpha}}, \end{array} \tag{4}$$

where $C > 0$ is a constant and $f(z)$ is a probability density.

Evidently, we can see that for $\alpha > 1$ we obtain $P_F \in \mathcal{M}_1(R^1)$. Furthermore, for $\beta \in (0, \frac{1}{2})$ and $R := R(N) = N^{\gamma}$, $\gamma > 0$ it holds that

$$\begin{array}{l} N^{\beta} \int\limits_{R(N)}^{\infty} [1 - F(z)]dz = N^{\beta}[C(-\alpha+1)z^{-\alpha+1}]_{R(N)}^{\infty} = -C(-\alpha+1)N^{\beta}N^{\gamma(-\alpha+1)}, \\ N[1 - F(R(N))] = NCN^{-\alpha\gamma} = CN^{1-\alpha\gamma}. \end{array} \tag{5}$$

Consequently we can obtain the following assertion.

Proposition 4. 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, $\{\xi^i\}_{i=1}^{\infty}$ be an independent random sample corresponding to the (Pareto) distribution function F. If

1. P_F is absolutely continuous with respect to the Lebesgue measure on R^1 (we denote by f the probability density corresponding to F),

2. there exists constants $C > 0$, $T > 0$ such that

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

then

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

Evidently, the assumptions of Proposition 4 can be fulfilled only for $\alpha > 2$.

2.3 Bivariate Pareto Distributions

A few definitions of slightly different univariate Pareto probability distributions have been introduced in [16]. The Pareto(I) distribution (introduced there) is very similar to the definition corresponding to the relation (4).The exact form of the definition from [16] is following.

Definition. The random value ξ is said to have a univariate Pareto(I) distribution if

$$P\{\xi > z\} = (\frac{z}{\sigma})^{-\alpha} \quad \text{for} \quad z \geq \sigma, \, \alpha > 0.$$

σ is referred as the Pareto index of inequality. The distribution is denoted by $P(I)(\sigma, \alpha)$.

The random element is mostly (in optimization problems) represented by an s–dimensional random vector ($s > 1$). In [16] a bivariate and multivariate Pareto corresponding to $P(I)(\sigma, \alpha)$ distribution are introduced. We recall the bivariate case only.

Definition. The random two dimensional vector $\xi = (\xi_1, \xi_2)$ with the joint density function

$$\begin{aligned} p_{\xi_1,\xi_2}(z_1, z_2) \;=\;& (a+1)a(\theta_1\theta_2)^{a+1}(\theta_2 z_1 + \theta_1 z_2 - \theta_1\theta_2)^{-(a+1)}, \\ & z_1 \geq \theta_1 > 0, \, z_2 \geq \theta_2 > 0, \, a > 0 \end{aligned}$$

may be called a bivariate Pareto distribution of the first kind. Since the marginal distribution have density functions

$$p_{\xi_i}(z_i) = a\theta_i^a z_i^{-(a+1)}, \quad z_i \geq \theta_i > 0, \quad i = 1, 2,$$

that is $\xi_i =_d PI(\frac{1}{\theta_i}, a)$.

Remark. There exists a survey of Pareto distributions applications in [18]. Moreover, there is also mentioned a usual assumption that all components of s–dimensional random vector have the same parameter α. However according to an analysis in [18] this assumption is not fulfilled in many applications. Employing a reduction (from the mathematical point of view) of multivariate problem to one–dimensional marginal case, we can present some new properties of the "empirical" estimates.

3 Main Results

We shall introduce a system of the assumptions.

1. A.1 $g_0(x, z)$ is
   - a uniformly continuous function on $X \times Z_F$,
   - for every $x \in X$ a Lipschitz function of z with the Lipschitz constant L (corresponding to the $\mathcal{L}_1$ norm),

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

A.3 for every $x \in X$ there exists finite $\mathsf{E}_F g_0(x, \xi)$.

3.1 Consistency

Theorem 2. Let X be a nonempty compact set, the assumptions A.1, A.2 and A.3 be fulfilled. If $P_{F_i} \in \mathcal{M}_1(R^1)$, $i = 1, \ldots, s$, then

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

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

Since, first moment exists in the case of many univariate stable distributions with heavy tails, we can see that $\varphi(F^N)$ (under rather general conditions) is a consistent estimate of $\varphi(F)$. More general assumptions guaranteing consistency are introduced in [21]. However, it is much simple to verify our assumptions.

3.2 Convergence Rate

Theorem 3. [10] Let $t > 0$, X be a compact set. Let, moreover, the assumptions A.1, A.2 and A.3 be fulfilled. If

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

2. there exist constants $C_1, C_2 > 0$ and $T > 0$ such that

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

then

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

Theorem 4. Let $t > 0$, $\alpha_i > 1$, $i = 1, \ldots, s$, X be a compact set. Let, moreover, the assumptions A.1, A.2 and A.3 be fulfilled. If

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

2. there exists a constants $C_i > 0$, $i = 1, \ldots, s$ and $T > 0$ such that

$$f_i(z) \le C_i \alpha_i z_i^{-\alpha_i - 1} \quad \text{for} \quad z \in (-\infty, -T) \cup (T, \infty), \quad i = 1, \ldots, s,$$

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

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

then

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

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

4 Application to More General Problems

Furthermore, we introduce some types of problems that do not fulfilled the assumptions of problem (1), however, for which the former presented results can be employed.

4.1 Portfolio Selection Problem

A simple "underlying" classical portfolio problem is known as the following one:

Find

$$\max \sum_{k=1}^{n} \xi_k x_k \quad \text{subject to} \sum_{k=1}^{n} x_k \le 1, \quad x_k \ge 0, \quad k = 1, \ldots, n, \quad s = n, \tag{6}$$

where x_k is a fraction of the unit wealth invested in the asset k, $\xi_k x_k$ denotes the return of the value x_k invested in the asset $k \in \{1, 2, \ldots n\}$.

If ξ_k, $k = 1, \ldots, n$ are supposed to be random, then surely it is reasonable (and quite usual) to set to the portfolio selection two–objective optimization problem:

Find

$$\begin{array}{ll} & \max \sum\limits_{k=1}^{n} \mu_k x_k, \quad \min \sum\limits_{k=1}^{n} \sum\limits_{j=1}^{n} x_k c_{k,j} x_j \\ \text{subject to} & \sum\limits_{k=1}^{n} x_k \le 1, \quad x_k \ge 0, \quad k = 1, \ldots, n, \end{array} \tag{7}$$

where $\mu_k = \mathsf{E}_F \xi_k, \quad c_{k,j} = \mathsf{E}_F(\xi_k - \mu_k)(\xi_j - \mu_j), \quad k, j = 1, \ldots n.$

Remark. It is mentioned in [1] that the efficient points of the problem (7) and the problem;

Find

$$\begin{array}{ll} & \max \sum\limits_{k=1}^{n} \mu_k x_k, \quad \min \sqrt{\sum\limits_{k=1}^{n} \sum\limits_{j=1}^{n} x_k c_{k,j} x_j} \\ \text{subject to} & \sum\limits_{k=1}^{n} x_k \le 1, \quad x_k \ge 0, \quad k = 1, \ldots, n \end{array} \tag{8}$$

are the same. (For the definition of efficient points of the multicriteria problems see e.g. [4].)

Furthermore, evidently,

$$\sigma^2(x) = \sum_{k=1}^{n} \sum_{j=1}^{n} x_k c_{k,j} x_j = \mathsf{E}_F \{ \sum_{j=1}^{n} \xi_j x_j - \mathsf{E}_F [\sum_{j=1}^{n} \xi_j x_j] \}^2, \; x = (x_1, \, \ldots, \, x_n)$$

can be considered as a risk measure. Konno and Yamazaki recalled in [15] another risk measure

$$w(x) = \mathsf{E}_F | \sum_{k=1}^{n} \xi_k x_k - \mathsf{E}_F [\sum_{k=1}^{n} \xi_k x_k] |. \tag{9}$$

Moreover, they have recalled that $w(x) = \sqrt{\frac{2}{\pi}} \sigma(x)$ in the case of mutually normally distributed random vector $(\xi_1, \, \ldots, \, \xi_n)$. Other risk measures can be considered in the form:

$$w^+(x) = \mathsf{E}_F | \sum_{k=1}^{n} \xi_k x_k - \mathsf{E}_F [\sum_{k=1}^{n} \xi_k x_k] |^+, \quad w^-(x) = \mathsf{E}_F | \sum_{k=1}^{n} \xi_k x_k - \mathsf{E}_F [\sum_{k=1}^{n} \xi_k x_k] |^-. \tag{10}$$

Employing a "weight" approach (first employed by Markowitz) we obtain the following one objective problems:

$$\begin{array}{l} \max\limits_{x \in X} [\sum\limits_{k=1}^{n} \mu_k x_k - K \mathsf{E}_F | \sum\limits_{k=1}^{n} \xi_k x_k - \mathsf{E}_F [\sum\limits_{k=1}^{n} \xi_k x_k] |, \\ \max\limits_{x \in X} [\sum\limits_{k=1}^{n} \mu_k x_k - K \mathsf{E}_F | \sum\limits_{k=1}^{n} \xi_k x_k - \mathsf{E}_F [\sum\limits_{k=1}^{n} \xi_k x_k] |^+, \\ \max\limits_{x \in X} [\sum\limits_{k=1}^{n} \mu_k x_k - K \mathsf{E}_F | \sum\limits_{k=1}^{n} \xi_k x_k - \mathsf{E}_F [\sum\limits_{k=1}^{n} \xi_k x_k] |^-, \\ X \, := \, \bar{X} = \{ x \in R^n : \sum\limits_{k=1}^{n} x_k \le 1, \quad x_k \ge 0, \quad k = 1, \ldots, n, \}, \quad K > 0. \end{array} \tag{11}$$

To introduce more general problems covering problems of the type (1), let $h(x, \, z) = (h_1(x, \, z), \, \ldots, \, h_{m_1}(x, \, z))$ be m_1–dimensional vector function defined on $R^n \times R^s$, $g_0^1(x, \, z, \, y)$ be a real valued function defined on $R^n \times R^s \times R^{m_1}$. If we replace the problem (1) by a stochastic programming problem:

Find

$$\varphi(F) := \varphi^1(F) = \inf \{ \mathsf{E}_F g_0^1(x, \, \xi, \mathsf{E}_F h(x, \, \xi)) | x \in X \}, \tag{12}$$

then to investigate the stability and empirical estimates of the problem (12), evidently, we can employ a stability approach employed in [19] (pp. 449). We recall the following assertion.

Proposition 4. [10] Let $X \subset R^n$ be nonempty compact set, $P_F, P_G \in \mathcal{M}_1(R^s)$. If

1. $h_i(x, z)$, $i = 1, \ldots, m_1$ are for every $x \in X$ Lipschitz functions of z with the Lipschitz constants L_h^i (corresponding to $\mathcal{L}_1$ norm),
2. $g_0^1(x, z, y)$ is for $x \in X$, $y \in R^{m_1}$ a Lipschitz function of $z \in R^s$ with the Lipschitz constant $L^z(x, y)$ (corresponding to $\mathcal{L}_1$ norm) and for every $x \in X$, $z \in Z_F$ a Lipp. fumction of y with Lips. constant L^y,

then there exists a constant $C > 0$ such that for every $x \in X$

$$|\inf_X \mathsf{E}_F g_0^1(x, \xi, \mathsf{E}_F h(x, \xi)) - \inf_X \mathsf{E}_G g_0^1(x, \xi, \mathsf{E}_G h(x, \xi))| \leq$$
$$\{ \mathsf{E}_F[L^y(x, \xi)]\, C \sum_{i=1}^{m_1} L_h^i + L^z(x, \mathsf{E}_G h(x, \xi)) \} \sum_{j=1}^{s} \int_{-\infty}^{\infty} |F_j(z_j) - G_j(z_j)| dz_j.$$

4.2 Stochastic Linear Programming with Recourse

The following auxiliary assertion can be useful for stochastic linear programming problems with recourse. To this end we consider the case $s = 2$ and we set $\xi_1 := \bar{\xi}$, $\xi_2 := \bar{\eta}$, where $\bar{\xi}(:= \bar{\xi}(\omega))$, $\bar{\eta}(:= \bar{\eta}(\omega))$ are random values defined on $(\Omega, \mathcal{S}, P)$ with finite second moments. If we denote by the symbols $F(:= F_{(\bar{\xi}, \bar{\eta})}), F_{\bar{\xi}}, F_{\bar{\eta}}$ the distribution functions of the random vector $(\bar{\xi}, \bar{\eta})$ and marginal distribution functions of the random valuables $\bar{\xi}$ and $\bar{\eta}$, then we can recall the following auxiliary assertions.

Proposition 5. [12] Let $\bar{\zeta} = \bar{\xi}\bar{\eta}\,(:= \bar{\xi}(\omega)\bar{\eta}(\omega))$, $F_{\bar{\zeta}}$ denote the distribution function of $\bar{\zeta}$. If

1. $P_{F_{\bar{\xi}}}, P_{F_{\bar{\eta}}}$ are absolutely continuous with respect to the Lebesgue measure on R^1 (we denote by $f_{\bar{\xi}}, f_{\bar{\eta}}$ the probability densities corresponding to $F_{\bar{\xi}}, F_{\bar{\eta}}$),
2. there exist constants $C_1^{\bar{\xi}}, C_2^{\bar{\xi}}, C_1^{\bar{\eta}}, C_2^{\bar{\eta}} > 0$ and $T > 0$ such that

$$f_{\bar{\xi}}(z) \leq C_1^{\bar{\xi}} \exp\{-C_2^{\bar{\xi}}|z|\} \quad \text{for} \quad z \in (-\infty, -T) \bigcup (T, \infty),$$
$$f_{\bar{\eta}}(z) \leq C_1^{\bar{\eta}} \exp\{-C_2^{\bar{\eta}}|z|\} \quad \text{for} \quad z \in (-\infty, -T) \bigcup (T, \infty),$$

then, there exist constants $C_1^{\bar{\zeta}}, C_2^{\bar{\zeta}} > 0$, $D_1, D_2, D_3 > 0$, $\bar{T} > 1$ such that

$$\begin{aligned} &F_{\bar{\zeta}}(-z) \leq \frac{C_1^{\bar{\zeta}}}{C_2^{\bar{\zeta}}} \exp\{-C_2^{\bar{\zeta}}\sqrt{z}\}, \quad (1 - F_{\bar{\zeta}}(z)) \leq \frac{C_1^{\bar{\zeta}}}{C_2^{\bar{\zeta}}} \exp\{-C_2^{\bar{\zeta}}\sqrt{z}\} \quad \text{for} \quad z > \bar{T} \\ &\int_z^{+\infty} (1 - F_{\bar{\zeta}}(u))du \leq D_1\sqrt{z}e^{-D_2\sqrt{z}} + D_3 e^{-D_2\sqrt{z}} \quad \text{for} \quad z > \bar{T}. \end{aligned} \tag{13}$$

4.3 Multiobjective Stochastic Programming Problems

A rather general multiobjective stochastic programming problem can be introduced as the following problem [8]:

Find

$$\inf \mathsf{E}_F g_i(x,\, \xi), \quad i = 1,\, \ldots,\, l \quad \text{subject ro} \quad x \in X, \tag{14}$$

where $g_i(x,\, z)$, $i = 1,\, \ldots,\, l$ are functions defined on $R^n \times R^s$.

It is known from the multiobjective optimization theory (see e.g. [4]) that the problem (14) can be in the convex case (under some additional assumptions) solved by one criterion parametric optimization problem:

Find

$$\inf_{x \in X} \mathsf{E}_F \sum_{i=1}^{l} \lambda_i g_i(x,\, \xi), \quad \sum_{i=1}^{l} \lambda_i = 1, \quad \lambda_i > 0,\, i = 1,\, \ldots,\, l. \tag{15}$$

4.4 Multistage Stochastic Programming Problems

Multistage (generally nonlinear) stochastic problems. It follows, from a recursive definition, that in the case of autoregressive random sequence or at least Markov sequence, the results achieved for one–stage problem can be employed for the multistage problems. Namely, the multistage stochastic programming problem is (according to this approach) introduced as a finite system of parametric (one–stage) optimization problems with an inner type of dependence (for more details see e.g. [3] or [11]). The multistage stochastic (generally nonlinear) programming problem is then introduced in the form:

Find

$$\varphi_{\mathcal{F}}(M) = \inf \{\mathsf{E}_{F^{\xi^0}} g^0_{\mathcal{F}}(x^0,\, \xi^0) |\; x^0 \in \mathcal{K}^0\}, \tag{16}$$

where the function $g^0_{\mathcal{F}}(x^0,\, z^0)$ is defined recursively

$$\begin{aligned} g^k_{\mathcal{F}}(\bar{x}^k,\, \bar{z}^k) \quad &= \quad \inf\{\mathsf{E}_{F^{\xi^{k+1}|\bar{\xi}^k = \bar{z}^k}}\; g^{k+1}_{\mathcal{F}}(\bar{x}^{k+1},\, \bar{\xi}^{k+1})\, | x^{k+1} \in \mathcal{K}^{k+1}(\bar{x}^k,\, \bar{z}^k)\}, \\ k = 0,\, 1,\, \ldots,&\, M-1 \\ g^M_{\mathcal{F}}(\bar{x}^M,\, \bar{z}^M) \quad &:= \quad g^M_0(\bar{x}^M,\, \bar{z}^M), \quad \mathcal{K}_0 := X^0. \end{aligned} \tag{17}$$

$\xi^j := \xi^j(\omega)$, $j = 0,\, 1,\, \ldots,\, M$ denotes an s–dimensional random vector defined on a probability space $(\Omega,\, \mathcal{S},\, P)$; $F^{\xi^j}(z^j)$, $z^j \in R^s$, $j = 0,\, 1\, \ldots,\, M$ the distribution function of the ξ^j and $F^{\xi^k|\bar{\xi}^{k-1}}(z^k|\bar{z}^{k-1})$, $z^k \in R^s$, $\bar{z}^{k-1} \in R^{(k-1)s}$, $k = 1,\, \ldots,\, M$ the conditional distribution function (ξ^k conditioned by $\bar{\xi}^{k-1}$); $P_{F^{\xi^j}}$, $P_{F^{\xi^{k+1}|\bar{\xi}^k}}$, $j = 0,\, 1,\, \ldots,\, M$, $k = 0,\, 1,\, \ldots,\, M-1$ the corresponding probability measures. Furthermore, $g^M_0(\bar{x}^M,\, \bar{z}^M)$ denotes a continuous function defined on $R^{n(M+1)} \times R^{s(M+1)}$; the symbol $\mathcal{K}^{k+1}(\bar{x}^k,\, \bar{z}^k)$, $k = 0,\, 1,\, \ldots,\, M-1$ denotes a multifunction mapping $R^{n(k+1)} \times R^{s(k+1)}$ into the space of subsets of R^n. $\bar{\xi}^k (:= \bar{\xi}^k(\omega)) = [\xi^0,\, \ldots,\, \xi^k]$; $\bar{z}^k = [z^0,\, \ldots,\, z^k]$, $z^j \in R^s$; $\bar{x}^k = [x^0,\, \ldots,\, x^k]$, $x^j \in R^n$. Symbols $\mathsf{E}_{F^{\xi^0}}$, $\mathsf{E}_{F^{\xi^{k+1}|\bar{\xi}^k = \bar{z}^k}}$, $k = 0,\, 1,\, \ldots,\, M-1$ denote the operators of mathematical expectation corresponding to F^{ξ^0}, $F^{\xi^{k+1}|\bar{\xi}^k = \bar{z}^k}$, $k = 0,\, \ldots,\, M-1$.

The problem (16) is a "classical" one–stage, one–objective stochastic problem, the problems (17) are (generally) parametric one–stage, one–objective stochastic optimization problems.

4.5 Multistage Mlltiobjective Stochastic Programming Problems

Assuming the same inner time dependence as it was assumed in the problem (15) and (16), we obtain formally the following multistage, multiobjective problem:

Find

$$\varphi^i_{\mathcal{F}}(M) = \inf\, \mathsf{E}_{F^{\xi^0}} g^{0,\, i}_{\mathcal{F}}(x^0,\, \xi^0),\; i = 1,\, \ldots,\, l \quad \text{subject to} \quad x^0 \in \mathcal{K}^0, \tag{18}$$

where the function $g^{0,\,i}_{\mathcal{F}}(x^0,\, z^0),\ i = 1,\, \ldots,\, l$ are defined recursively

$$
\begin{aligned}
g^{k,\,i}_{\mathcal{F}}(\bar{x}^k,\, \bar{z}^k) \quad &= \quad \inf \mathsf{E}_{F^{\xi^{k+1}|\bar{\xi}^k=\bar{z}^k}}\, g^{k+1,\,i}_{\mathcal{F}}(\bar{x}^{k+1},\, \bar{\xi}^{k+1}),\ i = 1,\, \ldots,\, l \\
& \qquad \text{subject to} \quad x^{k+1} \in \mathcal{K}^{k+1}(\bar{x}^k,\, \bar{z}^k), \quad k = 0,\, 1,\, \ldots,\, M-1, \\
g^{M,\,i}_{\mathcal{F}}(\bar{x}^M,\, \bar{z}^M) \quad &:= \quad \sum_{j=0}^{M} \bar{g}^j_i(x^j,\, z^j), \quad i = 1,\, \ldots,\, l, \quad \mathcal{K}_0 := X^0.
\end{aligned}
\tag{19}
$$

$g^j_i(x^j,\, z^j),\ i = 1,\, \ldots,\, l,\ j = 0,\, 1,\, \ldots,\, M$ are real valued functions defined on $R^n \times R^s$.

The problem (18) is a problem of one–stage multiobjective optimization theory. The problems (19) are one–stage multiobjective parametric optimization problems. However, it is known that there doesn't exist (mostly) an optimal solution simultaneously with respect to all criteria. Consequently, the optimal solution has to be mostly replaced by a set of efficient points. For more details and other notions see [13].

5 Conclusion

We have recalled, in section 3, the convergence rate for empirical estimates in the case of the problem (1). In section 4 we have introduced some special types of problems that are not covered by the type of the problem (1), however, it follows from the analysis presented in the section 4 that under a little modified assertion can be employed too.

Acknowledgement

This work was supported by the Grant Agency of the Czech Republic under grants P402/10/0956, 402/08/0107 and 402/06/0990.

References

[1] R. Caballero, E. Cerdá, M. M. Muñoz, L. Rey and I. M. Stancu–Minasian: Afficient Solution Concepts and their Relations in Stochastic Multiobjective Programming. Journal of Optimization Theory and Applications *110* (2001), 1, 53–74.

[2] L. Dai, C. H. Chen, and J. R. Birge: Convergence Properties of Two–Stage Stochastic Programming. J. Optim. Theory Appl. *106* (2000), 489–509.

[3] J. Dupačová: Multistage Stochastic Programs: the State–of–the–Art and Selected Bliography. Kybernetika *31* (1995),2, 151–174.

[4] M. Ehrgott: Multicriteria Optimization. Springer, Berlin 2005.

[5] W. Hoeffding: Probability Inequalities for Sums of Bounded Random Variables. Journal of Americ. Statist. Assoc. *58* (1963), 301, 13–30.

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

[7] V. Kaňková: An Approximative Solution of Stochastic Optimization Problem. In: Trans. 8th Prague Conf. 1977. Academia, Prague 1978, 349–353.

[8] V. Kaňková: A Note on the Relationship between Strongly Convex Functions and Multiobjective Stochastic Programming Problems. In: Operatios Resarch Proceedings 2004. (H. Fleuren, D. den Hertog and P. Kort, eds.) Springer, Berlin 2005.

[9] 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.

[10] V. Kaňková: Empirical Estimates via Stability in Stochastic Programming. Research Report UTIA 2007, No 2192, Prague 2007.

[11] V. Kaňková: Multistage Stochastic Programs via Autoregressive Sequences and Individual Probability Constraints. Kybernetika *44* (2008), 2, 151–170.

[12] V. Kaňková: A Remark on Nonlinear Functionals and Empirical Estimates. In: Proceedings of Quantitative Methods in Economics (Multiple Criteria Decision Making XIV). The Slovak Society for Operations Research and University of Economics in Bratislava 2008,124–133.

[13] V. Kaňková: Multiobjective Stochastic Programming via Multistage Problems. In: Proceedins of Mathematical Methods in Economics 2008 (P. Řehořová and K. Maršíková, eds.), Faculty of Economics, Technical University of Liberec (Czech Republic) 2008, 50250–256, CD–ROM.

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

[15] H. Konno and H. Yamazaki: Mean–Absolute Deviation Portfolio Optimization Model and Its Application to Tokyo Stock Markt. Magement Science *37* (1991), 5, 519–531.

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

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

[18] 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, ae.) Elsevier, Amsterdam 2003, 595–640.

[19] G. Ch. Pflug: Stochastic Optimization and Statistical Inference. In: *Stochastic Programming* (A. Ruszczynski and A. A. Shapiro, eds.) Handbooks in Operations Research and Managemennt Science, Vol 10, Elsevier, Amsterdam 2003, 427–480.

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

[21] W. Römisch: Stability of Stochastic Programming Problems. In: Stochastic Programming (A. Ruszczynski and A. Shapiro, eds.) Handbooks in Operations Research Management Science, Vol10, Elsevier, Amsterdam 2003, 483–554.

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

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