A REMARK ON EMPIRICAL ESTIMATES VIA ECONOMIC PROBLEMS

Vlasta Kaňková

Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
kankova@utia.cas.cz

Abstract. Optimization problems depending on a probability measure correspond to many economic applications. Since the "underlying" measure is usually unknown the decision is mostly determined on the date basis, it means on statistical (mostly empirical) estimates of the probability measure. Properties of the optimal value (and solution) estimates have been investigated many times. There were introduced assumptions under which the asymptotic distribution estimate is normal and the convergence rate is at least exponential. We generalize the assertions concerning the rate convergence. Especially we shall consider distributions with the Pareto tail.

Keywords. Economic problems, stochastic optimization, stochastic estimates, exponential rate convergence, exponential tails, Pareto distribution.

JEL Classification: C44 AMS Classification: 90 C 15

1 Introduction

Economic processes are often influenced by a random factor and a decision parameter. Constructing mathematical models we obtain mostly models depending on a probability measure. They can be static (one stage) or dynamic. A multistage stochastic programming technique can treat an essential class of dynamic cases. Employing a recursive definition (see e.g. [9], [11]), we obtain a system of one–stage (mostly) parametric problems. Consequently, the investigation of one problems is crucial also for multistage cases.

To introduce "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 $(\Omega, \mathcal{S}, P)$; $F (:= F(z),\, z \in R^s)$ and P_F the distribution function and the probability measure corresponding to ξ; $F_i,\, i = 1, \ldots, s$ one–dimensional marginal distribution functions of $\xi_i,\, i = 1, 2, \ldots, s$. 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 set. If the symbol E_F denotes the operator of mathematical expectation corresponding to F, then a rather general "classical" one–stage stochastic programming problem can be introduced in the form:

Find

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

Since in applications very often the problem (1) has to be solved on the basis of empirical data we have to replace the measure P_F by its estimate. An empirical probability measure is a very suitable candidate for it. Consequently, the solution of the problem (1) has to be often sought (in applications) 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 (under rather general assumptions) $\varphi(F^N)$, $\mathcal{X}(F^N)$ are "good" stochastic estimates of $\varphi(F)$, $\mathcal{X}(F)$.

The investigation of the empirical (above introduced) estimates started in 1974 by [16]; followed by many papers (see e.g. [3], [15]). The investigation of the convergence rate appeared in [5] and followed e.g. by [1], [4], [12]; for weak dependent random samples e.g. by [6]. The exponential rate convergence has been proven under some relatively strong assumptions on the objective function and on the "underlying" distribution function F. Employing the stability results [8] we can see that the normal distribution is covered by this approach. However, the distribution functions with heavy tails correspond to many random factor in economic problems (see e.g. [13], [14]). The aim of this contribution is to genaralize results [5] to the case of distribution functions with the Pareto tail.

2 Auxiliary Assertion

2.1 Stability

To recall a suitable stability result, let $\mathcal{P}(R^s)$ denote the set of all Borel probability measures on $R^s, s \geq 1$ and let $\mathcal{M}_1(R^s)$ be defined by

$$\mathcal{M}_1(R^s) = \{P \in \mathcal{P}(R^s) : \int\limits_{R^s} \|z\|_s^1 P(dz) < \infty\},$$

$\| \cdot \|_s^1$ denotes the $\mathcal{L}_1$ norm in R^s. We shall introduce a little generalized assertion of [7].

Proposition 1. [7] Let $P_F, P_G \in \mathcal{M}_1(R^s)$, X be a compact set. If for every $x \in X$

1. $g_0(x, z)$ is a Lipschitz function of $z \in R^s$ with the Lipschitz constant L,
2. finite $\mathsf{E}_F g_0(x, \xi)$, $\quad \mathsf{E}_G g_0(x, \xi)$ exist,
3. $g_0(x,\, z)$ is a uniformly continuous function on $X \times R^s$,

then

$$|\inf_{x \in X} \mathsf{E}_F g_0(x, \xi) - \inf_{x \in X} \mathsf{E}_G g_0(x, \xi)| \leq L \sum_{i=1}^{s} \int\limits_{-\infty}^{+\infty} |F_i(z_i) - G_i(z_i)| dz_i. \tag{3}$$

Evidently, the right hand side of the relation (3) is in the form of a product of the Lipschitz constant and the Wasserstein metric. Replacing the distribution 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

Proposition 2. Let $s = 1,\, t > 0$. Let moreover $P_F \in \mathcal{M}_1(R^1)$. 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

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

If, moreover,

3. there exists $R := R(N)$ defined on $\mathcal{N}$ such that $R(N) \longrightarrow_{(N \longrightarrow \infty)} \infty$ and simultaneously for $\beta \in (0,\, \frac{1}{2})$

$$N^\beta \int\limits_{-\infty}^{-R(N)} F(z)dz \longrightarrow_{N \longrightarrow \infty} 0, \qquad N^\beta \int\limits_{R(N)}^{\infty} [1 - F(z)]dz \longrightarrow_{(N \longrightarrow \infty)} 0,$$
$$2NF(-R(N)) \longrightarrow_{(N \longrightarrow \infty)} 0, \qquad 2N[1 - F(R(N))] \longrightarrow_{(N \longrightarrow \infty)} 0, \tag{4}$$
$$(\tfrac{12R(N)N^\beta}{t} + 1)\psi(N,\, \tfrac{t}{12R(N)N^\beta},\, R(N)) \longrightarrow_{(N \longrightarrow \infty)} 0,$$

then also

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

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

Proof. The assertion can be proven by the proof technique employed in [8]. □

It is well known (see e.g. [2]) that if $\{\xi^i\}_{i=1}^{\infty}$ is a sequence of independent random values with a common distribution function F, then we can set

$$\Psi(N,\, t,\, R) = \exp\{-2Nt^2\}, \quad t > 0,\, R > 0,\, N = 1,\, 2,\, \ldots,. \tag{5}$$

Setting $R(N) = N^\gamma$, $\gamma + \beta \in (0,\, \frac{1}{2})$ we can see that the following assertion holds.

Corollary 1. Let $s = 1,\, t > 0$. Let, moreover, $P_F \in \mathcal{M}_1(R^1)$, $\{\xi^i\}_{i=1}^{\infty}$ 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 exists constants C_1, C_2 and $T > 0$ such that

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

then

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

Proof. The assertion follows from Proposition 2 and the assumptions. □

The assumption 2 of Corollary 1 covers the normal and empirical distributions. However, many random elements corresponding to economic applications correspond to distributions with heavy tails. We employ the uniform Pareto distribution introduced in [13]. A random value ξ has a Pareto distribution if its probability measure P_F and its probability density f fulfils the relation

$$\begin{aligned} P_F\{\xi > z\} = Cz^{-\alpha}, \quad f(z) &= C\alpha z^{-\alpha-1} \text{ for } z > C^{\frac{1}{\alpha}}, \\ &0 \qquad\qquad z \le C^{\frac{1}{\alpha}}. \end{aligned} \tag{6}$$

Evidently, the Pareto distribution has only one tail. Moreover, we can see that for $\alpha > 1$ it holds $P_F \in \mathcal{M}_1(R^1)$ and for $\beta \in (0, \frac{1}{2})$ and $R := R(N) = N^{\gamma}$, $\gamma > 0$ it holds

$$N^{\beta} \int\limits_{R(N)}^{\infty} [1 - F(z)]dz = N^{\beta}[C(-\alpha+1)z^{-\alpha+1}]_{R(N)}^{\infty} = -C(-\alpha+1)N^{\beta}N^{\gamma(-\alpha+1)},$$

$$N[1 - F(R(N))] = NCN^{-\alpha\gamma} = CN^{1-\alpha\gamma}.$$

Consequently for $\gamma > \max[\frac{\beta}{\alpha-1}, \frac{1}{\alpha}]$ we can obtain that

$$N^{\beta} \int\limits_{R(N)}^{\infty} [1 - F(z)]dz \longrightarrow_{(N \longrightarrow \infty)} 0, \quad \text{and, simultaneously,} \quad 2N[1 - R(N)] \longrightarrow_{(N \longrightarrow \infty)} 0.$$

Employing the relations (4), (5) we can see that the following assertion holds.

Corollary 2. 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 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 $C > 0$ and $T > 0$ such that

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

then

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

3 Main Results

In this section we shall try to introduce the results that guarantee the exponential convergence rate of $\varphi(F^N)$ to $\varphi(F)$. First result covers the known case when the tails of the probability density are at least exponential. Evidently, this assertion covers the classical case of normal distribution. The second case will try to cover some new arising economic applications when one dimensional marginal distribution functions have the Pareto tails. It is known that this case appears for example in finance or river flow (for more details see e.g. [13]). The corresponding form of multivariate case can be found e.g. in [10]. However it is over the possibility of this contribution to deal with this case in more detail.

Theorem 1. [8] Let $t > 0$, $\{\xi^i\}_{i=1}^{\infty}$ be a sequence of independent s–dimensional random vectors with a common distribution function F. Let moreover X be a compact set. If

1. F^N is an empirical distribution function determined by $\{\xi^i\}_{i=1}^{N}$, $N = 1, 2, \ldots$,
2. 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),

3. there exist constants C_1, $C_2 > 0$ and $T > 0$ such that
$$f_i(z_i) \le C_1 \exp\{-C_2|z_i|\} \quad \text{for} \quad z_i \in (-\infty, -T) \cup (T, \infty), \quad i = 1, \ldots, s,$$
4. $g_0(x, z)$ is a uniformly continuous, Lipschitz (with respect to $\mathcal{L}_1$ norm) function of $z \in R^s$, the Lipschitz constant L is not depending on $x \in X$,

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

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

Theorem 2. Let $t > 0$, $\alpha > 1$, $\beta, \gamma > 0$ 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 s–dimensional independent random sample corresponding to the distribution function F, X be a compact set. If

1. F^N is an empirical distribution function determined by $\{\xi^i\}_{i=1}^{N}$, $N = 1, 2, \ldots$,
2. 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),
3. there exist constants $C > 0$ and $T > 0$ such that
$$f_i(z) \le C\alpha z_i^{-\alpha-1} \quad \text{for} \quad z \in (-\infty, -T) \cup (T, \infty), \quad i = 1, \ldots, s,$$
4. $g_0(x, z)$ is a uniformly continuous, Lipschitz (with respect to $\mathcal{L}_1$ norm) function of $z \in R^s$, the Lipschitz constant L is not depending on $x \in X$,

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

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

4 Conclusion

The aim of the paper has been to investigate properties of the empirical estimates of the optimal value in the case of one–stage optimization problems depending on a probability measure. The introduced results are based on the stability results corresponding to the Wasserstein metric and $\mathcal{L}_1$ norm in R^s, $s \ge 1$. They do not cover only the normal distribution corresponding to many "classical" approaches in finance, however also the case of Pareto distribution. This result is crucial, namely, it is known that the distributions with "heavy" tails correspond to many new applications (for more see e.g. [10] and [13]). The achieved convergence rate is the best as possible in the case of exponential tails, in the case of Pareto distribution the introduced convergence rate is worse and depends on the parameter α.

Acknowledgement This work was supported by the Grant Agency of the Czech Republic under grants 402/07/1113, 402/08/0107 and the Grant Agency of MŠMT CR under grant LC 06075.

References

1. L. Dai, C. H. Chen, and J. R. Birge: Convergence Properties of Two–Stage Stochastic Programming. J. Optim. Theory Appl. *106* (2000), 489–509.
2. A. Dvoretzky, J. Kiefer, and J. Wolfowitz: Asymptotic minimax character of the sample distribution and the classical multinomial estimate. Ann. Math. Stat. *27* (1956) 642–669.
3. J. Dupačová and R. Wets: Asymptotic Behaviour of Statistical E stimates and Optimal Solutions of Stochastic Optimization Problems. Ann. Statist. *16* (1984), 1517–1549.
4. Y. M. Kaniovski, A King and R. Wets: Probabilistic bounds (via large deviations) for the solutions of stochastic programming problems. Annals of Oper. Res. *56* (1995), 189–208.
5. V. Kaňková: An approximative solution of stochastic optimization problem. In: *Trans. 8th Prague Conf. 1977* Academia, Prague 1978, 349–353.
6. V. Kaňková: A note on estimation in stochastic pProgramming. J Comput. Math. *56*(1994), 97–112.
7. 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.
8. V. Kaňková: Empirical Estimates via Stability in Stochastic Programming. Research Report UTIA 2007, No 2192, Prague 2007.

9. V. Kaňková: Multistage stochastic programs via autoregressive sequences and individual probability constraints. Kybernetika *44* (2008), 2, 151–170.
10. S. Kotz, N. Balakrishnan, and N. L. Johnson: Continuous Multiviariate Distributions (Volume 1: Models and Applications). Wiley, New York 2000.
11. D. Kuhn: Generalized Bounds for Convex Multistage Programs. Springer, Berlin 2005.
12. Tito Homen de Mello: On Rates of Convergence for Stochastic Optimization Problems under Non–i.i.d. Sampling. Speps 2006.
13. 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.
14. S. Rachev and S. Mitting: Stable Paretian Models in Finance (Series in Financial Economics and Quantitative Analysis). Wiley, Chichester 2000.
15. A. Shapiro: Quantitative stability in Stochastic Programming. Math. Program. *67* (1994), 99–108.
16. R. J. B. Wets: A statistical Approach to the Solution of Stochastic Programs with (Convex) Simple Recourse. Research Report, University Kentucky, USA 1974.