Empirical Estimates in Economic and Financial Problems via Heavy Tails

Vlasta Kaňková[1]

Abstract. Optimization problems depending on a probability measure correspond to many economic and financial applications. Complete knowledge of this measure is necessary to solve exactly these problems. Since this condition is fulfilled only seldom, the problem has to be usually solved on the data basis to obtain statistical estimates of an optimal value and optimal solutions. Great effort has been paid to investigate properties of these estimates; first under assumptions of distribution with thin tails and linear dependence on the probability measure. Recently, it has appeared an investigation in the case of nonlinear dependence on the probability measure and heavy tailed distributions with shape parameter greater two. We focus on the case of the stable and Pareto distributions with shape parameter in the interval (1, 2).

Keywords: Stochastic optimization problems, empirical estimates, thin and heavy tailed distributions, stable distributions, Pareto distribution, Lipschitz property.

JEL classification: C44
AMS classification: 90C15

1 Introduction

Let $\xi\,(:= \xi(\omega) = [\xi_1(\omega), \ldots, \xi_s(\omega)])$ be s–dimensional random vector defined on a probability space $(\Omega, \mathcal{S}, P)$; $F(:= F(z),\, z \in R^s)$ the distribution function of ξ; P_F the probability measure corresponding to F. Let, moreover, $g_0(:= g_0(x, z))$ be a real–valued 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

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

In applications very often the "underlying" probability measure P_F has to be replaced by empirical one. Evidently, then the solution is sought w.r.t. the "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 $\{\xi^i\}_{i=1}^N$ 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 $\varphi(F^N)$, $\mathcal{X}(F^N)$ are stochastic estimates of $\varphi(F)$, $\mathcal{X}(F)$.

The investigation of these estimates started in 1974 by R. Wets (see [27]). In the same time consistency has been investigated under ergodic assumption in [10]. These papers have been followed by many others (see e.g. [3], [14], [20]). The investigation of the convergence rate started in [11], and followed e.g. by [2], [9], [16], [23], [24],[26]. However these results have been obtained under the assumptions of an linear dependence of objective functions on the probability measure and an "underlying" distribution with "thin" tails. Later have appeared works with underlying distribution functions with heavy tails (see e.g.

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

[7], [14] and [22]). In this paper we focus on the case of the stable and Pareto distributions with shape parameter from the interval (1, 2).

2 Heavy Tails in Economic and Financial Applications

First, it has been assumed (in the stochastic optimization problems) that the "underlying" probability measure corresponds to the class of distribution functions with thin tails. In particular, the "best" convergence rate has been first proven for the case when an "underlying" objective function has been bounded (see e.g. [11]), later for the case when a finite moment generating function (corresponding to an "underlying" objective function) exits (see e.g. [24]). Consequently, the normal distribution has approximated often the real one. This assumption appeared in the Markowitz model of portfolio selection (see e.g. [4]). However, relatively soon it has been recognized that many data correspond to the distributions for which a finite moment generating function does not exist, it means to the distribution functions with "heavy" tails. This case corresponds often to economic and financial applications.

A relatively good analysis of the "heavy" tailed distribution in economy and finance is presented e.g. in [18]. 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 [8] that some data 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 [15], between the stable heavy tailed distributions and the Pareto tails is known and can be found e.g in [15] and [17]. The relationship between tails of stable distributions and Pareto tails distribution is shown in [25].

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 valid about a convergence rate and an asymptotic distribution? Some results about consistency are already known as well as some results about convergence rate. It has been proven that a "best" convergence rate is valid everywhere when all finite absolute moments exist. Weaker results have been obtained generally in dependence on the finite absolute moments existence (for more details see [7]). In this paper we focus on the case of heavy tailed distributions with a shape parameter $\alpha \in (1, 2)$. To this end we employ the stability results obtained by the Wasserstein metric corresponding to $\mathcal{L}_1$ norm.

3 Some Definition and Auxiliary Assertions

Let $F_i, i = 1, \dots, s$ denote one–dimensional marginal distribution functions corresponding to F; $\mathcal{P}(R^s)$ denote the set of Borel probability measures on $R^s, s \geq 1$; $\mathcal{M}_1(R^s) = \{P \in \mathcal{P}(R^s) : \int\limits_{R^s} \|z\|_s^1 P(dz) < \infty\}$, $\| \cdot \|_s^1$ denote $\mathcal{L}_1$ norm in R^s.

We introduce a system of assumptions:

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

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

A.3 $P_{F_i}, i = 1, \dots, s$ are absolutely continuous w. r. t. the Lebesgue measure on R^1.

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

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

Proposition 1 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.

Proposition 2 ([14])**.** *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 $\mathcal{N}$ such that $R(N) \xrightarrow[N \to \infty]{} \infty$ and, moreover,*

$$\begin{array}{ll} N^{\beta} \int\limits_{-\infty}^{-R(N)} F(z)dz \xrightarrow[N\to\infty]{} 0, & N^{\beta} \int\limits_{R(N)}^{\infty} [1 - F(z)]dz \xrightarrow[N\to\infty]{} 0, \\ 2NF(-R(N)) \xrightarrow[N\to\infty]{} 0, & 2N[1 - F(R(N))] \xrightarrow[N\to\infty]{} 0, \\ (\frac{12N^{\beta}R(N)}{t} + 1) \exp\{-2N(\frac{t}{12R(N)N^{\beta}})^2\} \xrightarrow[N\longrightarrow\infty]{} 0, & \end{array} \tag{3}$$

then

$$P\{\omega:\, N^{\beta} \int\limits_{-\infty}^{\infty} |F(z) - F^N(z)|dz > t\} \xrightarrow[N\to\infty]{} 0. \tag{4}$$

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

Setting $R(N) = N^{\gamma}$, $\beta > 0$, $\beta + \gamma \in (0, \frac{1}{2})$, it follows (under A.2, A.3) from [6] that

$$(\frac{12N^{\beta}R(N)}{t} + 1) \exp\{-2N(\frac{t}{12R(N)N^{\beta}})^2\} \longrightarrow_{N\longrightarrow\infty} 0.$$

Consequently the validity of the relation (4) depends on the tails behaviour.

Proposition 3 ([7])**.** *Let $s = 1, t > 0, r > 0,$ the assumptions A.2, A.3 be fulfilled. Let, moreover, ξ be a random variable such that $\mathsf{E}_F|\xi|^r < \infty$. If constants $\beta, \gamma > 0$ fulfil the inequalities $0 < \beta + \gamma < 1/2, \quad \gamma > 1/r, \quad \beta + (1 - r)\gamma < 0$, then*

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

Analyzing the assertion of Proposition 3 we can obtain for $\beta := \beta(r)$ fulfilling this assertion that

$$\beta \longrightarrow_{r\longrightarrow\infty} 1/2.$$

However, when there exist only first two moments, then there does not exist β (determined by Proposition 3) for which the relation (4) is fulfilled.

Proposition 4 ([1])**.** *Let $s = 1, \quad \{\xi^i\}_{i=1}^N, N = 1, 2, \ldots$ be a sequence of independent random values corresponding to a heavy tailed distribution F with the shape parameter $\alpha \in (1, 2)$. Then the sequence*

$$\frac{N}{N^{1/\alpha}} \int\limits_{-\infty}^{\infty} |F^N(z) - F(z)|dz,\, N = 1, \ldots, \tag{5}$$

is stochastically bounded if and only if

$$\sup_{t>0} t^{\alpha}\, P\{\omega : |\xi| > t\} < \infty. \tag{6}$$

The assertion of Proposition 4 follows from Theorem 2.2 [1]. According to the definition of the stochastically bounded random sequences it holds (under the relation (6)) that

$$\lim_{M\longrightarrow\infty} \sup_{N} P\{\omega : \frac{N}{N^{1/\alpha}} \int\limits_{-\infty}^{\infty} |F(z) - F^N(z)| > M\} = 0. \tag{7}$$

4 Convergence Rate

Employing the auxiliary assertions of the former part we can obtain new results. However, first we recall some already known assertions.

Theorem 5 ([7]). *Let the assumptions A.1, A.2 and A.3 be fulfilled, $P_F \in \mathcal{M}_1(R^s)$. Then*

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

Theorem 6 ([7]). *Let the assumptions A.1, A.2 and A.3 be fulfilled, $P_F \in \mathcal{M}_1(R^s)$, $t > 0$. If*

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

then

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

The last theorem cannot be applied to the stable distributions with the shape parameter $\alpha \in (1, 2)$. Namely, then there does not exist a finite second moment. The following weaker assertion can be proven.

Theorem 7. *Let the assumptions A.1, A.2 and A.3 be fulfilled, $P_F \in \mathcal{M}_1(R^s), M > 0$. If one dimensional components ξ_i, $i = 1, \ldots, s$ of the random vector ξ have the distribution functions F_i with tails parameters $\alpha_i \in (1, 2)$ fulfilling the relations*

$$\sup_{t>0} t^{\alpha} P\{\omega : |\xi_i| > t\} < \infty, \qquad i = 1, 2, \ldots, s,$$

then

$$\lim_{M \longrightarrow \infty} \sup_{N} P\{\omega : \frac{N}{N^{1/\alpha}}|\varphi(F^N) - \varphi(F)| > M\} = 0 \quad \text{with} \quad \alpha = \min(\alpha_1, \ldots, \alpha_s). \tag{8}$$

Proof. Let $M > 0$, $\alpha \in (1, 2)$. First, it follows from Proposition 1 successively that

$$\begin{aligned} &\sup_N P\{\omega : \tfrac{N}{N^{1/\alpha}}|\varphi(F^N) - \varphi(F)| > M\} && \leq \\ &\sup_N P\{\omega : \tfrac{N}{N^{1/\alpha}} L \sum_{i=1}^{s} \int_{-\infty}^{\infty} |F^N(z) - F(z)| dz > M\} && \leq \\ &\sum_{i=1}^{s} \sup_N P\{\omega : \tfrac{N}{N^{1/\alpha_i}} \int_{-\infty}^{\infty} |F^N(z) - F(z)| dz > M/Ls\}. \end{aligned}$$

Consequently, according to Proposition 4 we can obtain

$$\begin{aligned} &\lim_{M \longrightarrow \infty} \sup_N P\{\omega : \tfrac{N}{N^{1/\alpha}}|\varphi(F^N) - \varphi(F)| > M\} \leq \\ &\sum_{i=1}^{s} \lim_{M' \longrightarrow \infty} \sup_N P\{\omega : \tfrac{N}{N^{1/\alpha_i}} \int_{-\infty}^{\infty} |F^N(z) - F(z)| dz > M'\}, \quad M' = M/Ls. \end{aligned}$$

Now already we can see that the assertion of Theorem 7 holds. □

Remark 1. Let us assume that the assumptions of Theorem 7 are fulfilled and $\bar{\beta} \in (0, 1 - 1/\alpha)$, then

$$\lim_{M \longrightarrow \infty} \sup_N P\{\omega : N^{\bar{\beta}}|\varphi(F^N) - \varphi(F)| > M\} = 0. \tag{9}$$

Setting $\bar{\beta} := \bar{\beta}(\alpha) = 1 - \frac{1}{\alpha}$, we can see that $\bar{\beta}(\alpha)$ is an increasing function of α. Moreover, it holds that

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

5 Special Cases

In this section, first, we recall the definitions of the stable and Pareto univariate distributions.

Definition 1 ([15])**.** Suppose that $\bar{\xi}^1$, $\bar{\xi}^2$ are i.i.d. non–degenerate random variables. We shall say that $\bar{\xi}^1$ has stable distribution if for any b_1, $b_2 > 0$ there exists $b > 0$ and $a \in R^1$ such that

$$b\bar{\xi}^1 + a \, =_d b_1\bar{\xi}^1 + b_2\bar{\xi}^2 \quad (=_d \text{ denotes the equality in distribution}).$$

Definition 2 ([8])**.** A random variable ξ has a Pareto distribution if

$$P\{\omega : \xi > z\} = (\tfrac{C}{z})^{\alpha}, \qquad f(z) = \alpha C^{\alpha} z^{-\alpha-1} \qquad \text{for } z \geq C,$$
$$f(z) = 0 \qquad \text{for } z < C, \quad C > 0, \, \alpha > 0 \quad \text{constants.}$$

Let us now analyze these two special cases.

1. If $s = 1$ and ξ has a stable distribution with $1 < \alpha < 2$ (shape parameter), $|\beta| \leq 1$ (skewness parameter) , $\sigma > 0$ (scale parameter) and $\mu \in R$ (location parameter), then according to the fact published e.g. in [25] (page 191) we have
$$\lim_{t \longrightarrow \infty} t^{\alpha} P\{\xi > t]\} = C_{\alpha} \frac{1+\beta}{2} \sigma^{\alpha}, \quad \lim_{t \longrightarrow \infty} t^{\alpha} P\{\xi < -t\} = C_{\alpha} \frac{1+\beta}{2} \sigma^{\alpha}, \tag{10}$$
where C_{α} fulfils for $1 < \alpha < 2$ the relation $C_{\alpha} = \frac{1-\alpha}{\Gamma(2-\alpha)\cos\frac{\pi\alpha}{2}}$. According to the relations (10) and to the fact that the corresponding distribution function belonging to ξ is continuous one we can see that the relation (6) is fulfilled.

2. If $s = 1$ and ξ has a Pareto distribution with a shape parameter α, then
$$t^{\alpha} \, P\{\omega : |\xi| > t\} = C^{\alpha}, \quad C^{\alpha} \quad \text{constant.} \tag{11}$$
Evidently, the relation (6) is fulfilled in the case of Pareto distribution with $\alpha \in (1, 2)$ too.

Consequently the results of Theorem 7 can be applied to the above mentioned.

Theorem 8. *Let the assumptions A.1, A.2 and A.3 be fulfilled, $P_F \in \mathcal{M}_1(R^s), M > 0$. If one dimensional components ξ_i, $i = 1, \ldots, s$ of the random vector ξ have either Pareto or stable distributions with tails parameters $\alpha_i \in (1, 2)$, $i = 1, \ldots, s$, then*

$$\lim_{M \longrightarrow \infty} \sup_N P\{\omega : \frac{N}{N^{1/\alpha}} |\varphi(F^N) \, - \, \varphi(F)| > M\} = 0, \quad \alpha = \min(\alpha_1, \ldots, \alpha_s). \tag{12}$$

6 Conclusion

The paper deals with the investigation of the empirical estimates of the optimal value in the case of one–stage stochastic programming problems with the "underlying" heavy tailed one dimensional marginal distributions and a shape parameters in the interval (1, 2). To this end the results on the stability corresponding to the Wasserstein metric (determined by $\mathcal{L}_1$ norm) have been employed. Evidently, while the "classical" results on the convergence rate have been achieved in the case of finite moments existence, in this case we have proven only that the sequence of difference between "theoretical" optimal value and its empirical approximation corresponds to stochastically bounded random sequence with a convergence rate parameter defined by the value of the shape parameter. However, surely, this assertion generalizes known results.

Acknowledgment This work was supported by the Grant Agency of the Czech Republic under grants P402/10/0956, P402/11/0150 and P402/10/1610.

References

[1] Barrio, E.,. Giné, E., and Matrán, C.: Central limit theorems for the Wassersten distance between the empirical and the true distributions. *Annals of Probability* **27** (1999), 2, 1009–1071.

[2] Dai, L., Chen, C., and Birge, J.H.: Convergence properties of two–stage stochastic programming. *Journal of Optimization Theory Applications* **106** (2000), 489–509.

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

[4] Dupačová, J., HurtT, J., and Štěpán, J.: *Stochastic Modelling in Economics and Finance.* Kluwer, London, 2002.

[5] Dupačová, J.: *Portfolio Optimization and Risk Management via Stochastic Programming.* Osaka University Press, Osaka, 2009.

[6] Dvoretski, A., Kiefer, J. and Wolfowitz, J.: Asymptotic minimax character of tesample distribution function and the classical multinominal estimate. *Ann. Math. Statist.* **27** (1956), 642–669.

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

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

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

[10] Kaňková, V: Optimum solution of a stochastic optimization problem with unknown parameters. In: *Trans. 7th Prague Conf. 1974.* Academia, Prague 1977, 239–244.

[11] Kaňková, V.: An approximative solution of stochastic optimization problem. In: *Trans. 8th Prague Conf. 1974.* Academia, Prague 1978, 349–353.

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

[13] Kaňková, V. and Houda M.: Empirical estimates in stochastic programming. In: *Proceedings of Prague Stochastics 2006* (M. Hušková and M. Janžura, eds.), MATFYZPRESS, Prague 2006, 426–436.

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

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

[16] Homen–de–Mello, T.: On rates of convergence for stochastic optimization problems under non–i.i.d. sampling. *SIAM J. Optimization* **19** (2008), 524–551.

[17] Kozubowski, T.J., Panorska A. K., and Rachev, S. T.: Statistical issues in modelling stable portfolios. In: *Handbook of Heavy Tailed Distributions in Finance* (S.T. Rachev, ed.), Elsevier, Amsterdam 2003, 134–168.

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

[19] Omelchenko, V.: Stable distributions and application to finnace. *Diploma Thesis* (supervisor L. Klebanov) Faculty of Mathematics and Physics, Charles University Prague, Prague 2007.

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

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

[22] Rachev, S.T., and Römisch, W.: Quantitative stability in stochastic programming: the method of probability metrics. *Math. of Oper. Res.* **27** (2002), 4, 792–818.

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

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

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

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

[27] Wets, R. J. B.: A statistical approach to the solution of stochastic programs with (convex) simple recourse. *Research Report*, University Kentucky, USA 1974.