




Estimation of the Necessary Sample Size for Approximation of Stochastic Optimization Problems with Probabilistic Criteria

Sergey V. Ivanov^(✉)  and Irina D. Zhenevskaya

Moscow Aviation Institute (National Research University),
Volokolamskoe Shosse, 4, A-80, GSP-3, 125993 Moscow, Russia
sergeyivanov89@mail.ru, genevaskaia@gmail.com

Abstract. Stochastic optimization problems with probabilistic and quantile objective functions are considered. The probability objective function is defined as the probability that the value of losses does not exceed a fixed level. The quantile function is defined as the minimal value of losses that cannot be exceeded with a fixed probability. Sample approximations of the considered problems are formulated. A method to estimate the accuracy of the approximation of the probability maximization and quantile minimization is described for the case of a finite set of feasible strategies. Based on this method, we estimate the necessary sample size to obtain (with a given probability) an epsilon-optimal strategy to the original problems by solving their approximations in the cases of finite set of feasible strategies. Also, the necessary sample size is obtained for the probability maximization in the case of a bounded infinite set of feasible strategies and a Lipschitz continuous probability function.

Keywords: Stochastic optimization · Sample approximation · Probability function · Quantile function

1 Introduction

Problems of stochastic programming with probabilistic and quantile objective functions are encountered in many applied problems, where special attention is paid to the reliability of the system. These problems and methods for solving them are well covered in [1].

In this paper, we research the sample average approximation (SAA) method for solving stochastic programming problems with probabilistic and quantile objective functions. This method is based on statistical estimation of the objective function. For the expectation objective function, the convergence of the SAA method is proved in [2]. The SAA method was applied to stochastic programming problems with probabilistic constraints in [3], where the convergence of the method was proved for a special case of the problem. In [4], the possibility of approximating a stochastic programming problem with probabilistic and

quantile objective function was investigated. The hypo-convergence of sampling probability functions is proved, which in turn guarantees the convergence of the approximation of the probability maximization and quantile minimization problems with respect to the value of the objective function and with respect to the optimization strategy.

From recent works on the SAA, [5] can be noted, where general approach to study approximations of stochastic programming problems is suggested. In [6], confidence bounds on the optimal objective value are constructed.

When the SAA method is applied, the reduced problem can be considered as a stochastic optimization problem with discrete distribution of the vector of random parameters. These problems can be reduced to mixed integer programming problems [7], which can be solved by using available software.

To apply the SAA method, it is useful to know the quality of the obtained approximate solution. In [8–10], in the case of a finite set of feasible strategies, an estimate of the required sample size was obtained to approximate an expectation minimization problem. This result was extended for the case of a Lipschitz continuous expectation function in [8]. To estimate the required number of realizations of the random vector, the exponential estimation of large deviations was used. In [11], the rate of convergence is studied for stochastic programming problems with probabilistic constraints.

This paper presents sample size estimates for problems of stochastic programming with probabilistic and quantile objective function. For the probability maximization, we consider cases of finite and bounded set of feasible strategies. For the quantile minimization, the case of finite set of feasible strategies is considered.

2 Statement

Let $(\mathcal{X}, \mathcal{F}, \mathbf{P})$ be a complete probability space. Let X be a random vector defined on this probability space. For simplicity, we assume that $\mathcal{X} \subset \mathbb{R}^m$ is a closed set.

Let us denote by $\Phi(\cdot): U \times \mathcal{X} \rightarrow (-\infty, +\infty)$ a loss function, where $U \subset \mathbb{R}^n$ is a nonempty compact set of strategies u . We assume that the function $(u, x) \mapsto \Phi(u, x)$ is lower semi-continuous in $u \in U$ and $\mathcal{B}(U) \times \mathcal{F}$ -measurable in $x \in \mathcal{X}$, where $\mathcal{B}(U)$ is the Borel σ -algebra of subsets U . These conditions guarantees that the function $(u, x) \mapsto \Phi(u, x)$ is a normal integrand [12].

Let us introduce the probability function

$$P_\varphi(u) \triangleq \mathbf{P}\{\Phi(u, X) \leq \varphi\},$$

and quantile function

$$\varphi_\alpha(u) \triangleq \min\{\varphi : P_\varphi(u) \geq \alpha\},$$

where $\alpha \in (0, P^*)$ is a given reliability level,

$$P^* = \sup_{u \in U} \mathbf{P}\{\Phi(u, X) < +\infty\}.$$

We consider the probability maximization problem

$$\alpha^* \triangleq \sup_{u \in U} P_\varphi(u) \tag{1}$$

and the quantile minimization problem

$$\varphi^* \triangleq \inf_{u \in U} \varphi_\alpha(u). \tag{2}$$

The sets of ε -optimal solutions to problems (1) and (2) are denoted by

$$U_\varphi^\varepsilon \triangleq \{u \in U: P_\varphi(u) \geq \alpha^* - \varepsilon\},$$

$$V_\alpha^\varepsilon \triangleq \{u \in U: \varphi_\alpha(u) \leq \varphi^* + \varepsilon\}$$

respectively.

3 Sample Approximation

Let X_1, \dots, X_N be a sample generated by random vector X , i.e., random vectors $X_k, k = \overline{1, N}$, are independent identically distributed with distribution function $F(x) = \mathbf{P}\{X \leq x\}$. We assume that the sample is defined on a complete probability space $(\Omega, \mathcal{F}, \mathbf{P}')$. This probability space may differ from the space $(\mathcal{X}, \mathcal{F}, \mathbf{P})$. However, below we will use the same letter \mathbf{P} for the probability \mathbf{P}' , because it is clear which probability space is considered. Then we can write the frequency of the event $\{\Phi(u, X) \leq \varphi\}$ as

$$P_\varphi^{(N)} \triangleq \frac{1}{N} \sum_{k=1}^N \chi_{(-\infty, \varphi]}(\Phi(u, X_k)). \tag{3}$$

where

$$\chi_A(x) \triangleq \begin{cases} 0, & x \in A; \\ 1, & x \notin A. \end{cases}$$

By using (3), the quantile function can be estimated by

$$\varphi_\alpha^{(N)}(u) \triangleq \min\{\varphi: P_\varphi^{(N)}(u) \geq \alpha\}. \tag{4}$$

The sample approximation of the probability maximization problem is formulated as

$$\hat{U}_\varphi^{(N)} \triangleq \text{Arg max}_{u \in U} P_\varphi^{(N)}(u), \quad N \in \mathbb{N}; \tag{5}$$

and the sample approximation of the quantile minimization problem is formulated as

$$\hat{V}_\alpha^{(N)} \triangleq \text{Arg min}_{u \in U} \varphi_\alpha^{(N)}(u), \quad N \in \mathbb{N}. \tag{6}$$

From [4] it follows that quantile function (4) coincides with the order statistics of sample values $\{\Phi_k\}_{k=1}^N$ for the random variable $\Phi \triangleq \Phi(u, X)$ that has index $[\alpha N]$ and is called the sample quantile, where $[x] \triangleq \min\{k \in \mathbb{N}: x \leq k\}$.

4 Estimation of the Necessary Sample Size for Probability Maximization

In [4] the convergence of the approximating stochastic programming problems (5) and (6). These results, however, do not show the quality of solutions for a given sample of size N . In this section, we find upper bounds on the necessary sample size to consider a solution to (5) as an approximate solution to problem (1).

4.1 Case of Finite Set of Feasible Strategies

Let us begin with the case when the set U is finite. Its cardinality is denoted by $|U|$.

We consider the event

$$\left\{ \hat{U}_\varphi^{(N)} \not\subset U_\varphi^\varepsilon \right\} = \bigcup_{u \in U \setminus U_\varphi^\varepsilon} \bigcap_{y \in U} \left\{ \hat{P}_\varphi^{(N)}(u) \geq \hat{P}_\varphi^{(N)}(y) \right\}.$$

The event $\left\{ \hat{U}_\varphi^{(N)} \not\subset U_\varphi^\varepsilon \right\}$ means that there exists an optimal solution $u_\varphi^{(N)}$ to the approximation problem (5) such that $u_\varphi^{(N)}$ is not ε -optimal solution to the true problem (1).

Then, given that the set U is finite, we can find an upper bound for the probability

$$\mathbf{P} \left\{ \hat{U}_\varphi^{(N)} \not\subset U_\varphi^\varepsilon \right\} \leq \sum_{u \in U \setminus U_\varphi^\varepsilon} \mathbf{P} \left(\bigcap_{y \in U} \left\{ \hat{P}_\varphi^{(N)}(u) \geq \hat{P}_\varphi^{(N)}(y) \right\} \right). \tag{7}$$

Let u^* be an optimal solution to the true problem (1). If there several optimal solution to problem (1), then u^* can be taken arbitrarily. It follows from (7) that

$$\begin{aligned} \mathbf{P} \left\{ \hat{U}_\varphi^{(N)} \not\subset U_\varphi^\varepsilon \right\} &\leq \sum_{u \in U \setminus U_\varphi^\varepsilon} \mathbf{P} \left\{ P_\varphi^{(N)}(u) \geq P_\varphi^{(N)}(u^*) \right\} \\ &\leq |U| \max_{u \in U \setminus U_\varphi^\varepsilon} \mathbf{P} \left\{ \hat{P}_\varphi^{(N)}(u) \geq \hat{P}_\varphi^{(N)}(u^*) \right\}. \end{aligned} \tag{8}$$

Let us introduce the random variables

$$\xi_k = \chi_{(-\infty, \varphi]}(\Phi(u, X_k)) - \chi_{(-\infty, \varphi]}(\Phi(u^*, X_k)).$$

Notice that ξ_k are independent. Then we can write

$$\mathbf{P} \left\{ \hat{P}_\varphi^{(N)}(u) \geq \hat{P}_\varphi^{(N)}(u^*) \right\} = \mathbf{P} \left\{ \sum_{k=1}^N \xi_k \geq 0 \right\} = \mathbf{P} \left\{ \exp \left(t \sum_{k=1}^N \xi_k \right) \geq 1 \right\},$$

where $t > 0$. By Chebyshev's inequality,

$$\mathbf{P} \left\{ \exp \left(t \sum_{k=1}^N \xi_k \right) \geq 1 \right\} \leq \mathbf{E} \left[\exp \left(t \sum_{k=1}^N \xi_k \right) \right] = (M(t))^N, \tag{9}$$

where

$$M(t) \triangleq \mathbf{E} [\exp (t\xi_1)] \tag{10}$$

is the moment-generating function of the random variable ξ_1 .

Lemma 1. *Let $M(t)$ be the function defined by (10). Let $0 < \varepsilon < \alpha^*$. Then*

$$\inf_{t>0} M(t) \leq \sqrt{1 - \varepsilon^2}.$$

If $\alpha^* \leq \frac{1+\varepsilon}{2}$, then

$$\inf_{t>0} M(t) \leq 2\sqrt{(\alpha^* - \varepsilon)\alpha^*} + 1 + \varepsilon - 2\alpha^* \leq \sqrt{1 - \varepsilon^2}.$$

Proof. Let us introduce the events

$$\begin{aligned} A &\triangleq \{\Phi(u, X) \leq \varphi\}, \\ B &\triangleq \{\Phi(u^*, X) \leq \varphi\}. \end{aligned}$$

Then

$$M(t) = p_+ e^t + p_- e^{-t} + 1 - p_+ - p_-,$$

where

$$p_+ = \mathbf{P}(A \cap \bar{B}), \quad p_- = \mathbf{P}(\bar{A} \cap B). \tag{11}$$

Since u^* is an optimal solution to the true problem (1) and $u \in U \setminus U_\varphi^\varepsilon$, it holds that

$$\mathbf{P}(A) \leq \alpha^* - \varepsilon, \quad \mathbf{P}(B) = \alpha^*. \tag{12}$$

From (11) and (12), it follows that

$$p_+ \in [0, \alpha^* - \varepsilon], \quad p_- \in [\varepsilon, \alpha^*]. \tag{13}$$

Therefore,

$$\varepsilon \leq \mathbf{P}(B) - \mathbf{P}(A) = \mathbf{P}(\bar{A} \cap B) + \mathbf{P}(A \cap B) - (\mathbf{P}(A \cap \bar{B}) + \mathbf{P}(A \cap B)) = p_- - p_+. \tag{14}$$

The function $t \mapsto M(t)$ is convex. From the optimality conditions, we obtain

$$\text{Arg min}_{t>0} M(t) = \left\{ \frac{1}{2} \ln \frac{p_-}{p_+} \right\}$$

if $p_+ > 0$. From (14), it follows that

$$\frac{1}{2} \ln \frac{p_-}{p_+} > 0.$$

Thus,

$$Q(p_+, p_-) = \inf_{t>0} M(t) = 2\sqrt{p_- p_+} + 1 - p_+ - p_-. \tag{15}$$

If $p_+ = 0$, then equality (15) holds too.

The function $(p_+, p_-) \mapsto Q(p_+, p_-)$ is concave (see, for example, [13, P. 74]). Let us find

$$\max_{p_+, p_-} Q(p_+, p_-)$$

subject to (13), (14), and (15). Since the unconditional maximum of $Q(p_+, p_-)$ is attained when $p_+ = p_-$, the conditional maximum is attained when the constraint (14) is active. Taking into account the constraint $p_+ + p_- \leq 1$, it is easy to see that

$$\begin{aligned} \inf_{t>0} M(t) &\leq \\ \max_{p_+, p_-} \{Q(p_+, p_-) \mid p_- - p_+ &= \varepsilon, p_+ + p_- \leq 1, p_+ \in [0, \alpha^* - \varepsilon], p_- \in [\varepsilon, \alpha^*]\} \\ &= \max_{p_+} \left\{ 2\sqrt{p_+(p_+ + \varepsilon)} + 1 - 2p_+ - \varepsilon \mid p_+ + p_+ + \varepsilon \leq 1, p_+ \in [0, \alpha^* - \varepsilon] \right\} \\ &= \begin{cases} 2\sqrt{\frac{1-\varepsilon}{2} \cdot \frac{1+\varepsilon}{2}} + 1 - (1 - \varepsilon) - \varepsilon = \sqrt{1 - \varepsilon^2} & \text{if } \alpha^* - \varepsilon > \frac{1-\varepsilon}{2}, \\ 2\sqrt{(\alpha^* - \varepsilon)\alpha^*} + 1 + \varepsilon - 2\alpha^* \leq \sqrt{1 - \varepsilon^2} & \text{if } \alpha^* - \varepsilon \leq \frac{1-\varepsilon}{2}. \end{cases} \end{aligned}$$

Thus, Lemma 1 is proved.

Let us prove a theorem on the necessary sample size to approximate the true problem (1).

Theorem 1. *Let $\beta \in (0, 1)$. If the set U is finite and*

$$N \geq 2 \frac{\ln|U| - \ln(1 - \beta)}{|\ln(1 - \varepsilon^2)|}, \tag{16}$$

then

$$\mathbf{P} \left\{ \hat{U}_\varphi^{(N)} \subset U_\varphi^\varepsilon \right\} \geq \beta. \tag{17}$$

Moreover, if it is known that $\alpha^* \leq \frac{1+\varepsilon}{2}$, then inequality (17) holds if

$$N \geq \left\lceil \frac{\ln|U| - \ln(1 - \beta)}{\ln \left(2\sqrt{(\alpha^* - \varepsilon)\alpha^*} + 1 + \varepsilon - 2\alpha^* \right)} \right\rceil. \tag{18}$$

Proof. First, let us consider the case $\alpha^* \leq \varepsilon$. Then, it is obvious that

$$\mathbf{P} \left\{ \hat{U}_\varphi^{(N)} \subset U_\varphi^\varepsilon \right\} = 1,$$

hence, the assertion of the theorem is true.

If $\alpha^* > \varepsilon$ and $\alpha^* \leq \frac{1+\varepsilon}{2}$, then, from (8), (9), and Lemma 1, it follows that

$$\mathbf{P} \left\{ \hat{U}_\varphi^{(N)} \not\subset U_\varphi^\varepsilon \right\} \leq \inf_{t>0} |U|(M(t))^N \leq \left(2\sqrt{(\alpha^* - \varepsilon)\alpha^*} + 1 + \varepsilon - 2\alpha^* \right)^N.$$

Thus, inequality (17) holds if

$$|U| \left(2\sqrt{(\alpha^* - \varepsilon)\alpha^*} + 1 + \varepsilon - 2\alpha^* \right)^N \leq 1 - \beta \Leftrightarrow$$

$$N \geq \frac{\ln(1 - \beta) - \ln |U|}{\ln \left(2\sqrt{(\alpha^* - \varepsilon)\alpha^*} + 1 + \varepsilon - 2\alpha^* \right)} = \frac{\ln |U| - \ln(1 - \beta)}{\left| \ln \left(2\sqrt{(\alpha^* - \varepsilon)\alpha^*} + 1 + \varepsilon - 2\alpha^* \right) \right|}.$$

Since

$$2\sqrt{(\alpha^* - \varepsilon)\alpha^*} + 1 + \varepsilon - 2\alpha^* \leq \sqrt{1 - \varepsilon^2},$$

we obtain that inequality (17) holds for

$$N \geq \frac{\ln |U| - \ln(1 - \beta)}{\left| \ln \sqrt{1 - \varepsilon^2} \right|} = 2 \frac{\ln |U| - \ln(1 - \beta)}{\left| \ln(1 - \varepsilon^2) \right|}.$$

In the case when $\alpha^* > \varepsilon$ and $\alpha^* > \frac{1+\varepsilon}{2}$, the theorem is proved in the same manner. Theorem 1 is proved.

Remark 1. In [9, Theorem 5.17], a result similar to Theorem 1 is proved for the maximization of the expectation function. By applying this result to problem (1), it can be obtained that inequality (17) holds for

$$N \geq 2 \frac{\ln |U| - \ln(1 - \beta)}{\varepsilon^2}.$$

It is easy to check that

$$\varepsilon^2 < \left| \ln(1 - \varepsilon^2) \right|$$

for $\varepsilon \in (0, 1)$. Thus, the sample estimate (16) improves the result in [8] for maximization of the probability function.

Remark 2. To apply the sample estimate (18), we need to know exact solution to problem (1). However, the sample approximation is construct to estimate α^* . Sometimes, it possible to find an upper bound $\bar{\alpha} \geq \alpha^*$. If $\bar{\alpha} \leq \frac{1+\varepsilon}{2}$, then we can improve the sample estimate (16). It is guaranteed that inequality (17) holds if

$$N \geq \frac{\ln |U| - \ln(1 - \beta)}{\left| \ln \left(2\sqrt{(\bar{\alpha} - \varepsilon)\bar{\alpha}} + 1 + \varepsilon - 2\bar{\alpha} \right) \right|}.$$

4.2 Case of Bounded Set of Feasible Strategies

Let us consider the case when U is a bounded, not necessarily finite, subset of \mathbb{R}^n . The diameter of U is denoted by

$$D \triangleq \sup_{u, v \in U} \|u - v\|,$$

where the norm $\|u\| = \max\{|u_1|, \dots, |u_n|\}$ is used.

We suppose that the probability function $u \mapsto P_\varphi(u)$ is Lipschitz continuous on U with a Lipschitz constant L , i.e.

$$|P_\varphi(u) - P_\varphi(v)| \leq L\|u - v\| \tag{19}$$

for all $u, v \in U$.

Theorem 2. *Let $\beta \in (0, 1)$. If the assumption (19) is satisfied and*

$$N \geq 2 \inf_{\gamma \in (0,1)} \frac{n \ln \left\lceil \frac{DL}{(1-\gamma)\varepsilon} \right\rceil - \ln(1 - \beta)}{|\ln(1 - \gamma^2\varepsilon^2)|}, \tag{20}$$

then

$$\mathbf{P} \left\{ \hat{U}_\varphi^{(N)} \subset U_\varphi^\varepsilon \right\} \geq \beta. \tag{21}$$

Proof. First, let us check that the event $\left\{ \hat{U}_\varphi^{(N)} \subset U_\varphi^\varepsilon \right\} \in \mathcal{F}'$. Since the function $(u, x) \mapsto \Phi(u, x)$ is a normal integrand, the function $u \mapsto P_\varphi(u)$ is upper semi-continuous and the set U_φ^ε is compact [4]. Also, the $u \mapsto \hat{P}_\varphi^{(N)}(u)$ is upper semi-continuous for all fixed realizations of the sample. The compactness of U and semi-continuity of these function imply that the supremum

$$\sup_{u \in U} \hat{P}_\varphi^{(N)}(u)$$

is attained and is a measurable function of the sample. Thus, the considered event can be represented as

$$\begin{aligned} \left\{ \hat{U}_\varphi^{(N)} \subset U_\varphi^\varepsilon \right\} &= \bigcap_{u \in U \setminus U_\varphi^\varepsilon} \left\{ \hat{P}_\varphi^{(N)}(u) < \sup_{v \in U} \hat{P}_\varphi^{(N)}(v) \right\} \\ &= \bigcap_{k \in \mathbb{N}} \left\{ \sup_{u \in U_k} \hat{P}_\varphi^{(N)}(u) < \sup_{u \in U} \hat{P}_\varphi^{(N)}(u) \right\}, \end{aligned}$$

where

$$U_k = \left\{ u \in U : \inf_{v \in U_\varphi^\varepsilon} \|u - v\| \geq \frac{1}{k} \right\}.$$

The set U_k is compact, so the function

$$(x_1, \dots, x_N) \mapsto \sup_{u \in U_k} \hat{P}_\varphi^{(N)}(u)$$

is measurable and, hence,

$$\left\{ \hat{U}_\varphi^{(N)} \subset U_\varphi^\varepsilon \right\} \in \mathcal{F}'.$$

Let \tilde{U} be a finite subset of U . Let

$$\nu \triangleq \sup_{u \in U} \inf_{v \in \tilde{U}} \|u - v\|.$$

The value of ν shows the maximal distance between an arbitrary point of U and the nearest point of \tilde{U} . The set \tilde{U} can be selected in such a way that

$$|\tilde{U}| \leq \left\lceil \frac{D}{\nu} \right\rceil^n. \tag{22}$$

It will be assumed that condition (22) is satisfied.

Let

$$\tilde{U}_\varphi^\varepsilon \triangleq \{u \in \tilde{U} : P_\varphi(u) \geq \alpha_{\tilde{U}}^* - \varepsilon\},$$

where

$$\alpha_{\tilde{U}}^* \triangleq \sup_{u \in \tilde{U}} P_\varphi(u).$$

Since the function $u \mapsto P_\varphi(u)$ is Lipschitz continuous, the condition $u \in \tilde{U}_\varphi^\varepsilon$ implies $u \in U_\varphi^{\varepsilon+L\nu}$. If $\gamma \in (0, 1)$ is a fixed number and $L\nu = (1 - \gamma)\varepsilon$, then

$$\left\{ \hat{U}_\varphi^{(N)} \subset \tilde{U}_\varphi^{\gamma\varepsilon} \right\} \subset \left\{ \hat{U}_\varphi^{(N)} \subset U_\varphi^\varepsilon \right\}. \tag{23}$$

From (23) and Theorem 1, it follows that

$$\mathbf{P} \left\{ \hat{U}_\varphi^{(N)} \subset U_\varphi^\varepsilon \right\} \geq \mathbf{P} \left\{ \hat{U}_\varphi^{(N)} \subset \tilde{U}_\varphi^{\gamma\varepsilon} \right\} \geq \beta$$

if

$$N \geq 2 \frac{\ln |\tilde{U}| - \ln(1 - \beta)}{|\ln(1 - \gamma^2\varepsilon^2)|} \leq 2 \frac{\ln \left\lceil \frac{D}{\nu} \right\rceil^n - \ln(1 - \beta)}{|\ln(1 - \gamma^2\varepsilon^2)|} = 2 \frac{n \ln \left\lceil \frac{DL}{(1-\gamma)\varepsilon} \right\rceil - \ln(1 - \beta)}{|\ln(1 - \gamma^2\varepsilon^2)|}.$$

Since γ is selected arbitrarily, the theorem is proved.

Remark 3. A similar result for minimization of the expectation function is proved in [8]. This result can be obtained from Theorem 2 if $t = 1/2$. Additional optimization in $t \in (0, 1)$ can improve the result [8] for the special case of probability maximization. If the exact value of the infimum is difficult to find, then the sample estimate N can be found by substituting several values of $t \in (0, 1)$ into (20). The minimal value of the obtained numbers can be set as the sample estimate N .

Remark 4. To apply the sample estimate (20), the Lipschitz constant is need to know. If the function $u \mapsto P_\varphi(u)$ is continuously differentiable on the set U , then, from the mean value theorem, it follows that

$$|P_\varphi(u) - P_\varphi(v)| \leq \sup_{w \in U} \|\nabla P_\varphi(w)\| \|u - v\|.$$

Therefore, we can take $L = \sup_{w \in U} \|\nabla P_\varphi(w)\|$. Methods to find the gradient of the probability function are described in [14, 15].

5 Estimation of the Necessary Sample Size for Quantile Minimization

In this section, we find upper bounds on the necessary sample size to consider a solution to (6) as an approximate solution to problem (2). We assume that the set U is finite.

We suppose that the following assumption is satisfied.

Assumption 1. *The random variable $\Phi(u, X)$ is absolutely continuous for all $u \in U$ with probability density function $p_u(\cdot)$ continuous at the point $\varphi_\alpha(u)$ with its neighborhood. Also, there exists a number $C > 0$ such that*

$$\min_{u \in U} p_u(\varphi_\alpha(u)) > C.$$

As for the probability function, consider the event

$$\left\{ \hat{V}_\alpha^{(N)} \notin V_\alpha^\varepsilon \right\} = \bigcup_{u \in U \setminus V_\alpha^\varepsilon} \bigcap_{y \in U} \left\{ \hat{\varphi}_\alpha^{(N)}(u) \leq \hat{\varphi}_\alpha^{(N)}(y) \right\}.$$

Then we can find an upper bound for the probability

$$\mathbf{P} \left\{ \hat{V}_\alpha^{(N)} \notin V_\alpha^\varepsilon \right\} \leq \sum_{u \in U \setminus V_\alpha^\varepsilon} \mathbf{P} \left(\bigcap_{y \in U} \left\{ \hat{\varphi}_\alpha^{(N)}(u) \leq \hat{\varphi}_\alpha^{(N)}(y) \right\} \right). \tag{24}$$

Let us fix an optimal solution to the true problem (2) u^* . From (24), we obtain

$$\begin{aligned} \mathbf{P} \left\{ \hat{V}_\alpha^{(N)} \notin V_\alpha^\varepsilon \right\} &\leq \sum_{u \in U \setminus V_\alpha^\varepsilon} \mathbf{P} \left\{ \hat{\varphi}_\alpha^{(N)}(u) \leq \hat{\varphi}_\alpha^{(N)}(u^*) \right\} \\ &\leq |U| \max_{u \in U \setminus V_\alpha^\varepsilon} \mathbf{P} \left\{ \hat{\varphi}_\alpha^{(N)}(u) \leq \hat{\varphi}_\alpha^{(N)}(u^*) \right\}. \end{aligned} \tag{25}$$

Let us define the random variables

$$\eta_N = \hat{\varphi}_\alpha^{(N)}(u^*) - \hat{\varphi}_\alpha^{(N)}(u), \quad N \in \mathbb{N}.$$

Notice that the random variables $\varphi_\alpha^{(N)}(u^*)$ and $\varphi_\alpha^{(N)}(u)$ can be dependent. We need to find an upper bound on the probability

$$\mathbf{P}\{\eta_N \geq 0\}.$$

By the Mosteller theorem [16], the distribution of the order statistics $\varphi_\alpha^{(N)}(u^*)$ and $\varphi_\alpha^{(N)}(u)$ converges to a normal distribution:

$$\begin{aligned} \sqrt{N} \left(\varphi_\alpha^{(N)}(u^*) - \varphi^* \right) &\xrightarrow{d} Z_{u^*} \sim \mathcal{N} \left(0, \frac{\alpha(1-\alpha)}{p_{u^*}^2(\varphi^*)} \right), \\ \sqrt{N} \left(\varphi_\alpha^{(N)}(u) - \varphi_\alpha(u) \right) &\xrightarrow{d} Z_u \sim \mathcal{N} \left(0, \frac{\alpha(1-\alpha)}{p_u^2(\varphi_\alpha(u))} \right). \end{aligned}$$

Therefore, we can write

$$\begin{aligned} \lim_{N \rightarrow \infty} \mathbf{E}\eta_N &\leq -\varepsilon, \\ \limsup_{N \rightarrow \infty} \mathbf{var} [\eta_N \sqrt{N}] &< \frac{4\alpha(1-\alpha)}{C}. \end{aligned}$$

Thus, for any $0 < \varepsilon' < \varepsilon$, there exists $\tilde{N}(\varepsilon') \in \mathbb{N}$ such that

$$\mathbf{E}\eta_N \leq -\varepsilon + \varepsilon', \tag{26}$$

$$\mathbf{var} [\eta_N] < \frac{4\alpha(1-\alpha)}{CN} \tag{27}$$

for all $N > \tilde{N}(\varepsilon')$. We would like to notice that $\tilde{N}(\varepsilon')$ can depend on u . So, the maximal value of $\tilde{N}(\varepsilon')$ should be taken.

By Cantelli's inequality, for $N > \tilde{N}(\varepsilon')$,

$$\begin{aligned} \mathbf{P}\{\eta_N \geq 0\} &\leq \mathbf{P}\{\eta_N - \mathbf{E}\eta_N \geq \varepsilon - \varepsilon'\} \leq \frac{\mathbf{var} [\eta_N]}{\mathbf{var} [\eta_N] + (\varepsilon - \varepsilon')^2} \\ &< \frac{4\alpha(1-\alpha)}{4\alpha(1-\alpha) + (\varepsilon - \varepsilon')^2 CN}. \end{aligned}$$

Thus, from (25), the theorem follows.

Theorem 3. *Let U be a finite set, $\beta \in (0, 1)$. Assumption 1 is supposed to be satisfied. Then*

$$\mathbf{P} \left\{ \hat{V}_\alpha^{(N)} \notin V_\alpha^\varepsilon \right\} \leq |U| \frac{4\alpha(1-\alpha)}{4\alpha(1-\alpha) + (\varepsilon - \varepsilon')^2 CN}. \tag{28}$$

for sufficiently large N .

Now, we can obtain the corollary from Theorem 3.

Corollary 1. *Let assumptions of Theorem 3 be satisfied. Then*

$$\mathbf{P} \left\{ \hat{V}_\alpha^{(N)} \subset V_\alpha^\varepsilon \right\} \geq \beta$$

if

$$N \geq \frac{4\alpha(1-\alpha)(|U| + \beta - 1)}{(1-\beta)(\varepsilon - \varepsilon')^2 C} \tag{29}$$

and $N > \tilde{N}(\varepsilon')$.

Proof. From (28), it follows that the assertion of the theorem is true if

$$|U| \frac{4\alpha(1-\alpha)}{4\alpha(1-\alpha) + (\varepsilon - \varepsilon')^2 CN} \leq 1 - \beta.$$

By solving this inequality, we obtain (29). The corollary is proved.

Remark 5. Unfortunately, it is difficult find bounds on the value $\tilde{N}(\varepsilon')$. To use the estimate (29), inequalities (26) and (27) should be checked. It can be made by statistical methods.

6 Conclusion

In the paper, sample size estimates for approximation of stochastic optimization problems with probabilistic and quantile objective functions are obtained. These estimates are quite rough for practical use, but they allow us to judge the complexity of the solution to the original problem. In future research, it is planned to improve this result for special cases of stochastic optimization problems. We hope that it is possible to describe a class of problems for which exponential bounds can be obtained instead of (28). Also, a more general case of quantile minimization problem on a bounded set should be studied.

Acknowledgements. The work is supported by the Russian Foundation for Basic Research (project 19-07-00436).

References

1. Kibzun, A.I., Kan, Y.S.: Stochastic Programming Problems with Probability and Quantile Functions. Wiley, Chichester, New York, Brisbane, Toronto, Singapore (1996)
2. Artstein, Z., Wets, R.J.-B.: Consistency of minimizers and the SLLN for stochastic programs. *J. Convex Anal.* **2**, 1–17 (1996)
3. Pagnoncelli, B.K., Ahmed, S., Shapiro, A.: Sample average approximation method for chance constrained programming: theory and applications. *J. Optim. Theory Appl.* **142**, 399–416 (2009). <https://doi.org/10.1007/s10957-009-9523-6>
4. Kibzun, A.I., Ivanov, S.V.: On the convergence of sample approximations for stochastic programming problems with probabilistic criteria. *Automat. Remote Control.* **79**(2), 216–228 (2018). <https://doi.org/10.1134/S0005117918020029>
5. Hess, C., Seri, R.: Generic consistency for approximate stochastic programming and statistical problems. *SIAM J. Optim.* **29**(1), 290–317 (2019). <https://doi.org/10.1137/17M1156769>
6. Guigues, V., Juditsky, A., Nemirovski, A.: Non-asymptotic confidence bounds for the optimal value of a stochastic program. *Optim. Method. Softw.* **32**(5), 1033–1058 (2017). <https://doi.org/10.1080/10556788.2017.1350177>
7. Kibzun, A.I., Naumov, A.V., Norkin, V.I.: On reducing a quantile optimization problem with discrete distribution to a mixed integer programming problem. *Automat. Remote Control.* **74**(6), 951–967 (2013). <https://doi.org/10.1134/S0005117913060064>
8. Shapiro, A.: Monte Carlo sampling methods. In: Ruszczyński, A., Shapiro, A. (eds.) *Handbooks in OR Handbooks in Operations Research and Management Science & MS*, North-Holland, Dordrecht, The Netherlands, vol. 10, pp. 353–425 (2003). [https://doi.org/10.1016/S0927-0507\(03\)10006-0](https://doi.org/10.1016/S0927-0507(03)10006-0)
9. Shapiro, A., Dentcheva, D., Ruszczyński, A.: *Lectures on Stochastic Programming*. Society for Industrial and Applied Mathematics, Philadelphia (2009)
10. Kleywegt, A.J., Shapiro, A., Homem-De-Mello, T.: The sample average approximation method for stochastic discrete optimization. *SIAM J. Optim.* **12**(2), 479–502 (2001). <https://doi.org/10.1137/S1052623499363220>
11. Luedtke, J., Ahmed, S.: A sample approximation approach for optimization with probabilistic constraints. *SIAM J. Optim.* **19**(2), 674–699 (2008). <https://doi.org/10.1137/070702928>

12. Rockafellar, R.T., Wets, R.J.-B.: Variational Analysis. Springer, Heidelberg (2009). <https://doi.org/10.1007/978-3-642-02431-3>
13. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press, New York (2004)
14. Kibzun, A., Uryasev, S.: Differentiability of probability function. Stochast. Anal. Appl. **16**(6), 1101–1128 (1998). <https://doi.org/10.1080/07362999808809581>
15. Van Ackooij, W., Henrion, R.: Gradient formulae for nonlinear probabilistic constraints with Gaussian and Gaussian-like distributions. SIAM J. Optim. **24**(4), 1864–1889 (2014). <https://doi.org/10.1137/130922689>
16. Mosteller, F.: One some useful inefficient statistics. Ann. Math. Stat. **17**, 317–408 (1946)