

# Sample Average Approximation of Expected Value Constrained Stochastic Programs

Wei Wang and Shabbir Ahmed<sup>1</sup>

*H. Milton Stewart School of Industrial and Systems Engineering,  
Georgia Institute of Technology, Atlanta, GA 30332, USA*

---

## Abstract

We propose a sample average approximation (SAA) method for stochastic programming problems involving an expected value constraint. Such problems arise, for example, in portfolio selection with constraints on conditional value-at-risk (CVaR). Our contributions include an analysis of the convergence rate and a statistical validation scheme for the proposed SAA method. Computational results using a portfolio selection problem with a CVaR constraint are presented.

*Key words:* Sample average approximation; Expected value constrained stochastic program; Conditional value-at-risk; Convergence rate; Validation scheme; Portfolio optimization

---

## 1 Introduction

We consider expected value constrained stochastic programs of the form

$$\min_{x \in X} \{f(x) : \mathbb{E}[G(x, \omega)] \leq q\} \quad (1)$$

where  $X \subseteq \mathbb{R}^k$  is a nonempty set of feasible decisions,  $\omega$  is a random vector having probability distribution  $P$  and support  $\Omega$ ,  $f : X \rightarrow \mathbb{R}$ , and  $G : X \times \Omega \mapsto \mathbb{R}$  is a function such that  $\mathbb{E}[G(x, \omega)]$  is well-defined at all  $x \in X$ .

Problem (1) can arise in various situations. For example, it can be a two-stage stochastic program which aims to minimize the first-stage cost  $f(\cdot)$  while

---

<sup>1</sup> Corresponding author. Email: [sahmed@isye.gatech.edu](mailto:sahmed@isye.gatech.edu).

controlling the expected second-stage cost  $\mathbb{E}[G(\cdot, \omega)]$ . Another example is a conditional value-at-risk (CVaR) constrained problem [6],

$$\min_{x \in X} \{f(x) : \text{CVaR}_\alpha[G(x, \omega)] \leq q\} \quad (2)$$

where  $\alpha \in (0, 1)$ . It is well known [10] that for any random variable  $Z$  representing loss

$$\text{CVaR}_\alpha[Z] := \inf_{t \in \mathbb{R}} \left\{ t + (1 - \alpha)^{-1} \mathbb{E}[Z - t]_+ \right\}, \quad (3)$$

where  $[a]_+ = \max\{0, a\}$ . Problem (2) is then equivalent to

$$\min_{x \in X, t \in \mathbb{R}} \left\{ f(x) : \mathbb{E}[\tilde{G}_\alpha(x, t, \omega)] \leq q \right\},$$

where  $\tilde{G}_\alpha(x, t, \omega) := t + (1 - \alpha)^{-1}[G(x, \omega) - t]_+$ . Note that the minimizer in (3) is the  $\alpha$ -value-at-risk (VaR) defined as

$$\text{VaR}_\alpha[Z] := \min_{t \in \mathbb{R}} \{t : \Pr\{Z \leq t\} \geq \alpha\}.$$

Thus (2) is a conservative approximation of the much more difficult chance constrained problem

$$\min_{x \in X} \{f(x) : \Pr\{G(x, \omega) \leq q\} \geq \alpha\}, \quad (4)$$

in the sense that any solution feasible to (2) must be feasible to (4) [8,10].

In many cases, exact evaluation of the expected value  $\mathbb{E}[G(x, \omega)]$  in (1) for a given decision  $x$  is either impossible or prohibitively expensive. This difficulty is also inherent in traditional stochastic programming problems involving only expected value objectives

$$\min_{x \in X} \mathbb{E}[G(x, \omega)]. \quad (5)$$

The Sample Average Approximation (SAA) method [5] is a well-known approach for by-passing this difficulty. The idea is to generate a random sample  $\{\omega_1, \dots, \omega_N\}$  of  $\omega$  and solve a deterministic sample average approximate problem

$$\min_{x \in X} N^{-1} \sum_{n=1}^N G(x, \omega_n). \quad (6)$$

A solution to (6) then serves as an approximate solution of (5). Under mild conditions it has been shown that an optimal solution and the optimal value of (6) converge exponentially fast to their true counterparts in (5) as the sample size  $N$  increases. Using the convergence rate we can compute the sample size required in the approximating problem (6) to obtain solutions of the true problem (5) of desired quality with high confidence. Moreover the sampled approximation problem (6) also provides a simple way of statistically estimating lower bounds on the true optimal objective value, and hence it is possible to provide solution quality guarantees. A number of computational studies have

also verified the effectiveness of the SAA approach for stochastic programs of the form (5). See [11] and references therein for further details.

In this paper we investigate an SAA method for expected value constrained problems (1). We require the expected value constraint in (1) to be soft, i.e., after a slight adjustment to the right-hand-side  $q$ , the problem remains mathematically feasible and meaningful. This is a necessity of the nature of approximation. In cases where the approximation is “restrictive” (the region is smaller than that of the original problem), we assume the problem is still feasible. In cases where the approximation is “relaxed” (the region is larger), we hope the solution obtained is still meaningful in the sense that it can provide some useful information, e.g. optimality bounds. We analyze convergence rates of solutions of the approximating problem to the true counterpart, and also suggest a simple statistical validation scheme.

A number of authors have considered expected value constrained stochastic programs (1). O’Brien [9] studied (1) in the case where the random vector has a finite support. He proposed solution techniques including reformulating the problem as one with dual angular structure and using Benders decomposition. Kuhn [7] and Atlason, Epelman and Henderson [2] considered the case where the support of  $\omega$  is infinite. In [7], the author proposed bounding approximation schemes for multi-stage stochastic programs with expected value constraints. These schemes require that the function  $G(x, \omega)$  is convex in both  $x$  and  $\omega$ . In [2], the authors formulated a call center staffing problem as an integer expected value constrained problem and used sample average approximation together with a cutting plane method to solve it. They also analyzed the convergence rate with respect to the sample size of the simulation. Their results however require that  $x$  be discrete. By contrast our SAA analysis is for general expected value constrained stochastic programs without convexity or discreteness assumptions.

The remainder of this paper is organized as follows. Section 2 establishes the convergence results for general expected value constrained problems and as a special case, those for CVaR constrained problems. Section 3 presents an SAA scheme, where we discuss bounding techniques. Section 4 reports numerical results from using the proposed SAA scheme to solve a portfolio selection problem. Finally, conclusions are provided in Section 5.

## 2 Convergence Results

The main technique to derive the convergence rates is Large Deviations (LD) theory (cf. [13]), analogous to existing SAA analysis for traditional stochastic programs [5]. Consider an *iid* sequence  $Z_1, \dots, Z_N$  of replications of a

real-valued random variable  $Z$ . Let  $\mu := \mathbb{E}[Z]$ , which is finite, and  $\hat{Z}_N = N^{-1} \sum_{n=1}^N Z_n$  be the corresponding sample average. Then for any real number  $a > \mu$  we have the LD inequality

$$\Pr\{\hat{Z}_N \geq a\} \leq e^{-NI(a)}, \quad (7)$$

and similarly for  $a < \mu$ ,

$$\Pr\{\hat{Z}_N \leq a\} \leq e^{-NI(a)}, \quad (8)$$

where  $I(u) := \sup_{s \in \mathbb{R}} \{su - \log M(s)\}$  for  $u \in \mathbb{R}$ , is the LD rate function and  $M(s) := \mathbb{E}[e^{sZ}]$  is the moment generating function (MGF) of  $Z$ . Furthermore, suppose that the moment generating function  $M(s)$  is finite valued in a neighborhood of  $s = 0$ . Then by Taylor's expansion,

$$I(a) = \frac{(a - \mu)^2}{2\sigma^2} + o(|a - \mu|^2),$$

where  $\sigma^2 = \text{Var}[Z]$ , the variance of  $Z$ . This implies  $I(a) > 0$ .

Next we analyze convergence rates of the SAA approach for general expected value constrained programs (1) and then apply it to CVaR constrained problems (2).

### 2.1 The General Case

Let us introduce some notation and assumptions. Given  $\epsilon > 0$  define

$$X^\epsilon := \{x \in X : g(x) := \mathbb{E}[G(x, \omega)] \leq q + \epsilon\}.$$

Then  $X^0$  represents the feasible region of Problem (1). Let  $\{\omega_1, \dots, \omega_N\}$  be a sample of size  $N$  of  $\omega$ . Correspondingly we define

$$X_N^\epsilon := \{x \in X : g_N(x) := N^{-1} \sum_{n=1}^N G(x, \omega_n) \leq q + \epsilon\}.$$

Our goal is to estimate

$$\Pr\{X^{-\epsilon} \subseteq X_N^0 \subseteq X^\epsilon\}.$$

That is, we want to claim a feasible solution of the sample average approximation of (1) is  $\epsilon$ -feasible to the true problem, and at the same time the approximation is not too conservative, i.e., it is a relaxation of (1) with  $q$  replaced by  $q - \epsilon$ . The following assumptions will be required.

(C1)  $X \subset \mathbb{R}^k$  is a nonempty compact set.

- (C2) The expected value function  $g(x)$  is well-defined, i.e., for every  $x \in X$ , the function  $G(x, \cdot)$  is measurable and  $\mathbb{E}|G(x, \omega)| < +\infty$ .
- (C3) For any  $x \in X$ , the moment generating function  $M_x(\cdot)$  of  $G(x, \omega) - g(x)$  is finite in a neighborhood of zero.
- (C4) For any  $\omega \in \Omega$  there exists an integrable function  $\phi : \Omega \rightarrow \mathbb{R}_+$  such that

$$|G(x_1, \omega) - G(x_2, \omega)| \leq \phi(\omega) \|x_1 - x_2\|, \quad \forall x_1, x_2 \in X.$$

Denote  $\Phi := \mathbb{E}[\phi(\omega)]$ .

- (C5) The MGF  $M_\phi(\cdot)$  of  $\phi(\omega)$  is finite in a neighborhood of zero.

We begin with convergence analysis for expected value constrained problems where the cardinality  $|X|$  of the set  $X$  is finite. Then we extend it to the case where  $|X|$  is infinite.

**Proposition 1** *Suppose (C1)-(C3) hold and  $|X|$  is finite. Define*

$$\sigma^2 := \max_{x \in X} \text{Var}[G(x, \omega) - g(x)].$$

*Given  $\epsilon > 0$ , the following is true:*

(i) *convergence rate:*

$$\Pr \{X^{-\epsilon} \subseteq X_N^0 \subseteq X^\epsilon\} \geq 1 - 2|X|e^{-\frac{N\epsilon^2}{2\sigma^2}};$$

(ii) *estimate of the sample size for  $\Pr \{X^{-\epsilon} \subseteq X_N^0 \subseteq X^\epsilon\} \geq 1 - \beta$  to hold:*

$$N \geq \frac{2\sigma^2}{\epsilon^2} \log \left( \frac{2|X|}{\beta} \right).$$

**Proof.** We derive (i) as follows.

$$\begin{aligned} & \Pr \{X^{-\epsilon} \subseteq X_N^0 \subseteq X^\epsilon\} \\ &= 1 - \Pr \{ \exists x \in X \text{ s.t. } g(x) \leq q - \epsilon \text{ and } g_N(x) > q, \\ & \quad \text{or } \exists x \in X \text{ s.t. } g_N(x) \leq q \text{ and } g(x) > q + \epsilon \} \\ &\geq 1 - \Pr \{ \exists x \in X \text{ s.t. } g_N(x) - g(x) > \epsilon \} \\ & \quad - \Pr \{ \exists x \in X \text{ s.t. } g_N(x) - g(x) < -\epsilon \} \\ &\geq 1 - \sum_{x \in X} [\Pr \{g_N(x) - g(x) > \epsilon\} + \Pr \{g_N(x) - g(x) < -\epsilon\}] \\ &\geq 1 - \sum_{x \in X} [e^{-NI_x(\epsilon)} + e^{-NI_x(-\epsilon)}] \\ &\geq 1 - 2|X|e^{-Na(\epsilon)}, \end{aligned}$$

where  $a(\epsilon) := \min_{x \in X} \{I_x(\epsilon), I_x(-\epsilon)\}$ . Note that the second-last inequality follows from LD inequalities (7) and (8). Assumption (C3) implies that  $I_x(\epsilon)$

and  $I_x(-\epsilon)$  are no less than  $\epsilon^2(2\text{Var}[G(x, \omega) - g(x)])^{-1}$  at any  $x \in X$ , and as a result,  $a(\epsilon) \geq \epsilon^2(2\sigma^2)^{-1}$ .

To get  $\Pr\{X^{-\epsilon} \subseteq X_N^0 \subseteq X^\epsilon\} \geq 1 - \beta$ , it is sufficient to set  $2|X|e^{-\frac{N\epsilon^2}{2\sigma^2}} \leq \beta$ , which gives (ii).  $\square$

Now we consider the case where  $|X|$  is infinite. To deal with the infiniteness and obtain similar result as in Proposition 1, we use the idea of discretization as well as the assumption of Lipschitz continuity, as in existing SAA methods [1,12]. Given  $\nu > 0$ , build a finite subset  $X_\nu$  of  $X$  such that for any  $x \in X$  there exists  $x' \in X_\nu$  satisfying  $\|x - x'\| \leq \nu$ . Denoting by  $D$  the diameter of the set  $X$ , i.e.,  $D = \max_{x_1, x_2 \in X} \|x_1 - x_2\|$ , then such set  $X_\nu$  can be constructed with  $|X_\nu| \leq (D/\nu)^k$ . Now fix  $x \in X$  and  $x' \in X_\nu$  satisfying  $\|x - x'\| \leq \nu$ . Suppose Assumption (C4) holds. Then

$$|G(x, \omega) - G(x', \omega)| \leq \phi(\omega)\nu.$$

Consequently

$$|g(x) - g(x')| \leq \Phi\nu$$

and

$$|g_N(x) - g_N(x')| \leq \Phi_N\nu,$$

where  $\Phi_N = N^{-1} \sum_{n=1}^N \phi(\omega_n)$ .

**Proposition 2** *Suppose (C1)-(C5) hold and  $|X|$  is infinite. Define*

$$\nu := (4\Phi/\epsilon + 1)^{-1}$$

and

$$\sigma^2 := \max_{x \in X} \{\text{Var}[\phi(\omega)], \text{Var}[G(x, \omega) - g(x)]\}.$$

*Given  $\epsilon > 0$ , then the following holds:*

(1) *convergence rates:*

$$\Pr\{X^{-\epsilon} \subseteq X_N^0 \subseteq X^\epsilon\} \geq 1 - 2 \left(1 + \frac{D^k}{\nu^k}\right) e^{-\frac{N\epsilon^2}{8\sigma^2}};$$

(2) *estimate of the sample size for  $\Pr\{X^{-\epsilon} \subseteq X_N^0 \subseteq X^\epsilon\} \geq 1 - \beta$  to hold:*

$$N \geq \frac{8\sigma^2}{\epsilon^2} \log \left[ \frac{2}{\beta} \left(1 + \frac{D^k}{\nu^k}\right) \right].$$

**Proof.**

$$\begin{aligned}
& \Pr\{X^{-\epsilon} \subseteq X_N^0 \subseteq X^\epsilon\} \\
& \geq 1 - \Pr\{\exists x \in X \text{ s.t. } g_N(x) - g(x) > \epsilon\} \\
& \quad - \Pr\{\exists x \in X \text{ s.t. } g_N(x) - g(x) < -\epsilon\} \\
& \geq 1 - \Pr\{\exists x \in X_\nu \text{ s.t. } g_N(x) - g(x) > \epsilon - (\Phi + \Phi_N)\nu\} \\
& \quad - \Pr\{\exists x \in X_\nu \text{ s.t. } g(x) - g_N(x) > \epsilon - (\Phi + \Phi_N)\nu\} \\
& \geq 1 - 2\Pr\{\Phi_N > \Phi + \epsilon/2\} - \Pr\{\exists x \in X_\nu \text{ s.t. } g_N(x) - \\
& \quad g(x) > \epsilon/2\} - \Pr\{\exists x \in X_\nu \text{ s.t. } g(x) - g_N(x) > \epsilon/2\} \\
& \geq 1 - 2e^{-NI_\phi(\Phi + \epsilon/2)} - \sum_{x \in X_\nu} \left[ e^{-NI_x(\epsilon/2)} + e^{-NI_x(-\epsilon/2)} \right] \\
& \geq 1 - 2(1 + |X_\nu|)e^{-Nb(\epsilon)} \\
& \geq 1 - 2 \left[ 1 + \left( \frac{D}{\nu} \right)^k \right] e^{-Nb(\epsilon)}
\end{aligned}$$

where  $b(\epsilon) := \min_{x \in X_\nu} \{I_\phi(\Phi + \epsilon/2), I_x(\epsilon/2), I_x(-\epsilon/2)\}$ . By Assumptions (C3) and (C5),

$$\begin{aligned}
b(\epsilon) & \geq \min_{x \in X_\nu} \left\{ \frac{\epsilon^2}{8\text{Var}[\phi(\omega)]}, \frac{\epsilon^2}{8\text{Var}[G(x, \omega) - g(x)]} \right\} \\
& = \frac{\epsilon^2}{8 \max_{x \in X_\nu} \{\text{Var}[\phi(\omega)], \text{Var}[G(x, \omega) - g(x)]\}} \\
& \geq \frac{\epsilon^2}{8\sigma^2}.
\end{aligned}$$

The statements (1) and (2) now follow.  $\square$

## 2.2 CVaR Constrained Problems

Recall a CVaR constrained problem (2) can be reformulated as

$$\min_{x \in X, t \in \mathbb{R}} \{f(x) : \mathbb{E}[\tilde{G}_\alpha(x, t, \omega)] \leq q\}, \quad (9)$$

where

$$\tilde{G}_\alpha(x, t, \omega) = t + (1 - \alpha)^{-1}[G(x, \omega) - t]_+.$$

To apply the convergence results from the general case in Section 2.1, we need to guarantee that the conditions (C1)-(C5) hold for (9). To this end, we make the following assumptions.

- (D1)  $X \subset \mathbb{R}^k$  is a nonempty compact set. Without loss of generality, we assume  $X$  is continuous.
- (D2) For any  $x \in X$ ,  $G(x, \cdot)$  is measurable.
- (D3) There exists a measurable function  $\psi : \Omega \rightarrow \mathbb{R}_+$  such that  $|G(x, \omega)| \leq \psi(\omega)$  for every  $x \in X$  and  $\omega \in \Omega$ . Let  $\Psi := \mathbb{E}[\psi(\omega)]$ .

- (D4) The MGF of  $\psi(\omega)$ ,  $M_\psi(s) = \mathbb{E}[e^{s\psi(\omega)}]$  is finite valued in a neighborhood of zero.
- (D5) There exists a measurable function  $\phi : \Omega \rightarrow \mathbb{R}_+$  such that  $|G(x_1, \omega) - G(x_2, \omega)| \leq \phi(\omega)\|x_1 - x_2\|$ ,  $\forall \omega \in \Omega$  and  $x_1, x_2 \in X$ . Let  $\Phi := \mathbb{E}[\phi(\omega)]$ .
- (D6) The MGF of  $\phi(\omega)$ ,  $M_\phi(s) = \mathbb{E}[e^{s\phi(\omega)}]$  is finite valued in a neighborhood of zero.

**Lemma 1** *Let  $x_1, x_2 \in X$  and  $t_1, t_2 \in \mathbb{R}$ . Then under Assumptions (D1)-(D6), a.e.  $\omega \in \Omega$ ,*

$$|\tilde{G}_\alpha(x_1, t_1, \omega) - \tilde{G}_\alpha(x_2, t_2, \omega)| \leq \frac{\phi(\omega)}{1 - \alpha}\|x_1 - x_2\| + \frac{2 - \alpha}{1 - \alpha}|t_1 - t_2|. \quad (10)$$

**Proof.** For brevity, we use  $\tilde{G}$  denote  $\tilde{G}_\alpha(x, t, \omega)$ . The Lipschitz continuity of  $\tilde{G}$  with respect to  $x$  can be shown as follows.

$$\begin{aligned} & |\tilde{G}_\alpha(x_1, t, \omega) - \tilde{G}_\alpha(x_2, t, \omega)| \\ &= \frac{1}{1 - \alpha} |(G(x_1, \omega) - t)_+ - (G(x_2, \omega) - t)_+| \\ &\leq \frac{1}{1 - \alpha} |G(x_1, \omega) - G(x_2, \omega)| \\ &\leq \frac{\phi(\omega)}{1 - \alpha} \|x_1 - x_2\|. \end{aligned}$$

Similarly,  $\tilde{G}$  is Lipschitz continuous with respect to  $t$  since

$$\begin{aligned} & |\tilde{G}_\alpha(x, t_1, \omega) - \tilde{G}_\alpha(x, t_2, \omega)| \\ &\leq |t_1 - t_2| + \frac{1}{1 - \alpha} |(G(x, \omega) - t_1)_+ - (G(x, \omega) - t_2)_+| \\ &\leq |t_1 - t_2| + \frac{1}{1 - \alpha} |t_1 - t_2| \\ &= \frac{2 - \alpha}{1 - \alpha} |t_1 - t_2|. \end{aligned}$$

Hence by triangle inequality, inequality (10) holds.  $\square$

**Lemma 2** *Suppose Assumptions (D1)-(D4) hold. There exists a closed interval  $T \subset \mathbb{R}$  such that*

$$\min_{x \in X, t \in T} \{f(x) : \mathbb{E}[\tilde{G}_\alpha(x, t, \omega)] \leq q\} \quad (11)$$

*is equivalent to Problem (9) where  $t$  is a free variable.*



**Proof.** Fixing  $x \in X$ , we know

$$\text{VaR}_\alpha[G(x, \omega)] \in \text{Argmin}_{t \in \mathbb{R}} \mathbb{E}[\tilde{G}_\alpha(x, t, \omega)].$$

Therefore, it is sufficient to bound  $\text{VaR}_\alpha[G(x, \omega)]$  for all  $x \in X$  instead of bounding  $t$ . According to Assumption (D3),

$$-\psi(\omega) \leq G(x, \omega) \leq \psi(\omega), \quad \forall x \in X \text{ and } \omega \in \Omega.$$

It follows that

$$\text{VaR}_\alpha[-\psi(\omega)] \leq \text{VaR}_\alpha[G(x, \omega)] \leq \text{VaR}_\alpha[\psi(\omega)], \quad \forall x \in X.$$

By Assumption (D4),  $\mathbb{E}[\psi(\omega)] < +\infty$ . Together with the fact  $\psi(\omega) \geq 0$ , it follows that  $\text{VaR}_\alpha[\psi(\omega)] < +\infty$  and  $\text{VaR}_\alpha[-\psi(\omega)] > -\infty$  for any  $\alpha \in (0, 1)$ . Hence we can define

$$T := [\text{VaR}_\alpha[-\psi(\omega)], \text{VaR}_\alpha[\psi(\omega)]].$$

□

**Lemma 3** *Under Assumptions (D1)-(D4), for any  $x \in X$  and  $t < \infty$ , the moment generating function of  $M_{x,t,\alpha}(\cdot)$  of  $\tilde{G}_\alpha(x, t, \omega) - \mathbb{E}[\tilde{G}_\alpha(x, t, \omega)]$  is finite around zero.*

**Proof.**

$$M_{x,t,\alpha}(s) := e^{-s\mathbb{E}[\tilde{G}_\alpha(x,t,\omega)]} \mathbb{E} \left[ e^{s\tilde{G}_\alpha(x,t,\omega)} \right].$$

The finiteness of  $\mathbb{E}[\tilde{G}_\alpha(x, t, \omega)]$  is trivial since  $\mathbb{E}[G(x, \omega) - t]_+ < +\infty$ , which is implied by  $\mathbb{E}|G(x, \omega)| \leq \mathbb{E}[\psi(\omega)] < +\infty$ . We need to prove the finiteness of  $\mathbb{E} \left[ e^{s\tilde{G}_\alpha(x,t,\omega)} \right]$ , which can be split into two terms

$$e^{\frac{s\alpha t}{\alpha-1}} \int_{\omega: G(x,\omega) > t} e^{\frac{s}{1-\alpha}G(x,\omega)} dP(\omega) + e^{st} \Pr\{\omega : G(x, \omega) \leq t\}.$$

Since

$$\int_{\omega: G(x,\omega) > t} e^{\frac{s}{1-\alpha}G(x,\omega)} dP(\omega) \leq \begin{cases} \mathbb{E} \left[ e^{\frac{s}{1-\alpha}\psi(\omega)} \right] & \text{if } s > 0, \\ \mathbb{E} \left[ e^{\frac{-s}{1-\alpha}\psi(\omega)} \right] & \text{o.w.} \end{cases}$$

by Assumption (D4), it is finite. So  $M_{x,t,\alpha}(\cdot) < +\infty$  in a neighborhood of zero.

□

Denote the feasible region of Problem (11) by  $Y^0$  and that of the SAA problem by  $Y_N^0$ . The following result now immediately follows from Proposition 2.

**Proposition 3** *Suppose Assumptions (D1)-(D6) hold. Given  $\epsilon > 0$ , then it is true that*

$$\Pr\{Y^{-\epsilon} \subseteq Y_N^0 \subseteq Y^\epsilon\} \geq 1 - 2 \left[ 1 + \left( \frac{D_x}{\nu} \right)^k \left( \frac{D_t}{\nu} \right) \right] e^{-\frac{N\epsilon^2}{8\sigma^2}},$$

where

- (1)  $k$  and  $D_x$  are the dimension and the diameter of  $X$ , respectively,
- (2)  $D_t$  is the diameter of  $T$ , e.g.,  $T = [\text{VaR}_\alpha[-\psi(\omega)], \text{VaR}_\alpha[\psi(\omega)]]$ ,
- (3)  $\nu := \left\{ \frac{1}{1-\alpha} \left[ \frac{4(\Phi^{-\alpha+2})}{\epsilon} + 1 \right] \right\}^{-1}$ , and
- (4)  $\sigma^2 := \max_{x \in X, t \in T} \{\text{Var}[\tilde{G}_\alpha(x, t, \omega) - \mathbb{E}[\tilde{G}_\alpha(x, t, \omega)]], \text{Var}[\phi(\omega)]\}$ .

### 3 Validation Scheme

The sample size estimates obtained in the previous section, although theoretically appealing, can be overly conservative in practice. So the natural question is what can we say about a candidate solution obtained by solving an SAA problem with a practical sample size  $N$ . To this end we now discuss schemes for obtaining and validating candidate solutions using the SAA problem.

From Lagrange duality, we know

$$\min_{x \in X} \{f(x) : g(x) \leq q\} \tag{12}$$

$$\geq \min_{x \in X} \{f(x) + \pi[g(x) - q]\} \tag{13}$$

for any  $\pi \geq 0$  and the equality holds when  $\pi$  is an optimal Lagrangian multiplier and there is no duality gap. Since (13) takes the form of traditional stochastic programming, we know how to generate a lower bound for its optimal objective value, which then is a lower bound for (12). As to upper bounding, any  $x \in X$  satisfying  $g(x) \leq q$  can provide an upper bound  $f(x)$ . However, for expected value constrained problems, it is difficult to check if  $x \in X$  is feasible to the original problem, i.e., whether it satisfies  $g(x) \leq q$ . Therefore, we need to associate an upper bound obtained via evaluating the objective function value  $f(x)$  at some  $x$  with an estimated probability that it is a true upper bound.

We obtain  $\tilde{x}$  for upper bounding and  $\tilde{\pi}$  for lower bounding via solving one SAA problem of the original problem. Let  $\{\omega_1, \dots, \omega_N\}$  be a sample of size  $N$ . Solve the approximate SAA problem,

$$\min_{x \in X} \{f(x) : g_N(x) \leq \tilde{q}\} \tag{14}$$

and let  $(\tilde{x}, \tilde{\pi})$  be an optimal primal-dual pair. Note  $\tilde{q} \leq q$ . In order to solve (14) efficiently,  $N$  cannot be too large. Then it is very possible that  $\tilde{x}$  from solving (14) with  $\tilde{q} = q$  is infeasible to the original problem. Therefore a smaller  $\tilde{q}$ -value in (14) can improve the chance that  $\tilde{x}$  is feasible. Let  $\{\omega_1, \dots, \omega_{N_u}\}$  be another sample of size  $N_u$ , where  $N_u \gg N$ . Compute

$$\tilde{u} := \frac{1}{N_u} \sum_{n=1}^{N_u} G(\tilde{x}, \omega_n)$$

and

$$S_{\tilde{u}}^2 := \frac{1}{N_u(N_u - 1)} \sum_{n=1}^{N_u} [G(\tilde{x}, \omega_n) - \tilde{u}]^2.$$

For  $N_u$  large enough, we know

$$\frac{g(\tilde{x}) - \tilde{u}}{S_{\tilde{u}}} \sim N(0, 1).$$

Define  $z_\beta$  by  $\Pr\{Z \leq z_\beta\} = 1 - \beta$ , where  $Z$  is a standard normal random variable and  $\beta \in [0, 1]$ . By computing

$$z_\beta = \frac{q - \tilde{u}}{S_{\tilde{u}}},$$

we can claim that

$$\Pr\{g(\tilde{x}) \leq q\} = 1 - \beta.$$

So it suffices to check  $z_\beta$ . If  $z_\beta$  is big enough, then we accept  $\tilde{x}$ . Otherwise we decrease  $\tilde{q}$  by some small value  $\xi$ , and resolve (14) to get another  $\tilde{x}$  and check it. After obtaining a satisfactory  $\tilde{x}$  with characteristic  $z_\beta$ , we can calculate  $f(\tilde{x})$  and conclude that the probability that  $f(\tilde{x})$  is an upper bound is  $1 - \beta$ . Below we summarize the SAA scheme for expected value constrained problems.

*Step 0:* Set  $\tilde{z} > 0$ ,  $\xi > 0$ ,  $\gamma \in (0, 1)$  and  $\tilde{q} = q$ .

*Step 1:* Optimal  $\tilde{\pi}$  estimation.

Generate a sample of size  $N$ , i.e.,  $\{\omega_1, \dots, \omega_N\}$  and solve the SAA problem

$$\min_{x \in X} \{f(x) : g_N(x) \leq \tilde{q}\}.$$

Let  $\tilde{x}$  and  $\tilde{\pi}$  be its optimal solution and Lagrange multiplier, respectively.

*Step 2:* Upper bound estimation.

Generate another independent sample of size  $N_u$ , i.e.,  $\{\omega_1, \dots, \omega_{N_u}\}$ . Compute  $g_{N_u}(\tilde{x})$ , its variance  $S_{N_u}^2(\tilde{x})$ , and

$$z_\beta = \frac{q - g_{N_u}(\tilde{x})}{S_{N_u}(\tilde{x})}.$$

If  $z_\beta < \tilde{z}$ , let  $\tilde{q} = \tilde{q} - \xi$  and go back to step 1; otherwise, get an upper bound  $\tilde{u}$  and the associated probability

$$\tilde{u} = f(\tilde{x}),$$

$$\Pr\{\tilde{u} \text{ is an upper bound}\} = 1 - \beta.$$

*Step 3: Lower bound estimation.*

Generate  $M_l$  independent samples each of size  $N_l$ , i.e.,  $\{\omega_1^m, \dots, \omega_{N_l}^m\}$  for  $m = 1, \dots, M_l$ . For each sample, solve the SAA problem

$$\hat{l}^m := \min_{x \in X} \left\{ f(x) + \tilde{\pi} \left[ N_l^{-1} \sum_{n=1}^{N_l} G(x, \omega_n^m) - q \right] \right\}.$$

Compute the lower bound estimator  $\tilde{l}$  and its variance  $S_{\tilde{l}}^2$  as follows

$$\tilde{l} := \frac{1}{M_l} \sum_{m=1}^{M_l} \hat{l}^m,$$

$$S_{\tilde{l}}^2 := \frac{1}{M_l(M_l - 1)} \sum_{m=1}^{M_l} (\hat{l}^m - \tilde{l})^2.$$

And a  $(1 - \gamma)$  confidence interval for  $\mathbb{E}[\tilde{l}]$  is

$$[\tilde{l} - z_{\gamma/2} S_{\tilde{l}}, \tilde{l} + z_{\gamma/2} S_{\tilde{l}}].$$

## 4 A Portfolio Selection Application

In this section we report on the performance of the proposed SAA scheme for a portfolio selection problem involving a CVaR constraint. A similar problem has been considered in [6] however there the authors used historical returns as the distribution of the asset returns, whereas we consider a continuous distribution and use the proposed SAA approach. We assume that the asset returns are multi-variate normal (with parameters estimated from historical data). This assumption allows us to compare the results of the SAA method to the exact optimal solution, since in this case a CVaR constraint can be reformulated as a deterministic nonlinear inequality. We first describe the problem instance, and then the computational results.

#### 4.1 Problem Instance

We consider the following CVaR constrained portfolio optimization model:

$$\begin{aligned}
\min \quad & \mathbb{E}[G(x, r)] \\
\text{s.t.} \quad & e^\top x = 1 \\
& \text{CVaR}_\alpha[G(x, r)] \leq q \\
& ly \leq x \leq uy \\
& k_l \leq e^\top y \leq k_u \\
& x \in \mathbb{R}^n, \ y \in \{0, 1\}^n
\end{aligned} \tag{15}$$

where  $n$  is the number of stocks under consideration;  $r$  is the random one-month cent-per-dollar return vector;  $x$  is the vector of fractions of the initial capital invested in the stocks;  $e \in \mathbb{R}^n$  is a vector of ones;  $G(x, r) := -r^\top x$  is the portfolio “loss”, i.e. negative of the return;  $y$  is a binary decision vector whose  $j$ -th element  $y_j = 1$  if stock  $j$  is invested in, and  $y_j = 0$  otherwise;  $l \in \mathbb{R}^n$  and  $u \in \mathbb{R}^n$  are vectors of upper and lower bounds on allocations for the stocks being considered for investment;  $k_l$  and  $k_u$  are vectors of lower and upper bounds on the number of stocks allowed in the portfolio. Note that  $\mathbb{E}[G(x, r)] = -(\mathbb{E}[r])^\top x$ .

If the return vector  $r$  is multivariate normal with mean  $\mu$  and covariance matrix  $\Sigma$ , then for  $\alpha \geq 0.5$  [10],

$$\text{CVaR}_\alpha[G(x, r)] = -\mu^\top x + c(\alpha) \sqrt{x^\top \Sigma x},$$

with  $c(\alpha) := \{\sqrt{2\pi}(1 - \alpha)\exp[\text{erf}^{-1}(2\alpha - 1)]^2\}^{-1}$ , where  $\text{erf}^{-1}$  denotes the inverse of the error function

$$\text{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt.$$

In this case (15) reduces to deterministic mixed-integer quadratically constrained problem (MIQCP) which can be solved by commercial packages such as CPLEX. In our experiments we compare the results obtained by using the SAA method on (15) with those obtained from the MIQCP reformulation assuming normal returns.

We consider an instance of (15) consisting of 95 stocks from the S&P100, i.e.,  $n = 95$ . The 95 stocks are from S&P100, excluding SBC, ATI, GS, LU, and VIA-B due to insufficient data. We use historical monthly prices between 1996 to 2002 from <http://finance.yahoo.com>. Stock splitting is adjusted. Assuming normal returns, we estimate the mean  $\mu$  and covariance  $\Sigma$  of the return vector

$r$  from the seven-year historical data. We require that the fraction invested on any stock is either 0 or in the interval  $[0.05, 0.25]$ , i.e.,  $l = 0.05e$  and  $u = 0.25e$ , and the total number of stocks invested in should be no fewer than 10 and at most 20, i.e.,  $k_l = 10e$  and  $k_u = 20e$ .

#### 4.2 Computational Results

We solve (15) using the proposed SAA method for different values of  $\alpha$  and  $q$  with the following parameters. The sample size  $N$  for obtaining a candidate solution is 2000, the sample size  $N_u$  for testing its feasibility is 50000, the sample size  $N_l$  (sample number  $M_l$ ) for lower bounding is 1000 (10), the small value  $\xi$  used for decreasing the right-hand-side term in the expected value constraint is 0.2, and the level  $\tilde{z}$  for accepting a candidate solution is 2. The SAA results are compared to the exact optimal solution obtained from the MIQCP reformulation. The SAA subproblems (deterministic mixed-integer linear programs) and the MIQCP reformulation are solved using the commercial package CPLEX.

Tables 1 and 2 present the computational results of the CVaR constrained model for different values of  $\alpha$  and  $q$ , respectively. In these tables, “Opt.Obj.” and “Time” are the optimal objective value and the solution time, respectively, from the MIQCP reformulation. All other rows correspond to (15) solved by the SAA scheme. “Pr{UB}” is the estimated probability that “UB” is a statistical upper bound on the true optimal value of the CVaR constrained model, or equivalently, the estimated probability that the optimal solution  $\tilde{x}$  corresponding to “UB” is feasible to the true CVaR constrained problem; “95%-C.I.(LB)-l” is the left end of the 95% confidence interval of the lower bound “LB;” “# Iterations” is the number of iterations taken for getting a statistical feasible solution (Step 1 in SAA for EVC); and finally, “UB Time,” “LB Time,” and “Total Time” are the time in seconds for upper bounding, lower bounding and the entire SAA implementation, respectively.

We make the following observations. The first is regarding the upper bounding of SAA. In our SAA implementation, 97.7% is the smallest probability at which an obtained SAA solution is accepted as feasible solution to the true CVaR constrained problem. It turns out that the actual probability “Pr{UB}” is much higher. For example, over 99.5% in eight out of ten runs of the CVaR constrained problem. The second observation is regarding solution quality of SAA. The tables provide two relative gaps  $\frac{\text{UB}-95\text{-C.I.}(\text{LB})-l}{|\text{UB}|}(\%)$  and  $\frac{\text{UB}-\text{Opt.Obj.}}{|\text{Opt.Obj.}|}(\%)$  as indicators of the solution quality estimated by SAA and by comparing the solutions of the SAA and that of the MIQCP reformulation. Consider the case where  $\alpha = 90\%$  and  $q = 10$ . With 95% confidence, we estimate that the SAA solution is within 8.01% of optimality, whereas it is in

fact only within 1.06% of optimality. This suggests that the SAA solutions are quite good, on the other hand the validation scheme is overly conservative. A possible explanation of this is that since the considered model is non-convex (owing to discrete variables) the Lagrangian duality gap weakens the lower bound estimates. Finally we comment on the computation time of the SAA approach. We observe that the solution times are significantly more than that of the exact MIQCP approach. However recall that the exact approach is only possible due to the normality assumption, whereas the SAA approach is applicable for general distributions. Moreover, we solve the SAA subproblems as large-scale mixed-integer linear programs. Using a more sophisticated decomposition based cutting plane scheme (e.g. [3,4]) may significantly improve solution times.

## 5 Conclusions

This paper has proposed a sample average approximation method for general expected value constrained stochastic programs, and specialized it for CVaR constrained programs. In particular, we have proved that results of the SAA problem converge to their counterparts for the true problem with probability approaching one exponentially fast, and we designed a validation scheme which provides approximate solutions and statistically upper and lower bound estimates on the true optimal value. Probability estimate of the upper bound and variance estimate of the lower bound are constructed. Computational results on a portfolio optimization application demonstrate that the performance of the proposed SAA scheme is quite satisfactory.

## References

- [1] S. Ahmed and A. Shapiro. The sample average approximation method for stochastic programs with integer recourse. Available online <http://www.isye.gatech.edu/~sahmed>, 2002.
- [2] J. Atlason, M. Epelman and G. Henderson. Call center staffing with simulation and cutting plane methods, *Annals of Operations Research* 127 (2004) 333-358.
- [3] C.I. Fabian. Handling CVaR objectives and constraints in two-stage stochastic models. To appear in *European Journal of Operational Research*, 2007.
- [4] C.I. Fabian and A. Veszpremi. Algorithms for handling CVaR-constraints in dynamic stochastic programming models with applications to finance. To appear in *Journal of Risk*, 2007.

- [5] A.J. Kleywegt, A. Shapiro and T. Homem-De-Mello. The sample average approximation method for stochastic discrete optimization, *SIAM Journal on Optimization* 12 (2001) 479-502.
- [6] P. Krokmal, J. Palmquist and S. Uryasev. Portfolio optimization with conditional value-at-risk objective and constraints, *The Journal of Risk* 4 (2002) 11-27.
- [7] D. Kuhn. Convergent bounds for stochastic programs with expected value constraints. Available online <http://edoc.hu-berlin.de/series/speps/2006-22/PDF/22.pdf>, 2006.
- [8] A. Nemirovski and A. Shapiro. Convex approximations of chance constrained programs, *SIAM Journal on Optimization* 17 (2006) 969-996.
- [9] M. O'Brien, Techniques for Incorporating Expected Value Constraints into Stochastic Programs, PhD Thesis, Standard University, 2000.
- [10] R.T. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk, *The Journal of Risk* 2 (2000) 21-41.
- [11] A. Ruszczyński and A. Shapiro (eds.), Stochastic Programming, in: Handbooks in OR & MS, Vol.10, Elsevier Science, Amsterdam, 2003.
- [12] A. Shapiro and H. Xu. Stochastic mathematical programs with equilibrium constraints, modeling and sample average approximation. Available online <http://www.optimization-online.org>, 2005.
- [13] D.W. Stroock, An Introduction to the Theory of Large Deviations, Springer-Verlag, New York, 1984.



Table 1. Effects of  $\alpha$  ( $q = 10$ )

$\alpha(\%)$	90	91	92	93	94	95	96	97	98	99
Opt.Obj.	-3.217	-3.181	-3.140	-3.09	-3.044	-2.988	-2.923	-2.844	-2.750	-2.597
Time(sec.)	1.99	3.33	3.84	6.46	19.64	15.12	13.19	20.43	16.95	3.67
UB	-3.183	-3.144	-3.100	-3.051	-3.003	-2.924	-2.864	-2.800	-2.699	-2.529
Pr{UB}{%}	100.00	100.00	100.00	100.00	97.83	100.00	99.97	99.92	99.67	98.97
LB	-3.421	-3.493	-3.494	-3.477	-3.428	-3.340	-3.263	-3.094	-2.972	-2.920
LBSD( $\times 10^2$ )	0.750	0.846	0.853	0.866	0.897	0.914	0.886	0.790	1.047	1.324
95%-C.I.(LB)-l	-3.438	-3.512	-3.514	-3.496	-3.449	-3.361	-3.283	-3.111	-2.996	-2.950
$\frac{\text{UB}-95\text{-C.I.}(\text{LB})-l}{ \text{UB} }(\%)$	8.01	11.70	13.35	14.59	14.85	14.95	14.63	11.11	11.00	16.65
$\frac{\text{UB}-\text{Opt.Obj.}}{ \text{Opt.Obj.} }(\%)$	1.06	1.16	1.27	1.39	1.35	2.14	2.02	1.55	1.85	2.62
UB Time(sec.)	122.46	125.92	132.23	127.85	142.92	185.52	182.50	183.18	228.67	309.15
LB Time(sec.)	30.83	34.81	54.10	54.92	60.63	55.62	61.08	55.97	65.00	63.88
Total Time(sec.)	153.29	160.73	186.33	182.77	203.55	241.14	243.58	250.19	293.67	373.03

Table 2. Effects of  $q$  ( $\alpha = 95\%$ )

$q$	5	6	7	8	9	10	11	12	13	14
Opt.Obj.	-2.098	-2.360	-2.569	-2.734	-2.866	-2.988	-3.095	-3.190	-3.264	-3.316
Time(sec.)	28.09	15.06	2.92	12.33	16.17	15.11	8.01	3.29	3.39	15.87
UB	-2.015	-2.268	-2.484	-2.689	-2.822	-2.924	-3.060	-3.165	-3.245	-3.300
Pr{UB}{%}	99.91	100.00	100.00	99.92	100.00	100.00	99.95	99.94	99.09	100.00
LB	-2.352	-2.646	-2.851	-3.000	-3.209	-3.340	-3.427	-3.518	-3.312	-3.335
LBSD( $\times 10^2$ )	1.593	1.357	1.113	1.012	1.000	0.914	0.764	0.704	1.198	1.069
95%-C.I.(LB)-1	-2.388	-2.676	-2.876	-3.022	-3.232	-3.361	-3.444	-3.534	-3.339	-3.359
$\frac{\text{UB}-95\text{-C.I.}(\text{LB})-1}{ \text{UB} }(\%)$	18.51	17.99	15.78	12.38	14.53	14.95	12.55	11.66	2.90	1.79
$\frac{\text{UB}-\text{Opt.Obj.}}{ \text{Opt.Obj.} }(\%)$	3.96	3.90	3.31	1.65	1.54	2.14	1.13	0.78	0.58	0.48
UB Time(sec.)	246.89	297.68	246.04	145.16	143.56	187.94	127.89	122.70	117.10	115.29
LB Time(sec.)	123.78	98.55	82.38	70.45	64.56	64.50	52.01	36.44	20.72	14.52
Total Time(sec.)	370.67	396.23	328.42	215.61	208.12	252.44	179.90	159.14	137.82	129.81