Proceedings of the 2008 Winter Simulation Conference
S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, J. W. Fowler eds.

A RATE RESULT FOR SIMULATION OPTIMIZATION WITH CONDITIONAL VALUE-AT-RISK CONSTRAINTS

Soumyadip Ghosh

Business Analytics and Math Sciences Division
1101 Kitchawan Road
Yorktown Heights, NY 10598, USA.

ABSTRACT

We study a stochastic optimization problem that has its roots in financial portfolio design. The problem has a specified deterministic objective function and constraints on the conditional value-at-risk of the portfolio. Approximate optimal solutions to this problem are usually obtained by solving a sample-average approximation. We derive bounds on the gap in the objective value between the true optimal and an approximate solution so obtained. We show that under certain regularity conditions the approximate optimal value converges to the true optimal at the canonical rate $O(n^{-1/2})$, where n represents the sample size. The constants in the expression are explicitly defined.

1 INTRODUCTION

Financial markets have seen an explosive growth in the number of investment vehicles available, each of which comes with its own risk-to-reward tradeoff. As the industry gathers more knowledge and experience with various exotic investment opportunities, it has become increasingly clear that a portfolio manager must actively seek to assess and manage the risk inherent in a portfolio. Of particular interest is the fact that though each option's risk on returns might be easy to determine, the nature of the *joint* risk or volatility in a diverse portfolio is relatively less understood.

This article concerns itself with a portfolio manager's task of designing a portfolio by allocating all or part of a budget over a fixed set of high-return but also high-risk assets. Let random variables $\{\xi_k, k = 1, \ldots, d\}$ represent the change in value of investment vehicle k over a fixed time interval. Denote by $x \in \mathbb{R}^d$ how the marginal dollar is divided amongst the d investments. Then, we concern ourselves with the stochastic program

$$\max_x \{\, c^t x \quad | \quad x \in X \cap X^0 \text{ and } R(g(\xi,x)) \le b \,\}, \qquad (1)$$

where the R-constraints are risk-based. The allocation x is scaled to that of a nominal dollar, and so $x \in X^0 \stackrel{\triangle}{=} \{\, x \in \mathbb{R}^d_+, \quad \sum_k x_k \le 1\}$, a subset of the non-negative orthant $\mathbb{R}^d_+$. The set $X \subset \mathbb{R}^d$ is a convex polytope that represents any additional (deterministic) constraints on the chosen allocation. The cost coefficients c are known and deterministic and can be thought of as the total revenue from a portfolio with allocation x.

The function g defines a random outcome that depends both on the choice of x and an independent random variable ξ in space $\mathbb{R}^d$. It represents a notion of loss experienced with decisions x and the change ξ in values of underlying random variables. The function R is said to be the (deterministic) *risk* inherent in the loss measure g for a particular choice of x. We shall limit this article to the case when the function g has a one-dimensional linear functional form $g(\xi,x) = -\xi^t x$.

The stochastic program thus seeks to find the maximum revenue portfolio allocation from a set of feasible allocations where no x results in a risk of more than b. Markowitz (1952), who laid the foundation to the portfolio optimization and management theory, considered the mean-variance relation as representative of the risk-returns tradeoff. Since then various measures of risk have been studied in this framework. J P Morgan's RiskMetrics^{TM} (1996) was an important advocate for the use of the *value-at-risk* measure in financial portfolio management. The VaR $V_\beta(\cdot)$ at level β is the lowest potential loss that may occur with probability $1-\beta$, and is thus is a natural candidate as a risk measure. It is indeed widely used in the industry.

The VaR measure has been known to exhibit behavior that might run counter to expectations, which limits their effectiveness to special conditions. Artzner et al. (1999) argue that risk measures should satisfy the following conditions in order to be *coherent*: they should be translation invariant, sub-additive, positive homogeneous and monotone in the random variable. The VaR measure fails, for instance, the sub-additivity test.

The risk measure R we consider is the *conditional value-at-risk* (CVaR) at level β, denoted as $C_\beta(\cdot)$. The CVaR risk

978-1-4244-2708-6/08/$25.00

measure, defined in terms of the VaR risk measure $V_\beta(\cdot)$, and has garnered a lot of attention recently. (Definitions for both are given in Section 2.) The CVaR $C_\beta(\cdot)$ represents the average loss experienced given that the loss is greater than $V_\beta(\cdot)$. Pflug (2000) and Acerbi and Tasche (2002) show that the CVaR measure is coherent. Coherent measures possess strong functional properties that make them more amenable to use in a wide variety of applications, and are now widely accepted in the academic community.

The problem of estimating CVaR measures $C_\beta(\mathbf{Y})$ for a random variable $\mathbf{Y}$ is of interest in itself, and it is particularly so as a rare-event estimation problem (Juneja and Shahabuddin 2007, provide a good review). This is because the estimation requires generating samples from a low-probability set when β is close to 1, as is typical in practice. The estimation of VaR has been well studied under this framework: for instance, Glasserman et al. (2000) provide a general variance-reduction framework to estimate VaR for light-tailed distributions, while Glasserman et al. (2002) look at the estimation problem for heavy-tailed $\mathbf{Y}$. Focus is now beginning to shift to the CVaR measure, which awaits a similar thorough treatment.

Rockafellar and Uryasev (2000) introduce a different estimator that is perhaps more suited for optimization applications like (1). An approximation problem constructed using their estimator usually results in a convex or even a linear program (Rockafellar and Uryasev 2002), and thus leads to efficient implementations for large-scale problems of the type (1).

Problem (1) is a member of the general class of stochastic problems that include a constraint involving an expectation which, in general, can not be written down in closed form. The simulation literature provides a diverse set of tools to tackle such stochastic convex problems. A standard approach called *sample average approximation* estimates the expectation of the random function via samples of the underlying random variable and then constructs a constraint that approximates the true CVaR constraint in (1). One usually expects the solution to the approximated problem to be approximately close to the true optimal. In Section 3, we provide a bound on the *relative optimality gap* between the approximation solution generated by a sample-average algorithm and a true optimal of (1). For a sufficiently large sample size n, the solution found has an objective value within $O(n^{-1/2})$ of the optimal. Wang and Ahmed (2007) provide bounds of the same order on the quality of sample-average approximation solutions to general convex problems with stochastic constraints. Their results are derived using large deviations theory and require $R(g(\xi,x))$ to satisfy certain conditions, which do hold in the case we study. We derive bounds with similar rates of convergence, but we provide a geometric argument using the coherence properties of $C_\beta(x)$. In our case, the constant in the expression is defined in a manner that can be calculated analytically in some cases, or estimated. Moreover, the constant does not include the r.h.s. b in its definition.

The rest of the article is laid out thus: Section 2 provides the mathematical background on the problem we are interested in. Section 3 describes the main result of this paper, and Section 4 discusses some directions these results can be extended in and/or utilized.

2 THE OPTIMIZATION PROBLEM

The β-VaR $V_\beta(\mathbf{Y})$ of a random variable $\mathbf{Y}$ is the $(1-\beta)$-th quantile of $\mathbf{Y}$, and is defined as

$$V_\beta(\mathbf{Y}) \triangleq F_{\mathbf{Y}}^{-1}(1-\beta) = \inf_{y\in\mathbb{R}} \quad \{P(\mathbf{Y}>y) \le 1-\beta\}.$$

The conditional value-at-risk $C_\beta(\mathbf{Y})$ of a random variable $\mathbf{Y}$ *with a continuous distribution* is

$$C_\beta(\mathbf{Y}) = E[\mathbf{Y}|\mathbf{Y} \ge V_\beta(\mathbf{Y})].$$

In problem (1), we are interested in the CVaR $C_\beta(g(\xi,x))$ of a portfolio with the allocation x, and we treat only the case $g(\xi,x) = -\xi^t x$ here. We shall use the shorthand $C_\beta(x)$ for $C_\beta(-\xi^t x)$. The CVaR of a random variable with non-continuous distributions is harder to define. The continuous-distribution requirement on $\xi^t x$ is usually not overly restrictive. In general, the linear functional $\xi^t x$ has a continuous distribution if even one of the components of ξ have a continuous distribution. Henderson (2007) note that the distribution function of the convolution $\xi^t x$ can be obtained by conditioning on a component ξ_k that has a continuous distribution with a density, which leads to an expression for convolution $\xi^t x$'s density.

Function $C_\beta(\mathbf{Y})$ is sub-additive, positive homogeneous, translation invariant and monotone in $\mathbf{Y}$. The first two properties, in particular, imply that $C_\beta(\mathbf{Y})$ is convex in $\mathbf{Y}$. Since our choice of g is linear in x, $C_\beta(x)$ is also sub-additive and *concave* in x, and the feasible region carved out by the CVaR constraint, a level-set, is convex. The stochastic program (1) can thus be rewritten as a convex optimization problem:

$$\max \quad c^t x \quad \text{s.t.} \quad C_\beta(x) \le b, \quad x \in X^0 \cap X. \tag{2}$$

Random variables $\{\xi_k,\ k=1,\dots,d\}$ represent the change in real value of the assets under consideration over the fixed decision time-period. Let μ represent the mean and Σ the correlation structure of ξ. In this exercise, one would typically consider only those investments that have an expected net positive growth outcome $E\xi_k = \mu_k > 0$. To avoid trivialities, we additionally assume the ξ_k to satisfy

Assumption 1 *There exists a positive constant δ such that $C_\beta(\xi_k) \ge \delta > 0$, $\forall k = 1,\dots,d$.*

This implies that each asset considered is expected to net us a positive return μ_i, but it is accompanied with distribution tails that are fat enough to result in a risk of a positive loss at the β risk-tolerance level. This assumption is reasonable in most cases since instruments that violate it typically do not result in positive real returns. Instruments such as Treasury-bills that are traditionally considered "safe" or relatively risk-free are typically expected to only track inflation in value. From formulation (2) we see that an optimal allocation x^* might not sum to 1, and the rest of the marginal dollar is assumed to be invested in such safe instruments.

The program (2) is a convex problem, and thus large instances could potentially be solved efficiently. The principle difficulty lies in the fact that the $C_\beta(\cdot)$ constraints cannot be written down in an explicit form given the distribution of ξ. One approach to overcoming this difficulty is to construct a sample approximation of the original problem (2), in which the CVaR constraint is replaced with an estimate that is constructed based on samples of the random variable ξ. The *sample average approximation* of the problem (2) is of the general form:

$$\max c^t x \qquad \text{s.t.} \quad \hat{C}_\beta(x) \le b, \quad x \in X^0 \cap X, \tag{3}$$

where $\hat{C}_\beta(\cdot)$ is an estimator of $C_\beta(\cdot)$ constructed from samples of ξ. The approximation $\hat{C}_\beta(\mathbf{Y})$ can be provided using a canonical estimator of the expected value of the random variable $\mathbf{Y}|\mathbf{Y} \ge V_\beta(\mathbf{Y})$, because by definition $C_\beta(\mathbf{Y})$ is the expected value of a random variable that follows the same distribution as $\mathbf{Y}$ to the right of the (fixed) point $V_\beta(\mathbf{Y})$.

Rockafellar and Uryasev (2000) propose a new sample-average based estimator to the function $C_\beta(x)$:

$$\min_{\alpha \in \mathbb{R}} \left\{ \alpha + \frac{1}{1-\beta} \int_{y \in \mathbb{R}^d} [y^t x - \alpha]^+ p_\xi(y) dy. \right\}$$

This estimator has been designed with optimization in mind, and is convex in x for our choice of a linear g. The sample average approximation problem (3) is then a convex program.

The results we derive in Section 3 shall assume that the estimator $\hat{C}_\beta(x)$ satisfies the following conditions:

Assumption 2 *2a. The estimator $\hat{C}_\beta(\mathbf{Y})$ of $C_\beta(\mathbf{Y})$ is consistent and satisfies a central limit theorem of form:*

$$\sqrt{n}(\hat{C}_\beta(\mathbf{Y}) - C_\beta(\mathbf{Y})) \Rightarrow \sigma N(0,1) \tag{4}$$

where σ^2 is the variance associated with the estimation,

2b. The estimator $\hat{C}_\beta(x)$ of $C_\beta(x)$ is positive homogeneous in x, and

2c. Let $\sigma(x)$ be the CLT (4) variance for random variable $\xi^t x$ and set $\Theta \stackrel{\triangle}{=} \{\theta \in \mathbb{R}^d_+ \mid \sum_k \theta_k = 1\}$. Then, the supremum $\sup_{\theta \in \Theta} \sigma(\theta)$ exists.

Set Θ is the collection of allocations where the entire nominal dollar is invested in the d assets under consideration, and thus is one of the faces of the boundary of set X^0.

These assumptions are not overly restrictive, and reasonable estimators are expected to satisfy this condition. Lemma 1 shows that the canonical estimator complies. Let $\{\xi_i : i = 1, \ldots, N\}$ be N i.i.d. samples of ξ, and $\{\xi^t_{(i)} x\}$ be a non-decreasing ordering of the $\xi^t_i x$ values. Then, the canonical estimator of $C_\beta(x)$ at point x is

$$\hat{C}_\beta(x) \stackrel{\triangle}{=} \frac{1}{N - \lceil N\beta \rceil} \sum_{i=\lceil N\beta \rceil}^{N} \xi^t_{(i)} x. \tag{5}$$

Lemma 1 *The estimator (5) satisfies Assumption 2.*

Proof: For notational convenience, let $\mathbf{Y}(x) = -\xi^t x$, and we shall drop the argument x of $\mathbf{Y}(x)$ when the context makes the meaning clear. Note that $E[\mathbf{Y}(x)] = \mu^t x$ and $\text{Var}[\,\mathbf{Y}(x)\,] = x^t \Sigma x$. The set being optimized over is compact, so it will be sufficient to show that the variance of the random variables $\{\mathbf{Y}(\theta)|\mathbf{Y}(\theta) \ge V_\beta(\theta)\}$ are bounded above to prove 2c. The variance can be written as

$$\begin{aligned}
&= E[\mathbf{Y}^2|\mathbf{Y} \ge V_\beta(\theta)] - E^2[\mathbf{Y}|\mathbf{Y} \ge V_\beta(\theta)] \\
&\le E[\mathbf{Y}^2]\beta - \delta^2 \\
&= \{\text{Var}[\,\mathbf{Y}\,] + E^2[\mathbf{Y}]\}\beta - \delta^2 = \beta\{x^t\Sigma x + (\mu^t x)^2\} - \delta^2.
\end{aligned}$$

The first inequality uses Assumption 1. The upper bound is a quadratic in x, which attains a finite maximum within the compact set Θ. Thus, the variance term in assumption 2a is finite and assumption 2c is satisfied. Assumption 2b is true because of the linear form of (5) and the fact that the ordering in (5) does not change if x is scaled by a positive value.□

The Rockafellar and Uryasev (2000) estimator can be verified to satisfy Assumption 2b. It is not immediately clear that it obeys a limit theorem such as (4), though we suspect that this is the case.

3 OPTIMALITY GAP

The central limit theorem (4) obeyed by the CVaR estimator $\hat{C}_\beta(x)$ ultimately leads us to our central result Theorem 1, which demonstrates that the optimal objective value $c^t \hat{x}^*_n$ output for the approximation problem (3) is within a relative gap of $O(n^{-1/2})$ of the optimal objective value $c^t x^*$ of the original problem (2) for sufficiently large n. The convergence rate is as we should expect given the CLT (4), but interestingly the limit does not depend on b. Let $\phi(\cdot)$

represent the distribution function of the standard normal distribution $N(0,1)$ and x^* be an optimal solution to the original problem (2).

Theorem 1 *The objective value $c^t x_n^*$ of the optimal solution returned for the n-sample approximation problem (3) (with a sufficiently large n) satisfies*

$$a) \quad c^t x_n^* \leq c^t x^* \cdot \left\{1+\sqrt{\tfrac{M}{n}}\phi^{-1}(p)+O(\tfrac{1}{n})\right\}$$

$$b) \quad c^t x_n^* \geq c^t x^* \cdot \left\{1+\sqrt{\tfrac{M}{n}}\phi^{-1}(p)\right\}^{-1}$$

with probability p. The constant M is defined in Lemma 2, and is independent of parameters c or b in (3).
The inequalities should be interpreted in the same sense as when used in deriving standard confidence intervals for sampling-based estimators.

We shall provide a set of preliminary results that will in turn lead to the proof of Theorem 1. In the proof we show that the feasible set created by the sample average approximation problem, convex or not, is contained within a scaled-up version of the convex feasible set of the original problem (2). In turn, the approximation feasible set contains a scaled-down version of the original convex set. The scaling parameters are bounded by $O(n^{-1/2})$ terms, which gives us Theorem 1. The idea is demonstrated for the $\mathbb{R}^2$ case in Figure 1.

We start with the ratio of the variance $\sigma^2(x)$ and the square of the estimator $C_\beta(x)$, otherwise known as the *coefficient of variation* of the random variable $\xi^t x | \xi^t x \geq V_\beta(x)$.

Lemma 2 *Define*

$$M \stackrel{\triangle}{=} \max_\theta \{\frac{\sigma^2(\theta)}{C_\beta^2(\theta)} \ : \ \theta \in \Theta\}.$$

M exists and is finite.

Proof: The function $C_\beta(x)$ is concave in its argument x, and hence $C_\beta(\theta) \geq \sum_k \lambda_k C_\beta(e_k) \geq \delta$. Here the λ_k constitute a convex combination, and e_k represent the unit vector with one in the k^{th} component, and we also use Assumption 1. Thus, the term $1/C_\beta(\theta)$ is bounded above by the constant $1/\delta^2$. This, combined with Assumption 2c, gives the result. □

The constant M depends on the distribution of ξ in (2). In many cases, M can be explicitly evaluated; for instance when ξ are multi-variate normally distributed as $N(\mu, \sigma)$, explicit expressions for $\sigma(\theta)$ and $C_\beta(\theta)$ can be written down and shown to be quadratic functions of x, and the optimal value over Θ can then be determined. At first glance the fact that M does not use c or b in its definition seems remarkable, but this is to be expected given the strong positive scaling property of $C_\beta(\cdot)$.

We need some more notation to state our next result. Let Ω represent the intersection of the convex set defined by the CVaR constraint in (2) with the non-negative orthant $\mathbb{R}^d_+$, i.e. $\Omega \stackrel{\triangle}{=} \{C_\beta(x) \leq b, \quad x \in \mathbb{R}^d_+\}$, and Ω_n correspondingly its approximating set constructed by $\hat{C}_\beta(x)$ in (3). The set Ω_n need not necessarily be convex. We denote the boundary of a set A by ∂A. For an $x \in \partial\Omega$, let $\theta(x) = x/||x||_1$ and $\bar{r}(x) = ||x||_1$. (The l_1-norm of x is $||x||_1 = \sum_k |x_k|$.) Denote by $P_n(x)$ the point in $\partial\Omega_n$ that lies along $\theta(x)$. Let and $\bar{r}_n(x) = ||P_n(x)||_1$. Figure 1 pictorially depicts these definitions.

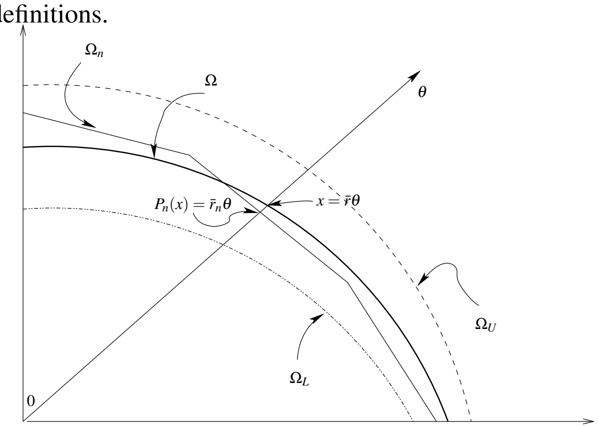

Figure 1: The sets Ω, Ω_n, Ω_L and Ω_U in $\mathbb{R}^2$.

Lemma 3 *For each $x \in \partial\Omega$,*

$$\bar{r}\left\{1+\sqrt{\frac{M}{n}}\phi^{-1}(p)\right\}^{-1} \leq \bar{r}_n \leq \bar{r}\left\{1-\sqrt{\frac{M}{n}}\phi^{-1}(p)\right\}^{-1},$$

with probability p.

Proof: The function $C_\beta(x)$ and its estimator $\hat{C}_\beta(x)$ (as defined in (5)) are both positive homogeneous in $x \in \mathbb{R}^d$. We have that for any $x \in \Omega$,

$$\begin{aligned} |C_\beta(x) - \hat{C}_\beta(x)| &= \bar{r}|C_\beta(\theta) - \hat{C}_\beta(\theta)| \\ &\lessapprox \bar{r} \cdot \frac{\sigma(\theta)}{\sqrt{n}}\phi^{-1}(p). \end{aligned} \quad (6)$$

The first equation uses the homogeneity of the functions (Assumption 2b). The second inequality holds for a probability p following the standard confidence-interval derivation using the CLT (4).

For an x in the boundary $\partial\Omega$, the constraint $C_\beta(x) \leq b$ is tight. Thus $|\hat{C}_\beta(x) - C_\beta(x)| = |\hat{C}_\beta(x) - b|$. Moreover from the definition of P_n and $\bar{r}_n$, $\hat{C}_\beta(x) = \bar{r}\hat{C}_\beta(\theta) = \bar{r}\, b/\bar{r}_n$. In other words,

$$\begin{aligned} \left|\frac{\bar r}{\bar r_n} - 1\right| &= \frac{1}{b}|\hat C_\beta(x) - b| \\ &\lessapprox \frac{1}{b}\left(\frac{b}{C_\beta(\theta)}\right)\cdot\frac{\sigma(\theta)}{\sqrt{n}}\phi^{-1}(p), \\ &\le \sqrt{M}\frac{1}{\sqrt{n}}\phi^{-1}(p), \end{aligned}$$

with probability p. The second inequality uses (6), while the last uses Lemma 2. The last inequality can be refashioned to give the relation required in the statement of the lemma. □

For a set S and a scalar a, let $aS = \{ax \mid x \in S\}$. Lemma 3 leads to the following corollary:

Corollary 2 *For n sufficiently large,*

$$a)\quad \Omega_L \triangleq \Omega\cdot\left\{1+\sqrt{\frac{M}{n}}\phi^{-1}(p)\right\}^{-1} \subseteq \Omega_n \quad (7)$$

$$b)\quad \Omega_n \subseteq \Omega_U \triangleq \Omega\cdot\left\{1-\sqrt{\frac{M}{n}}\phi^{-1}(p)\right\}^{-1} \quad (8)$$

with probability p.

Proof: We shall prove the upper bound and the lower bound follows similarly. For any $x \in \Omega_n$, we need to show that $x \in \Omega_U$. Write $x = r(x)\theta$ where θ is a unit vector from Θ. Consider the ray $\{r\theta,\ r > 0\}$ along θ. As before, let $\bar r$ represent the r-value that defines the ray-part contained within Ω (i.e., $\bar r = \max_r\{r\theta \in \Omega\}$ and $\bar r\theta \in \partial\Omega$), and similarly $\bar r_n$ and $\bar r_U$ for Ω_n and Ω_U respectively. From Lemma 3 we have that (w.p. p) $\bar r_n \le K\bar r$ for some constant K independent of θ. Moreover, $\bar r_U = K\bar r$. Thus, $r(x) \le \bar r_U$ and $x \in \Omega_U$. This holds for any $\theta \in \Theta$ and thus any $x \in \Omega_n$, and this establishes a). □

We now have all the results we need and shall proceed to prove Theorem 1.

Proof of Theorem 1: We prove part a) here; the other part follows in a similar fashion. The feasible region of the original problem (2) is given by $\Omega \cap (X^0 \cap X)$. Similarly, $\Omega_n \cap (X^0 \cap X)$ defines the feasible set of the approximation problem (3). The literature on *convex bodies* (closed, compact, convex sets; cf. Schneider 1993) tells us that if (7) holds, then so does

$$(\Omega_n \cap (X^0 \cap X)) \subseteq (\Omega_U \cap (X^0 \cap X)). \quad (9)$$

Let x^* be an optimal solution to (2). Now, for any scalar constant $a > 0$, $c^t(ax^*) \ge c^t(ax)$, $\forall x \in (\Omega \cap X^0 \cap X)$, i.e., ax^* is also optimal for objective $c^t x$ over $x \in a(\Omega \cap X^0 \cap X)$. Combining this with (9), we have that

$$\begin{aligned} c^t(x_n^*) &\le c^t x^*\cdot\left\{1-\sqrt{\frac{M}{n}}\phi^{-1}(p)\right\}^{-1} \\ &= c^t x^*\cdot\left\{1+\sqrt{\frac{M}{n}}\phi^{-1}(p)+O(\frac{1}{n})\right\}. \end{aligned}$$

The final expression uses the expansion for $(1-x)^{-1}$ when $0 \le x < 1$, which is true for sufficiently large n. □

Remark 1 *The arguments in this section primarily use the coherence properties of $C_\beta(\cdot)$ as described by Artzner et al. (1999), and so these results presumably hold if $C_\beta(\cdot)$ is interchanged with other coherent measures.*

Remark 2 *This article treats only the case of a linear uni-dimensional function $g(\xi,x) = -\xi^t x$. The results above can be generalized to multi-dimensional linear settings. Non-linear functions that satisfy certain properties might also be good candidates: for instance, functions that are convex and positive homogeneous, or Lipschitz continuous.*

4 FUTURE DIRECTIONS

The constant M that appears in the rate relation in Theorem 1 depends on the CVaR estimator used via the CLT (4). This leads to the natural suggestion that an estimator can be designed such that it has a lower value of M, or even minimizes it, which will in turn lead to better estimation of the true optimal value for the same sample size n. Finding such an estimator falls under the general purview of variance reduction, with the added twist that the M is defined as the maximum *coefficient-of-variation*. Whether this demands special attention when constructing variance-reduction estimators over applying standard techniques is not clear, but the question definitely merits further investigation.

REFERENCES

Acerbi, C., and D. Tasche. 2002. On the coherence of expected shortfall. *Journal of Banking and Finance* 26:1487–1503.

Artzner, P. F., J. M. Eber, and D. Heath. 1999. Coherence measures of risk. *Mathematical Finance* 9:203–228.

Glasserman, P., P. Heidelberger, and P. Shahabuddin. 2000. Variance reduction techniques for estimating Value-at-Risk. *Management Science* 46:1349–1364.

Glasserman, P., P. Heidelberger, and P. Shahabuddin. 2002. Portfolio value-at-risk with heavy-tailed risk factors. *Math. Finance* 12:239–269.

Henderson, S. G. 2007. Mathematics for simulation. In *Handbooks in Operations Research and Management*

Science: Simulation, ed. B. L. Nelson and S. G. Henderson, 19–54. Amsterdam: Elsevier Science.

Juneja, S., and P. Shahabuddin. 2007. Rare event simulation techniques: An introduction and recent advances. In *Handbooks in Operations Research and Management Science: Simulation*, ed. B. L. Nelson and S. G. Henderson, 291–350. Amsterdam: Elsevier Science.

Markowitz, H. M. 1952. Portfolio selection. *Journal of Finance* 7:77–91.

Pflug, G. 2000. Some remarks on the value-at-risk and conditional value-at-risk. In *Probabilistic Constrained Optimization: Methodology and Applications*, ed. S. Uryasev. Dordrecht: Kluwer.

RiskMetrics^{TM} 1996. *Technical document, 4-th edition.* J. P.Morgan.

Rockafellar, R. T., and S. Uryasev. 2000. Optimization of conditional value-at-risk. *Journal of Risk* 2:21–41.

Rockafellar, R. T., and S. Uryasev. 2002. Conditional Value-at-Risk for general loss distributions. *Journal of Banking and Finance* 26:1443–1471.

Schneider, R. 1993. Convex surfaces, curvature and surface area measures. In *Handbook of Convex Geometry*, ed. P. M. Gruber and J. M. Wills, Volume A, 273–299. Amsterdam: Elsevier Science.

Wang, W., and S. Ahmed. 2007. Sample average approximation of expected value constrained stochastic programs. Technical report, Submitted for publication.

AUTHOR BIOGRAPHIES

SOUMYADIP GHOSH is a Research Staff Member of the Mathematical Sciences Division at IBM T.J. Watson Research Center, Yorktown Heights, NY. His simulation research interests include theory and practice in particular on input dependence modeling and risk modeling, and simulation-based optimization techniques. His other interests are in the field of supply-chain analysis, and queueing theory based scheduling of large-scale production systems. He can be contacted at <ghoshs@us.ibm.com>.