

BUILDING METAMODELS FOR QUANTILE-BASED MEASURES USING SECTIONING

Xi Chen

Statistical Sciences and Operations Research
Virginia Commonwealth University
Richmond, VA 23284, USA

Kyoung-Kuk Kim

Industrial and Systems Engineering
KAIST
Daejeon 305-701, SOUTH KOREA

ABSTRACT

Simulation metamodeling has been used as an effective tool in predicting the mean performance of complex systems, reducing the computational burden of costly and time-consuming simulation runs. One of the successful metamodeling techniques developed is the recently proposed stochastic kriging. However, standard stochastic kriging is confined to the case where the sample averages and sample variances of the simulation outputs at design points are the main building blocks for creating a metamodel. In this paper, we show that if each simulation output is further comprised of i.i.d. observations, then it is possible to extend the original framework into a more general one. Such a generalization enables us to utilize estimation methods including sectioning for obtaining point and interval estimates in constructing stochastic kriging metamodels for performance measures such as quantiles and tail conditional expectations. We demonstrate the superior performance of stochastic kriging metamodels under the generalized framework through some examples.

1 INTRODUCTION

Simulation is known as a powerful tool for analyzing complex systems and predicting their performance to support decision making. However, fast-changing business environments and ever more complicated system designs are making maintain decision-making in a responsive and effective fashion increasingly difficult even with state-of-the-art computing power. In this regard, simulation metamodels have gained increasing popularity across engineering disciplines over recent years as efficient substitutes for simulation models, thanks to their abilities in accurately capturing the relationships between inputs and outputs of complex systems with less resources invested. Among the recent developments in the stochastic simulation metamodeling literature, stochastic kriging proposed by Ankenman et al. (2010) has drawn some attention as an effective metamodeling technique for approximating the mean response surface implied by a stochastic simulation. However, the mean response as a performance measure fails to take into account uncertainties or risks associated with the response of interest.

The main research question addressed in this paper is how to utilize existing methods for point and interval estimation to improve the performance of stochastic kriging in predicting quantile-based performance measures. To start, let us take quantile or percentile estimation as an example. The p th quantile is formally defined by $v_p = \inf\{x : F(x) \geq p\}$ where $F(\cdot)$ denotes the underlying distribution function of a random variable of interest, say L . There exists a rich body of literature in compliance with real-world applications of the estimation techniques. For example, JPMorgan and Reuters (1996) explain the use of a quantile-based risk measurement called value-at-risk (VaR) to monitor and control the riskiness of portfolios held by banks; and the announcement of Baseball Prospectus (2013) on March 28, 2013 provides the method and reasoning behind using percentiles in predicting baseball players' performances. The importance of quantile estimation has drawn considerable attention from the simulation research community over the last several decades. Seila (1982), Chen and Kelton (1999), and Alexopoulos et al. (2012) study quantile estimation

for regenerative or more general stationary processes; the applications of variance reduction techniques are studied by Hsu and Nelson (1990) and Avramidis and Wilson (1998). More closely related to this paper, Nakayama (2012) combines the sectioning method, introduced in Asmussen and Glynn (2007), with importance sampling for building confidence intervals for quantiles.

In this paper, we are concerned with the problem of building more efficient metamodels for quantile-based response measures in a design space of some design variables that parameterize the underlying distribution function $F(\cdot)$. Based on a sample of i.i.d. outputs drawn from $F(\cdot)$ at each design point, the default approach adopted by standard stochastic kriging in constructing response estimates and associated variance estimates for building a metamodel is identified as batching in Section 2. We generalize the standard formulation so that existing methods for point and interval estimation can be utilized to improve the performance of stochastic kriging in predicting quantile-based performance measures. With regard to estimation methods, we focus on sectioning, sectioning-batching, and jackknifing as they provide simple ways of obtaining point and interval estimates using the same sample of i.i.d. outputs as used in the default approach. Nakayama (2012) and Asmussen and Glynn (2007) are good references for sectioning and sectioning-batching. Parr and Schucany (1982) and Shao and Wu (1989) are important for understanding the jackknifing methods employed in this paper. Although quantile estimation is used as the primary example for ease of exposition, the proposed approach in this paper applies to other quantile-based performance measures such as conditional value-at-risk (CVaR) or tail conditional expectation (TCE). Such a statistical measure has become particularly popular and useful in the field of quantitative risk management in recent years. See, for instance, Acerbi and Tasche (2002) and Gordy and Juneja (2010) and references therein. However, applications of metamodeling to risk management have been studied only recently (Chen et al. 2012; Liu and Staum 2010) and the simulation literature on the topic is still scarce.

The remainder of this paper is organized as follows. In Section 2, we review standard stochastic kriging and generalize the model formulation. Estimation methods including sectioning, sectioning-batching and jackknifing that we seek to incorporate into the stochastic kriging framework are examined in Section 3. Two numerical examples, on a quadratic loss of a portfolio and a simple stochastic activity network, respectively, with detailed analysis of the results when different estimation methods applied are given in Section 4. Section 5 concludes the paper.

2 GENERALIZED STOCHASTIC KRIGING

Standard stochastic kriging models the simulation response estimate obtained at a design point $\mathbf{x} \in \Omega \subset \mathbb{R}^d$ on the j th simulation replication as

$$\mathcal{Y}_j(\mathbf{x}) = Y(\mathbf{x}) + \varepsilon_j(\mathbf{x}) = \mathbf{f}(\mathbf{x})^\top \boldsymbol{\beta} + M(\mathbf{x}) + \varepsilon_j(\mathbf{x}) , \quad (1)$$

where $Y(\mathbf{x})$ represents the unknown true response that we intend to estimate at point $\mathbf{x}_0 \in \Omega$, and the term $\varepsilon_j(\mathbf{x})$ represents the mean zero simulation error realized on the j th replication. The simulation errors $\varepsilon_1(\mathbf{x}), \varepsilon_2(\mathbf{x}), \dots$ are assumed to be independent and identically distributed across replications at a given design point. Notice that the variance of $\varepsilon_j(\mathbf{x})$ may depend on \mathbf{x} . The terms $\mathbf{f}(\cdot)$ and $\boldsymbol{\beta}$ are, respectively, a $p \times 1$ vector of known functions of \mathbf{x} and a $p \times 1$ vector of unknown parameters. The term $M(\cdot)$ represents a mean zero stationary Gaussian random field such that $E \left[|M(\mathbf{x})|^2 \right] < \infty$ for all $\mathbf{x} \in \Omega$. One can think of $M(\mathbf{x})$ as being sampled from a space of mappings $\mathbb{R}^d \rightarrow \mathbb{R}$, in which functions are assumed to exhibit spatial correlation. Ankenman et al. (2010) refer to the stochastic nature of M as *extrinsic uncertainty*, in contrast to the *intrinsic uncertainty* represented by $\varepsilon_j(\mathbf{x})$ that is inherent in a stochastic simulation output. Specifically, the spatial covariance function between two points in the random field is typically modeled as

$$\text{Cov}(M(\mathbf{x}), M(\mathbf{y})) = \tau^2 \mathcal{R}(\mathbf{x}, \mathbf{y}; \boldsymbol{\theta}) , \quad (2)$$

where τ^2 denotes the spatial variance of the random process and $\mathcal{R}(\cdot, \cdot; \boldsymbol{\theta})$ is the spatial correlation function. The function $\mathcal{R}(\mathbf{x}, \mathbf{y}; \boldsymbol{\theta})$ depends on \mathbf{x} and \mathbf{y} only through their difference; and the parameter vector

$\theta = (\theta_1, \theta_2, \dots, \theta_d)^\top$ controls how quickly the spatial correlation between the two points diminishes as they become farther apart in each direction.

An experimental design for stochastic kriging consists of $\{(\mathbf{x}_i, n_i)_{i=1}^k\}$, a set of design points from the design space Ω to conduct simulation experiments and the corresponding number of replications to apply (or, the number of simulation response estimates to obtain) at each design point. Denote the $k \times 1$ vector of the sample averages of simulation responses by $\bar{\mathcal{Y}} = (\bar{\mathcal{Y}}(\mathbf{x}_1), \bar{\mathcal{Y}}(\mathbf{x}_2), \dots, \bar{\mathcal{Y}}(\mathbf{x}_k))^\top$, and

$$\bar{\mathcal{Y}}(\mathbf{x}_i) = \frac{1}{n_i} \sum_{j=1}^{n_i} \mathcal{Y}_j(\mathbf{x}_i) = Y(\mathbf{x}_i) + \bar{\varepsilon}(\mathbf{x}_i), \quad i = 1, 2, \dots, k, \quad (3)$$

in which $\bar{\varepsilon}(\mathbf{x}_i) = n_i^{-1} \sum_{j=1}^{n_i} \varepsilon_j(\mathbf{x}_i)$. Standard stochastic kriging builds a linear predictor of the form $\lambda_0 + \lambda^\top \bar{\mathcal{Y}}$ to predict the true response $Y(\mathbf{x}_0)$ at any given point \mathbf{x}_0 , such that the location dependent weights λ_0 and λ are chosen to minimize the resulting MSE. Appendix EC.1 of Ankenman et al. (2010) shows that the MSE-optimal predictor of $Y(\mathbf{x}_0)$ is given by

$$\hat{Y}(\mathbf{x}_0) = \mathbf{f}(\mathbf{x}_0)^\top \beta + \Sigma_M(\mathbf{x}_0, \cdot)^\top \Sigma^{-1} (\bar{\mathcal{Y}} - \mathbf{F}\beta), \quad (4)$$

and its corresponding mean square error follows as

$$\text{MSE}(\hat{Y}(\mathbf{x}_0)) = \Sigma_M(\mathbf{x}_0, \mathbf{x}_0) - \Sigma_M(\mathbf{x}_0, \cdot)^\top \Sigma^{-1} \Sigma_M(\mathbf{x}_0, \cdot),$$

where $\Sigma = \Sigma_M + \Sigma_\varepsilon$, and $\mathbf{F} = (\mathbf{f}(\mathbf{x}_1)^\top, \mathbf{f}(\mathbf{x}_2)^\top, \dots, \mathbf{f}(\mathbf{x}_k)^\top)^\top$.

We now elaborate on the terms $\Sigma_M, \Sigma_M(\mathbf{x}_0, \cdot)$ and Σ_ε . The pairwise spatial covariances across the design points are recorded in the $k \times k$ matrix Σ_M , whose (i, h) th element is given by $\Sigma_M(\mathbf{x}_i, \mathbf{x}_h) = \text{Cov}(M(\mathbf{x}_i), M(\mathbf{x}_h))$ as specified in (2). The $k \times 1$ vector $\Sigma_M(\mathbf{x}_0, \cdot)$ contains the spatial covariances between the design points and the prediction point \mathbf{x}_0 . Lastly, the $k \times k$ matrix Σ_ε represents the variance-covariance matrix of the vector of the averaged simulation errors, $\bar{\varepsilon} = (\bar{\varepsilon}(\mathbf{x}_1), \bar{\varepsilon}(\mathbf{x}_2), \dots, \bar{\varepsilon}(\mathbf{x}_k))^\top$. As Chen, Ankenman, and Nelson (2012) show that the use of common random numbers (CRN) does not necessarily help improve the performance of the stochastic kriging predictor, in this paper we assume that CRN is not applied in simulation experiments. In addition, we assume that a common number of simulation replications n is used at all k design points, in which case Σ_ε is reduced to a $k \times k$ diagonal matrix $n^{-1} \text{diag}\{\sigma_{10}^2, \sigma_{20}^2, \dots, \sigma_{k0}^2\}$ with $\sigma_{i0}^2 := \text{Var}(\varepsilon_j(\mathbf{x}_i))$. In implementing stochastic kriging, one has to estimate the unknown model parameters first. Under the assumption that $(Y(\mathbf{x}_0), \bar{\mathcal{Y}}^\top)^\top$ follows a multivariate normal distribution, the standard practice is to obtain estimates of the parameters τ^2, θ and β through maximizing the resulting log-likelihood function. The first step of the estimation procedure is to replace Σ_ε with its estimate $\hat{\Sigma}_\varepsilon$ whose i th diagonal element is specified by $(n-1)^{-1} \sum_{j=1}^n (\mathcal{Y}_j(\mathbf{x}_i) - \bar{\mathcal{Y}}(\mathbf{x}_i))^2, i = 1, 2, \dots, k$. See Ankenman et al. (2010), Chen and Kim (2013) and references therein for details.

We notice that constructing a stochastic kriging predictor given in (4) requires two building blocks, namely, point estimates of the desired response measures at all k design points and the corresponding variance estimates. In the case of standard stochastic kriging, they are simply given by the vector $\bar{\mathcal{Y}}$ and the diagonal entries of $\hat{\Sigma}_\varepsilon$. Now let us consider alternative ways to create the two building blocks so that better stochastic kriging predictors of quantile-based performance measures can possibly be achieved.

Suppose that to obtain a single simulation response $\mathcal{Y}_j(\mathbf{x}_i)$ is required a sample of n_s i.i.d. *basic* simulation outputs at design point \mathbf{x}_i . Mathematically speaking,

$$\mathcal{Y}_j(\mathbf{x}_i) = \Phi(\mathbf{L}^{(j)}(\mathbf{x}_i)), \quad \mathbf{L}^{(j)}(\mathbf{x}_i) := \left\{ \mathcal{L}_h^{(j)}(\mathbf{x}_i) \right\}_{h=1}^{n_s}, \quad j = 1, 2, \dots, n; \quad i = 1, 2, \dots, k.$$

For example, when estimating the p th quantile via simulations, Φ is the generalized inverse of the empirical distribution constructed from $\mathbf{L}^{(j)}(\mathbf{x}_i)$ at the p level. We then realize that the vector of the sample average

responses $\tilde{\mathcal{Y}}$ given in (3) can be regarded as derived from a single long sequence of i.i.d. basic simulation outputs $\{\mathcal{L}_q(\mathbf{x}_i)\}_{q=1}^N$ with $\mathcal{L}_{(j-1)n_s+h} = \mathcal{L}_h^{(j)}(\mathbf{x}_i)$ for $1 \leq j \leq n$ and $1 \leq h \leq n_s$ (assuming $n \cdot n_s = N$). If multiple estimation methods are available, it is possible for the analyst to use a different point estimate $\tilde{\mathcal{Y}}(\mathbf{x}_i)$ other than $\tilde{\mathcal{Y}}(\mathbf{x}_i)$ based on $\{\mathcal{L}_q(\mathbf{x}_i)\}_{q=1}^N$ as long as a variance estimate of the proposed point estimate, say $\tilde{\sigma}^2(\mathbf{x}_i)$, is available as well. Therefore, we are motivated to generalize the model for the simulation response obtained at design point \mathbf{x}_i to

$$\tilde{\mathcal{Y}}(\mathbf{x}_i) = \mathbf{f}(\mathbf{x}_i)^\top \boldsymbol{\beta} + M(\mathbf{x}_i) + \tilde{\epsilon}(\mathbf{x}_i), \quad i = 1, 2, \dots, k, \quad (5)$$

where $\tilde{\epsilon}(\mathbf{x}_i)$ represents the simulation error associated with the response point estimate at \mathbf{x}_i having mean zero and variance $\tilde{\sigma}^2(\mathbf{x}_i)$. Given that simulation runs are conducted without using CRN across the design points, the intrinsic variance-covariance matrix $\tilde{\Sigma}_\epsilon$ under this generalized model is expected to maintain the diagonal form. Moreover, we recognize that (3) is just one instance of the generalized model (5). Clearly, as long as a method for processing *basic* simulation outputs is available to set up the vector of simulation response estimates $\tilde{\mathcal{Y}} := (\tilde{\mathcal{Y}}(\mathbf{x}_1), \tilde{\mathcal{Y}}(\mathbf{x}_2), \dots, \tilde{\mathcal{Y}}(\mathbf{x}_k))^\top$ and the corresponding estimated variance-covariance matrix $\tilde{\Sigma}_\epsilon = \text{diag}(\tilde{\sigma}^2(\mathbf{x}_1), \tilde{\sigma}^2(\mathbf{x}_2), \dots, \tilde{\sigma}^2(\mathbf{x}_k))$, the metamodel parameter estimation and subsequent prediction can be accomplished in a similar fashion as presented for standard stochastic kriging.

3 METHODS FOR POINT AND VARIANCE ESTIMATION

The generalized modeling formulation suggested by (5) in Section 2 opens up new avenues for exploiting existing methods to obtain point and interval estimates to build potentially better stochastic kriging predictors of quantile-based performance measures. In this section, we review some of the well-known techniques through investigating the problem of estimating quantiles of an unknown distribution of a random variable L whose realizations can be obtained via simulations. For simplicity of exposition, we assume that the underlying distribution function $F(\cdot)$ is continuous and strictly increasing.

Given a sample of i.i.d. simulation outputs $\mathbf{L} := \{\mathcal{L}_q\}_{q=1}^N$, the classical point estimate of the p th quantile based on the entire sequence \mathbf{L} as

$$\hat{v}_p^{\text{all}} = \mathcal{L}_{\lceil pN \rceil : N}, \quad (6)$$

where $\mathcal{L}_{1:N} \leq \mathcal{L}_{2:N} \leq \dots \leq \mathcal{L}_{N:N}$ represent the ordered sequence of elements in \mathbf{L} . If we define an operator Φ that maps a sample of size m to its $\lceil pm \rceil$ th order statistic, then (6) can be rewritten as $\hat{v}_p^{\text{all}} = \Phi(\mathbf{L})$. Although (6) provides a consistent quantile estimate, it is not immediate how to give a variance estimate for \hat{v}_p^{all} based on the single sequence \mathbf{L} .

Several techniques have been proposed to give point and interval estimates for quantiles in the literature. The method known as *batching* suggests to divide the sample \mathbf{L} into n non-overlapping batches each of size n_s (assuming $N = n \cdot n_s$), and construct the p th quantile point and interval estimates based on the n quantile estimates obtained from the n batches. Denoting the j th batch of simulation outputs by $\mathbf{L}^{(j)} := \{\mathcal{L}_{(j-1)n_s+h}\}_{h=1}^{n_s}$ and the resulting n quantile estimates by $\Phi(\mathbf{L}^{(j)})$, $j = 1, 2, \dots, n$, we have the quantile estimate and its associated variance estimate due to batching follow as

$$\hat{v}_p^{\text{batch}} = \frac{1}{n} \sum_{j=1}^n \Phi(\mathbf{L}^{(j)}), \quad \hat{\sigma}_{\text{batch}}^2 = \frac{1}{n(n-1)} \sum_{j=1}^n \left(\Phi(\mathbf{L}^{(j)}) - \hat{v}_p^{\text{batch}} \right)^2. \quad (7)$$

It is known that $\Phi(\mathbf{L}^{(j)})$ is asymptotically normal as long as the batch size n_s is sufficiently large, and $(\hat{v}_p^{\text{batch}} - v_p) / \sqrt{\hat{\sigma}_{\text{batch}}^2}$ approximately follows a t -distribution with $n-1$ degrees of freedom. This leads to an asymptotically valid confidence interval for v_p . We realize that batching is the (default) approach adopted by standard stochastic kriging in building metamodels for performance measures including quantiles. Given that the bias of a point estimator of v_p is of order $O(m^{-1})$ if the sample size m used to construct this

estimator is sufficiently large, we expect that the bias of the point estimator $\widehat{v}_p^{\text{batch}}$ due to batching is of order $O(n_s^{-1})$ whereas for $\widehat{v}_p^{\text{all}}$ its bias has an order of $O(N^{-1})$. Furthermore, the bias problem of $\widehat{v}_p^{\text{batch}}$ becomes more salient as the number of batches n increases given a fixed total sample size N .

Due to the bias consideration, we study the second estimation technique known as *sectioning*. See, for example, Asmussen and Glynn (2007) and Nakayama (2012). It is quite similar to batching in that it provides a variance estimate of the point estimate utilizing the n batches (or equivalently, sections). The main difference is that sectioning replaces the point estimate $\widehat{v}_p^{\text{batch}}$ with the quantile estimate $\widehat{v}_p^{\text{all}}$ given in (6). Specifically, the quantile estimate and its corresponding variance estimate due to sectioning are given by

$$\widehat{v}_p^{\text{sect}} = \Phi(\mathbf{L}), \quad \widehat{\sigma}_{\text{sect}}^2 = \frac{1}{n(n-1)} \sum_{j=1}^n \left(\Phi(\mathbf{L}^{(j)}) - \widehat{v}_p^{\text{sect}} \right)^2. \quad (8)$$

An asymptotically valid confidence interval due to sectioning can be established in a similar fashion as for batching. The third approach we consider is a variation of sectioning which is referred to as *sectioning-batching* in Nakayama (2012). This approach uses $\widehat{v}_p^{\text{sect}}$ as the p th quantile point estimate and adopts $\widehat{\sigma}_{\text{batch}}^2$ as the corresponding variance estimate.

The last two techniques of interest relate to *jackknifing*, which is known as a bias-reduction and variance estimation technique for more than half a century. Recently, jackknifing has been applied to risk management as well (Gordy and Juneja 2010). We consider two versions of jackknifing based on sections. One is the *jackknifing bias-corrected estimation* method and the other is the *jackknifing variance estimation* method. Given a sample of size N , say, i.i.d. outputs \mathbf{L} in our problem context, the basic idea is to generate n data sets of size $n_s(n-1)$ by systematically eliminating each section one at a time, and construct point and variance estimates from those data sets. Specifically, the estimates due to the jackknifing bias-corrected method (e.g., Chapter 1 of Efron 1982, Efron and Stein 1981) follow as

$$\widehat{v}_p^{\text{jack-b}} = \frac{1}{n} \sum_{j=1}^n \Psi^{(j)}, \quad \widehat{\sigma}_{\text{jack-b}}^2 = \frac{1}{n(n-1)} \sum_{j=1}^n \left(\Psi^{(j)} - \widehat{v}_p^{\text{jack-b}} \right)^2 \quad (9)$$

where $\Psi^{(j)} := n\Phi(\mathbf{L}) - (n-1)\Phi(\widetilde{\mathbf{L}}^{(j)})$ are sometimes called *pseudovalues* and $\widetilde{\mathbf{L}}^{(j)} := \mathbf{L} \setminus \mathbf{L}^{(j)}$. It can be shown that if the bias order is correctly specified, then the bias of $\widehat{v}_p^{\text{jack-b}}$ can be effectively reduced to less than $O(N^{-1})$. Nevertheless, attempting jackknifing bias reduction is not guaranteed to help; we will learn more about this from the examples in Section 4. On the other hand, the jackknifing variance estimation method (Efron and Stein 1981) adopts $\widehat{v}_p^{\text{jack}} = \Phi(\mathbf{L})$ as its point estimate and approximates its variance by

$$\widehat{\sigma}_{\text{jack}}^2 = \frac{n-1}{n} \sum_{j=1}^n \left(\Phi(\widetilde{\mathbf{L}}^{(j)}) - \widehat{v}_p^{\text{jack}} \right)^2. \quad (10)$$

In fact, one can show that $\widehat{\sigma}_{\text{jack-b}}^2$ given in (9) is asymptotically equivalent to $\widehat{\sigma}_{\text{jack}}^2$ given in (10), since

$$\widehat{\sigma}_{\text{jack-b}}^2 = \frac{1}{n(n-1)} \sum_{j=1}^n \left(\Psi^{(j)} - \widehat{v}_p^{\text{jack-b}} \right)^2 = \frac{n-1}{n} \sum_{j=1}^n \left(\Phi(\widetilde{\mathbf{L}}^{(j)}) - \frac{1}{n} \sum_{j=1}^n \Phi(\widetilde{\mathbf{L}}^{(j)}) \right)^2,$$

provided that sufficient conditions such as those specified in Proposition 1 and Corollary 3 in Shao and Wu (1989) hold.

In summary, the aforementioned approaches give both point estimates and their variance estimates; in particular, the point estimates \widehat{v}_p^\bullet provided above are approximately normal as long as n and n_s are suitably large. It follows that the two building blocks required to apply (5) are readily available. Lastly, we notice that the methods in this section can be applied to operators other than $\Phi(\cdot)$ for sample quantiles, as will be demonstrated by the examples in Section 4.

Before ending this section, we make some comments on the impact of the possible bias induced in finite-sample estimates $\mathcal{Y}(\mathbf{x}_i)$. Standard stochastic kriging assumes that the simulated responses obtained at the design points are unbiased and so does the generalized formulation given in (5). However, like quantiles, simulated point estimates can be biased and this might affect the predictive performance of the stochastic kriging predictor. Chen and Kim (2013) consider the impact of biased estimates by modifying (1) so that a bias term is explicitly incorporated. They show that the number of batches n cannot be too large in order to achieve a good predictive performance if batching is used. Through a parallel analysis we expect to obtain some results regarding performance differences of stochastic kriging when other methods are implemented.

4 NUMERICAL EXAMPLES

4.1 Quadratic Loss (QL) of Two Assets and Its Risk Measurements

In this subsection, we consider a two-dimensional problem borrowed from Hong and Liu (2009). Our objective is to construct the response surfaces of two popular risk measures, namely, VaR and CVaR. Here, we assume that the random loss $L(\boldsymbol{\mu})$ is given by $L(\boldsymbol{\mu}) = a_0 + a^\top \Delta S + \Delta S^\top H \Delta S$ where $\boldsymbol{\mu} = (\mu_1, \mu_2)^\top$, and $a_0 = 0.3$, $a = (0.8, 1.5)^\top$, $H = \begin{pmatrix} 1.2 & 0.6 \\ 0.6 & 1.5 \end{pmatrix}$. Further, we assume that the risk factor ΔS is a bivariate normal random vector with mean $\boldsymbol{\mu}$ and variance-covariance matrix $\Sigma_s = 0.02 \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$. Writing VaR and CVaR at the level p as $v_p(\boldsymbol{\mu})$ and $c_p(\boldsymbol{\mu})$ respectively, we have

$$c_p(\boldsymbol{\mu}) = \mathbb{E} \left[L(\boldsymbol{\mu}) \mid L(\boldsymbol{\mu}) \geq v_p(\boldsymbol{\mu}) \right] = v_p(\boldsymbol{\mu}) + \frac{1}{1-p} \mathbb{E} [L(\boldsymbol{\mu}) - v_p(\boldsymbol{\mu})]^+.$$

Hong and Liu (2009) suggest using the following consistent estimators: given n_s i.i.d. simulated random losses $\mathcal{L}_h(\boldsymbol{\mu})$ for a given mean vector $\boldsymbol{\mu}$,

$$\widehat{v}_p(\boldsymbol{\mu}) = \mathcal{L}_{\lceil n_s \alpha \rceil : n_s}(\boldsymbol{\mu}), \quad \widehat{c}_p(\boldsymbol{\mu}) = \widehat{v}_p(\boldsymbol{\mu}) + \frac{1}{n_s(1-p)} \sum_{h=1}^{n_s} [\mathcal{L}_h(\boldsymbol{\mu}) - \widehat{v}_p(\boldsymbol{\mu})]^+, \quad (11)$$

where $\mathcal{L}_{k:n_s}(\boldsymbol{\mu})$ is the k th order statistic of $\{\mathcal{L}_h(\boldsymbol{\mu})\}_{h=1}^{n_s}$.

Experiments. The experimental design space is $\Omega_{\boldsymbol{\mu}} = [0.001, 0.1]^2$ from which we select $k = 16$ design points, i.e., choose a set of $\boldsymbol{\mu}$ vectors, $\{\boldsymbol{\mu}_i\}_{i=1}^k$. Specifically, a ‘‘maxmin’’ Latin-hypercube sample of 12 design points from $\Omega_{\boldsymbol{\mu}}$ plus its four corner points are used. At each design point we conduct N simulation runs to obtain a sample of simulated random losses $\{\mathcal{L}_q(\boldsymbol{\mu})\}_{q=1}^N$, based on which we construct n pairs of estimates of $v_p(\boldsymbol{\mu})$ and $c_p(\boldsymbol{\mu})$ using the estimation methods discussed in Section 3, namely, batching (**Batch**), sectioning (**Section**), sectioning-batching (**SB**), and the two jackknifing methods. Notice that at a given design point $\boldsymbol{\mu}$ the j th pair of estimates of $v_p(\boldsymbol{\mu})$ and $c_p(\boldsymbol{\mu})$ is calculated based on $\mathbf{L}^{(j)} := \{\mathcal{L}_{(j-1)n_s+h}\}_{h=1}^{n_s}$ using (11), $j = 1, 2, \dots, n$. One decision to be made by the experimenter is how to utilize the entire collection of N simulated random losses at each design point, that is, how to allocate N between the number of pairs of estimates to obtain (i.e., n) and the sample size used to obtain one such pair (i.e., n_s). Different allocation rules of N are expected to lead to different predictive performances of standard stochastic kriging when each of the estimation methods is applied; this is investigated through varying the number of pair of estimates n in $\{5, 25, 100, 200, 400\}$ for a fixed N .

A total number of 1601 check-points are chosen for evaluation. They are regularly spaced in $\Omega_{\boldsymbol{\mu}}$ and we add the extra testing point from Hong and Liu (2009) for a sanity check. We use the estimated root

mean squared error (ERMSE) over the check-point grid as our predictive performance measure:

$$\text{ERMSE}(\widehat{v}_p) = \sqrt{\frac{1}{1601} \sum_{l=1}^{1601} (v_p(\mu_l) - \widehat{v}_p(\mu_l))^2}, \quad \text{ERMSE}(\widehat{c}_p) = \sqrt{\frac{1}{1601} \sum_{l=1}^{1601} (c_p(\mu_l) - \widehat{c}_p(\mu_l))^2}, \quad (12)$$

where $v_p(\mu_l)$, $c_p(\mu_l)$ are true VaR, CVaR, and $\widehat{v}_p(\mu_l)$, $\widehat{c}_p(\mu_l)$ are their predicted values at μ_l by stochastic kriging. As the closed-form expressions are not available, the true values are approximated by simulations with a sample size $n_s = 10^6$ at each of the check-points. Experiments for $p = 0.99$ are conducted and the results are as follows.

Results. The entire experiment is repeated for 100 macro-replications, and the resulting quartiles of the ERMSEs for estimating $v_{0.99}(\mu)$ and $c_{0.99}(\mu)$ with $N = 10^4$ are summarized in Tables 1 and 2. The results for $N = 5 \times 10^4$ and 10^5 convey similar information, hence for the sake of brevity we will only mention the impact of increasing N without showing details.

From Table 1, we observe that sectioning methods outperform the other ones, followed by the jackknifing variance estimation and then batching. When comparing sectioning-batching with sectioning, the latter seems to perform better when the total number of simulation runs N is 10^4 ; but as N increases, the performances of the two methods quickly become indistinguishable. Batching, on the other hand, leads to much larger ERMSEs relative to both sectioning methods and the jackknifing variance estimation method. Furthermore, the ERMSEs of batching seem to increase as n increases given a fixed sample size N . In fact, the minimum ERMSEs of batching are found to be obtained by using a moderate n value, which is consistent with the conclusion reached by Chen and Kim (2013). Regarding jackknifing, the ERMSEs given in the 6th and 12th columns of Table 1 are due to the jackknifing variance estimation method (**Jack**). This method seems as competitive as sectioning methods, especially in terms of estimating $c_{0.99}$. Furthermore, its performance relative to sectioning methods improves as N becomes larger. Table 2 provides some interesting results regarding the performances of **Jack** and the jackknifing bias-corrected estimation method (**Jack-b**). We see that in terms of estimating $v_{0.99}$ the results due to **Jack-b** deteriorate considerably as n becomes greater than 100. In sharp contrast, when estimating $c_{0.99}$ **Jack-b** performs consistently well given a fixed number of simulation runs N .

In fact, among the five estimation methods considered, **Section**, **SB** and **Jack** provide the same point estimates of v_p and c_p using the N simulated random losses at each design point; the difference lies in the way that the variance estimates are given, and hence disparate performances of the three methods follow. The variance estimates $\widehat{\sigma}_{\text{batch}}^2$ and $\widehat{\sigma}_{\text{sect}}^2$ given by (7) and (8) become fairly close as long as N is sufficiently large. As to the jackknifing variance estimate $\widehat{\sigma}_{\text{jack}}^2$ given in (10), the requirements to achieve good performance differ for different types of point estimators under consideration. Shao and Wu (1989) classify the sample p th quantiles as “nonsmooth” estimators, and they claim that more stringent conditions have to be satisfied by n_s and n as $N \rightarrow \infty$ for **Jack** to achieve consistency and asymptotic unbiasedness. On the other hand, the sample estimate \widehat{c}_p is considered to be “smoother” than sample quantiles, and hence its jackknifing variance estimator is consistent with bounded n_s . The interested reader is referred to Shao and Wu (1989) or Chapter 2 of Shao and Tu (1995) for details. This helps explain why **Jack** is dominated by sectioning methods in estimating v_p , but nevertheless seems equally competent in estimating c_p .

Regarding the poor performances of **Batch** and **Jack-b**, Figures 1 (a) and (b) provide some illuminating explanations. The boxplots in Figures 1 (a) and (b) summarize 1000 *point* estimates of $v_{0.99}$ and $c_{0.99}$ at design point $\mu = (10^{-3}, 10^{-3})^\top$, obtained by batching, sectioning (or equivalently, **SB** and **Jack**) and **Jack-b** with $n = 100$ and $N = 10^4$. Recall that a stochastic kriging metamodel is built upon point estimates at the design points, therefore the efficiency of the point estimates determines the goodness of the resulting stochastic kriging predictor. From Figure 1 (a), we see that the variances of the point estimates due to batching and sectioning are pretty similar, which to some extent justifies the sectioning-batching method. However, the point estimates of $v_{0.99}$ and $c_{0.99}$ due to batching are more biased. Although **Jack-b** yields point estimates that are less biased, the resulting bias reduction is not effective as expected. A closer look

Table 1: QL: Summary of ERMSEs ($\times 10^{-3}$) for estimating $v_{0.99}$ and $c_{0.99}$ with $N = 10^4$.

$v_{0.99}$						$c_{0.99}$					
n	quartiles	Batch	Section	SB	Jack	n	quartiles	Batch	Section	SB	Jack
2	25th	25.9	13.7	14.0	13.7	2	25th	33.2	22.0	21.9	22.0
	50th	31.6	19.1	19.5	19.1		50th	39.2	26.8	26.9	26.8
	75th	38.2	25.0	24.8	25.0		75th	48.1	33.6	33.6	33.6
5	25th	23.8	8.90	9.10	9.80	5	25th	30.1	11.9	11.9	11.9
	50th	28.3	11.8	12.0	13.3		50th	34.4	18.2	18.1	18.4
	75th	34.7	16.5	16.8	17.2		75th	43.3	24.0	24.0	24.3
25	25th	26.9	7.70	7.90	9.70	25	25th	28.4	10.6	10.6	10.1
	50th	29.7	10.0	10.4	11.7		50th	33.6	14.2	14.3	13.8
	75th	33.1	12.5	13.0	14.4		75th	39.5	18.2	19.0	17.8
100	25th	98.4	7.60	8.20	10.1	100	25th	92.2	10.1	10.4	10.2
	50th	101	9.60	10.3	13.4		50th	96.4	13.1	13.7	13.5
	75th	104	13.0	14.4	18.3		75th	100	17.5	18.3	18.2
200	25th	31.8	7.80	7.80	10.8	200	25th	248	10.6	12.3	10.2
	50th	34.9	9.90	9.90	14.0		50th	252	13.4	17.4	13.4
	75th	38.5	13.1	13.1	19.4		75th	255	18.0	23.6	17.4
400	25th	193	8.30	9.20	11.4	400	25th	411	10.9	18.3	10.1
	50th	196	10.4	13.7	15.2		50th	413	13.7	25.5	13.3
	75th	199	14.7	19.5	20.2		75th	416	18.7	32.6	17.2

reveals that this method introduces extra bias for estimating $v_{0.99}$ as n further increases for a fixed N . What's worse, **Jack-b** leads to point estimates of $v_{0.99}$ that are much more variable than those given by batching or sectioning. Shao and Tu (1995) point out two facts that deserve our attention. Firstly, the point estimate given by **Jack-b** is less biased than the estimate obtained without applying bias reduction only when the bias order is correctly specified. Hence the formula given in (9) may not apply to estimating $v_{0.99}$ with the values of N and n specified. Secondly, if the point estimate given by **Jack-b** has a much larger variance which results in a considerably inflated mean squared error, then the performance may be adversely affected, especially in situations where the bias is not a major concern. When estimating $v_{0.99}$ with a sufficiently large N , the impact of bias indeed diminishes much faster than variance. On the other hand, Figure 1 (b) shows that unlike in the case of estimating $v_{0.99}$, **Jack-b** is able to provide point estimates $\hat{c}_{0.99}$ as efficient as those due to sectioning. Therefore its good predictive performance follows.

In summary, we recommend to use sectioning in simulation experiments for building stochastic kriging metamodels for VaR and CVaR.

4.2 A Simple Stochastic Activity Network (SAN)

We consider the following example of a simple stochastic activity network (SAN), much of which was constructed based on the one used in Nakayama (2012), Chu and Nakayama (2012) and Hsu and Nelson (1990). There are 5 activities involved in the completion of the project; let T_i represent the time to finish activity i for $i = 1, 2, \dots, 5$. In this example we assume that the activity times T_i are i.i.d. exponential random variables with rate 1. Independent of the other activity times, T_3 is distributed as an exponential random variable with rate x . The duration of the longest one among 3 possible activity paths determines the project time. Let $L(x)$ represent the observed project time, i.e., $L(x) = \max\{T_1 + T_2, T_1 + T_3(x) + T_5, T_4 + T_5\}$. We are interested in estimating the p th quantile $v_p(x)$, and the conditional expectation of the project duration beyond the p th quantile (or, tail conditional expectation), $c_p(x)$, as functions of the rate parameter x . Given any $x > 0$, the CDF F_x of $L(x)$ can be derived via a sequence of conditioning arguments whose form is

Table 2: QL: Comparisons of ERMSEs ($\times 10^{-3}$) obtained by the two jackknifing estimation methods for estimating $v_{0.99}$ and $c_{0.99}$ with $N = 10^4$.

$v_{0.99}$				$c_{0.99}$			
n	quartiles	Jack	Jack-b	n	quartiles	Jack	Jack-b
2	25th	13.7	14.9	2	25th	22.0	21.7
	50th	19.1	21.0		50th	26.8	27.0
	75th	25.0	27.0		75th	33.6	33.6
5	25th	9.80	12.2	5	25th	11.9	11.9
	50th	13.3	16.3		50th	18.4	18.3
	75th	17.2	22.0		75th	24.3	23.9
25	25th	9.70	22.9	25	25th	10.1	10.2
	50th	11.7	31.0		50th	13.8	13.8
	75th	14.4	42.7		75th	17.8	18.5
100	25th	10.1	79.8	100	25th	10.2	10.2
	50th	13.4	97.4		50th	13.5	13.3
	75th	18.3	116		75th	18.2	18.0
200	25th	10.8	291	200	25th	10.2	10.2
	50th	14.0	359		50th	13.4	13.4
	75th	19.4	439		75th	17.4	17.1
400	25th	11.4	812	400	25th	10.1	10.2
	50th	15.2	1020		50th	13.3	13.2
	75th	20.2	1220		75th	17.2	17.0

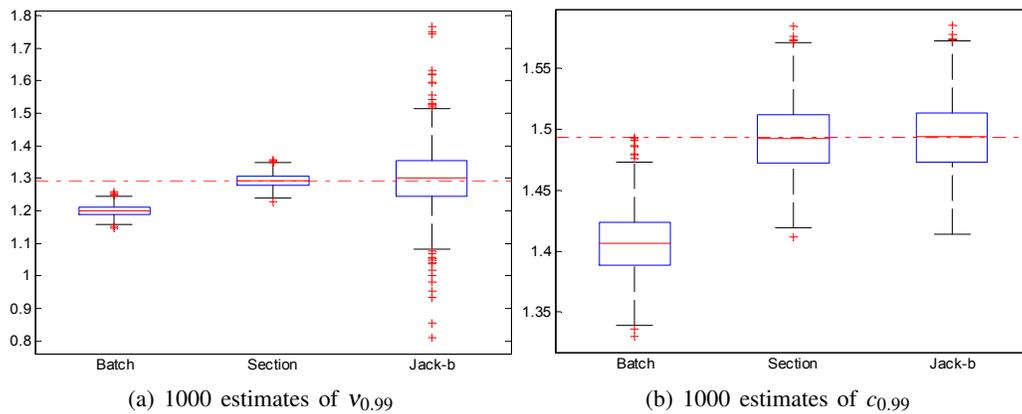


Figure 1: Boxplots of 1000 point estimates of $v_{0.99}$ and $c_{0.99}$ at design point $\mu = (10^{-3}, 10^{-3})^T$ with $n = 100$ and $N = 10^4$. The horizontal dash dotted lines show the respective true values.

given as follows. For $t \geq 0$,

$$F_x(t) = \begin{cases} 1 - (1-x)^{-2}e^{-xt} + ((1-x)^{-2} + (3x-2)(1-x)^{-1}t + 1 + 2x^{-1})e^{-t} \\ -2x^{-1}(1-x)^{-1}e^{-(1+x)t} + (2x(1-x)^{-1} + x^{-2} - tx^{-1} + t^2/2)e^{-2t} - x^{-2}e^{-(2+x)t} & \text{if } x \neq 1, \\ 1 + (3-3t-t^2/2)e^{-t} + (-3-3t+t^2/2)e^{-2t} - e^{-3t}, & \text{if } x = 1. \end{cases} \quad (13)$$

It can be shown that for a fixed $x > 0$, $F_x(t)$ is twice continuously differentiable with respect to t . Denote the derivative of $F_x(t)$ with respect to t by $f_x(t)$, we also know that both $F_x(t)$ and $f_x(t)$ are continuous in x for a fixed $t \geq 0$.

Experiments. The experimental design space for the rate parameter x is $\Omega_x = [1/2, 10/3]$, from which a grid of 7 equally spaced design points is chosen. At each design point N simulation runs are conducted to obtain N simulated project completion times $\{\mathcal{L}_q(x)\}_{q=1}^N$. We construct n pairs of estimates $\{v_p^i(x), c_p^i(x)\}_{i=1}^n$ for $p = 0.99$ based on $\{\mathcal{L}_q(x)\}_{q=1}^N$ using the estimation methods as discussed in Section 3. Given a fixed number of simulation runs N , we conduct simulations with varying n in $\{2, 5, 25, 100, 200, 400\}$ to investigate the impact of different allocation rules on the predictive performance of stochastic kriging with each estimation method applied. A total number of 193 equally spaced check-points are selected from Ω_x , the true $v_{0.99}$ and $c_{0.99}$ at each of which are obtained by numerically inverting the CDF given in (13) and integration afterwards. The ERMSEs over the 193 check-points are adopted as the performance measure, as defined in (12).

Results. The entire experiment is repeated for 100 macro-replications and the quartiles of the resulting ERMSEs obtained with $N = 10^4$ are summarized in Table 3. The results for $N = 5 \times 10^4$ and 10^5 are similar in spirit and hence are omitted. To economize on space, the results for **SB** are also omitted since they are close to those for sectioning and the performances of both methods become very similar as N increases. From Table 3 we observe that **Section** (also **SB**) leads to superior predictive performance followed by **Jack** and then **Batch**. In fact, **Jack** is a little less efficient in terms of estimating the 0.99th quantile $v_{0.99}$ when compared with sectioning methods, but it becomes more competitive in estimating the tail conditional expectation $c_{0.99}$. As to the impact of increasing n given a fixed simulation runs N , batching is found to result in even poorer predictive performances whereas those of **Section** and **Jack** seem to be pretty stable as long as N is sufficiently large and n is not too small. The last two columns of each of two panels in Table 3 provide a comparison of **Jack** and **Jack-b**. As in Section 4.1, we see that the latter leads to very poor performance in predicting $v_{0.99}$, especially when n becomes large. Nevertheless, it seems to perform consistently well for predicting $c_{0.99}$, as long as the total number of simulation runs N is sufficiently large.

We do, however, have to mention that jackknifing methods are typically not as economical as other estimation methods in terms of computational resources, i.e., the computation time and memory space needed to do a similar simulation experiment. Notice that the two jackknifing methods implemented in this paper are based on sections for efficiency improvement. As suggested in Section 4.1, when conducting simulations for building metamodels for quantiles and tail conditional expectations, we recommend to use sectioning rather than batching or jackknifing.

5 CONCLUSIONS

We proposed to generalize standard stochastic kriging to improve its performance in predicting quantile-based performance measures. More specifically, we modified the way in which standard stochastic kriging exploits a sample of i.i.d. simulation outputs from the underlying distribution. This allows us to take advantage of existing estimation methods to obtain point estimates of p th quantile and tail conditional expectation and their corresponding variance estimates. We investigated sectioning, sectioning-batching, and jackknifing methods and compared their performances with batching, which is recognized as the current approach adopted by standard stochastic kriging. Two examples, respectively, demonstrated a quadratic loss of two assets and a simple stochastic activity network. Sectioning methods (sectioning and

Table 3: SAN: Summary of 100 ERMSEs ($\times 10^{-2}$) for estimating $v_{0.99}$ and $c_{0.99}$ with $N = 10^4$.

$v_{0.99}$						$c_{0.99}$					
n	quartiles	Batch	Section	Jack	Jack-b	n	quartiles	Batch	Section	Jack	Jack-b
2	25th	10.1	8.29	8.29	8.19	2	25th	15.5	13.2	13.2	13.0
	50th	13.4	12.2	12.2	12.0		50th	19.3	16.2	16.2	16.1
	75th	16.3	15.2	15.2	16.0		75th	21.9	20.6	20.6	20.6
5	25th	9.77	6.61	7.03	8.23	5	25th	14.0	9.97	9.84	9.89
	50th	12.7	8.99	9.33	10.7		50th	16.7	13.1	13.1	13.2
	75th	14.8	11.2	11.9	14.1		75th	21.1	16.4	16.6	16.6
25	25th	17.2	6.82	6.94	13.1	25	25th	19.7	10.4	10.6	10.2
	50th	20.5	8.82	9.14	18.0		50th	23.1	13.4	13.1	13.2
	75th	25.0	12.5	13.3	26.5		75th	29.0	15.9	15.8	16.0
100	25th	58.4	5.56	6.56	37.0	100	25th	58.0	8.62	8.66	8.46
	50th	61.2	8.03	9.89	48.1		50th	61.5	12.3	12.5	12.1
	75th	64.2	12.1	13.7	65.2		75th	66.5	17.0	16.8	16.7
200	25th	16.3	6.59	8.61	117	200	25th	154	9.57	10.0	9.87
	50th	19.1	8.77	10.5	176		50th	157	13.1	13.2	13.2
	75th	23.1	11.4	12.4	257		75th	162	17.2	15.9	16.5
400	25th	114	7.48	9.22	370	400	25th	255	9.22	9.25	9.07
	50th	116	9.54	12.1	475		50th	257	12.9	12.5	12.6
	75th	119	12.2	16.8	599		75th	260	18.2	16.7	16.6

sectioning-batching) are identified as the most efficient ones to apply at least for building stochastic kriging metamodels for quantile-based measures; while batching is found to work best with a moderate number of batches used. The jackknifing bias-corrected method surprisingly leads to poor prediction results when applied to quantile estimation, which is largely due to the unduly large variance of the resulting point estimate. The proposed metamodeling technique combined with estimation methods such as sectioning is only one example of many other ways that standard stochastic kriging can be extended. Variance reduction techniques and quasi-Monte Carlo methods can be potentially incorporated to make further improvements.

REFERENCES

- Acerbi, C., and D. Tasche. 2002. "On the coherence of expected shortfall". *Journal of Banking & Finance* 26:1487–1503.
- Alexopoulos, C., D. Goldsman, and J. R. Wilson. 2012. "A new perspective on batched quantile estimation". In *Proceedings of the 2012 Winter Simulation Conference*, edited by C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A. M. Uhrmacher. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Ankenman, B. E., B. L. Nelson, and J. Staum. 2010. "Stochastic kriging for simulation metamodeling". *Operations Research* 58:371–382.
- Asmussen, S., and P. W. Glynn. 2007. *Stochastic Simulation*. New York: Springer.
- Avramidis, A. N., and J. R. Wilson. 1998. "Correlation-induction techniques for estimating quantiles in simulation experiments". *Operations Research* 46:574–591.
- Baseball Prospectus 2013. "PECOTA percentiles are here". available at <http://www.baseballprospectus.com/>.
- Chen, E. J., and W. D. Kelton. 1999. "Simulation-based estimation of quantiles". In *Proceedings of the 1999 Winter Simulation Conference*, edited by P. A. Farrington, U. B. Nembhard, D. T. Sturrock, and G. W. Evans, 428–434. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

- Chen, X., B. E. Ankenman, and B. L. Nelson. 2012. "The effects of common random numbers on stochastic kriging metamodels". *ACM Transactions on Modeling and Computer Simulation* 22:7/1–7/20.
- Chen, X., and K. Kim. 2013. "Stochastic kriging with biased sample estimates". Working Paper.
- Chen, X., K. Kim, and B. L. Nelson. 2012. "Stochastic kriging for conditional value-at-risk and its sensitivities". In *Proceedings of the 2012 Winter Simulation Conference*, edited by C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A. M. Uhrmacher. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Chu, F., and M. K. Nakayama. 2012. "Confidence intervals for quantiles when applying variance-reduction techniques". *ACM Transactions on Modeling and Computer Simulation* 22:10/1–10/25.
- Efron, B. 1982. *The Jackknife, the Bootstrap and Other Resampling Plans*. Philadelphia: SIAM.
- Efron, B., and C. Stein. 1981. "The jackknife estimate of variance". *The Annals of Statistics* 9:586–596.
- Gordy, M. B., and S. Juneja. 2010. "Nested simulation in portfolio risk measurement". *Management Science* 56:1833–1848.
- Hong, L. J., and G. Liu. 2009. "Simulation sensitivities of conditional value at risk". *Management Science* 55:281–293.
- Hsu, J. C., and B. L. Nelson. 1990. "Control variates for quantile estimation". *Management Science* 36:835–851.
- JPMorgan, and Reuters. 1996. "RiskMetricsTM - technical document". Fourth Edition, New York.
- Liu, M., and J. Staum. 2010. "Stochastic kriging for efficient nested simulation of expected shortfall". *Journal of Risk* 12:3–27.
- Nakayama, M. K. 2012. "Using sectioning to construct confidence intervals for quantiles when applying importance sampling". In *Proceedings of the 2012 Winter Simulation Conference*, edited by C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A. M. Uhrmacher. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Parr, W. C., and W. R. Schucany. 1982. "Jackknifing L-statistics with smooth weight functions". *Journal of the American Statistical Association* 77:629–638.
- Seila, A. F. 1982. "A batching approach to quantile estimation in regenerative simulations". *Management Science* 28:573–581.
- Shao, J., and D. S. Tu. 1995. *The Jackknife and Bootstrap*. New York: Springer-Verlag.
- Shao, J., and C. F. J. Wu. 1989. "A general theory for jackknife variance estimation". *The Annals of Statistics* 17:1176–1197.

ACKNOWLEDGMENTS

This work is supported by the Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Education, Science, and Technology (No. 2011-0007651).

AUTHOR BIOGRAPHIES

XI CHEN is an Assistant Professor in the Department of Statistical Sciences and Operations Research at Virginia Commonwealth University. She received her Ph.D. in Industrial Engineering and Management Sciences from Northwestern University. Her research interests include stochastic modeling and simulation, applied probability and statistics, computer experiment design and analysis, and simulation optimization. Her email address is xchen@vcu.edu and her web page is <http://www.people.vcu.edu/~xchen4/>.

KYOUNG-KUK KIM is an Assistant Professor in the Department of Industrial and Systems Engineering at KAIST. He received a Ph.D. from Columbia Business School, and B.S., M.S. in mathematics from Seoul National University and Stanford University, respectively. His research interests include stochastic simulation, financial engineering, and risk management. He can be reached at catednoid@kaist.ac.kr.