Proceedings of the 2015 Winter Simulation Conference
L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti, eds.

ESTIMATING A FAILURE PROBABILITY USING A COMBINATION OF VARIANCE-REDUCTION TECHNIQUES

Marvin K. Nakayama

Department of Computer Science
New Jersey Institute of Technology
Newark, NJ 07102, USA

ABSTRACT

Consider a system that is subjected to a random load and having a corresponding random capacity to withstand the load. The system fails when the load exceeds capacity, and we consider efficient simulation methods for estimating the failure probability. Our approaches employ various combinations of stratified sampling, Latin hypercube sampling, and conditional Monte Carlo. We construct asymptotically valid upper confidence bounds for the failure probability for each method considered. We present numerical results to evaluate the proposed techniques on a safety-analysis problem for nuclear power plants, and the simulation experiments show that some of our combined methods can greatly reduce variance.

1 INTRODUCTION

Consider a system that is subjected to a random load L and has a random capacity C to withstand the load. The system fails when the load exceeds capacity, and we define $\theta = P(L \geq C)$ as the failure probability. Our goal is to devise efficient Monte Carlo methods to estimate θ.

This problem framework encompasses many practical applications. For example, civil and mechanical engineers design devices, systems and structures that encounter random loads and have random capacities, with the aim of ensuring only a very small chance of failure (Wunderlich 2005). Catastrophe modeling plays a critical role in the insurance industry to assess the likelihood of infrastructure failures caused by hurricanes, floods, and earthquakes (Grossi and Kunreuther 2005).

Nuclear engineers perform risk assessments and safety analyses of nuclear power plants (NPPs) to determine if their facilities have acceptably low risks during hypothesized accidents. An international effort of the Nuclear Energy Agency Committee on the Safety of Nuclear Installations (2007) recently developed a Safety Margin Action Plan, introducing a framework called *risk-informed safety-margin characterization (RISMC)*. The basic RISMC problem is to ensure that the failure probability θ of an NPP is acceptably small. In this setting, the model complexity leads to extremely long simulation run times, making it crucial to apply variance-reduction techniques (VRTs).

To efficiently estimate θ, we consider applying various combinations of stratified sampling (SS), Latin hypercube sampling (LHS), and conditional Monte Carlo (CMC); see Chapter V of Asmussen and Glynn (2007) and Chapter 4 of Glasserman (2004) for overviews of these and other VRTs. In addition to constructing a point estimator for θ, we also want a confidence interval (CI) or upper confidence bound (UCB) to assess the statistical error in the point estimator. Indeed, current regulations of the Nuclear Regulatory Commission (NRC) require performing safety analyses of NPPs at a 95% confidence level; e.g., see Section 2.b of U.S. Nuclear Regulatory Commission (1995) and Section 24.9 of U.S. Nuclear Regulatory Commission (2011).

Sherry, Gabor, and Hess (2013) and Dube et al. (2014) carry out initial RISMC studies of postulated accidents in NPPs using combined SS+LHS. But these papers do not provide confidence intervals nor do they apply CMC, both of which we do. Our numerical experiments show that further incorporating CMC in a RISMC analysis can greatly reduce variance. Avramidis and Wilson (1996) examine combinations of different VRTs, including LHS, CMC, and control variates, but they do not consider SS, which plays a fundamental role in a RISMC analysis. Moreover, Avramidis and Wilson (1996) do not develop CIs, as needed in NRC licensing analyses, for their integrated methods.

The rest of the paper unfolds as follows. Section 2 develops the mathematical framework of the problem. Sections 3 and 4 review the uses of simple random sampling and stratified sampling to estimate the failure probability θ. We show how to combine SS+LHS in Section 5, further adding in CMC in Section 6. Section 7 presents numerical results of a stylized RISMC analysis, demonstrating the tremendous gains that may be obtained with our approaches. Finally, we provide concluding remarks in Section 8.

2 MATHEMATICAL FRAMEWORK

Let L denote the random load of a system, and let C be its corresponding random capacity. We initially allow L and C to be dependent, but we later will assume in Section 6 that they are independent. Define the joint CDF of (L,C) as H, which we assume is unknown. Let F and G denote the marginal CDFs of L and C, respectively. We assume that we have a simulation model that produces output $(L,C) \sim H$.

The system fails when $L \geq C$, and we want to determine if the *failure probability* $\theta = P(L \geq C)$ is acceptably small, i.e., that $\theta < \theta_0$, for some given constant θ_0. As H is unknown, we do not know the value of θ nor if $\theta < \theta_0$, and we use simulation to try to determine this. Because Monte Carlo simulation produces only noisy estimates of θ, we can not be certain that $\theta < \theta_0$ when our estimate of θ lies below θ_0. Thus, we further require the following:

$$\text{given constants } 0 < \theta_0 < 1 \text{ and } 0 < \beta < 1, \text{ determine with confidence level } \beta \text{ if } \theta < \theta_0. \tag{1}$$

For example, if $\theta_0 = 0.05$ and $\beta = 0.95$, we need to determine with 95% confidence if the failure probability θ is less than 0.05. In certain applications, the values of θ_0 and β are specified by a regulator.

We can analyze (1) through a one-sided hypothesis test at significance level $\alpha = 1 - \beta$. To do this, specify the null hypothesis $\mathscr{H}_0 : \theta \geq \theta_0$ and alternative hypothesis $\mathscr{H}_1 : \theta < \theta_0$. The null hypothesis states that the failure probability is unacceptably high, and rejecting $\mathscr{H}_0$ in favor of the alternative $\mathscr{H}_1$ requires significant evidence supporting $\mathscr{H}_1$. We carry out a hypothesis test by simulating the model n times, and we use the outputs to construct an asymptotic β-level upper confidence bound (UCB) $B(n)$, which has the property that

$$P(\theta < B(n)) \to \beta \quad \text{as} \quad n \to \infty. \tag{2}$$

One typically establishes that a UCB $B(n)$ satisfies (2) by showing that the point estimator of θ obeys a central limit theorem (CLT). With such a UCB, we then define the *decision rule*:

$$\text{at confidence level } \beta, \text{ conclude that } \theta < \theta_0 \text{ if and only if the } \beta\text{-level UCB } B(n) < \theta_0. \tag{3}$$

This then asymptotically satisfies (1). The particular form of the UCB depends on the simulation method applied, and we will examine several approaches, imposing additional assumptions as needed.

3 SIMPLE RANDOM SAMPLING

We first review how to apply *simple random sampling (SRS)*, which is also known as *naive simulation* or *crude Monte Carlo*, to construct a UCB satisfying (2). For SRS, we run the simulation model n independent and identically distributed (i.i.d.) times to obtain a sample (L_i, C_i), $i = 1, 2, \ldots, n$, of n i.i.d. replicates of $(L,C) \sim H$. We then compute the SRS (point) estimator

$$\hat{\theta}_{\mathrm{SRS}}(n) = \frac{1}{n} \sum_{i=1}^{n} I(L_i \geq C_i)$$

of the failure probability θ, where $I(\cdot)$ denotes the indicator function, which takes value 1 (resp., 0) when its argument is true (resp., false). The $I(L_i \geq C_i)$, $i = 1, 2, \ldots, n$, are i.i.d. with finite variance $\sigma^2_{\mathrm{SRS}} = \theta(1-\theta)$, which we can consistently estimate by $\hat{\sigma}^2_{\mathrm{SRS}}(n) = \hat{\theta}_{\mathrm{SRS}}(n)[1 - \hat{\theta}_{\mathrm{SRS}}(n)]$. We then have the following CLT:

$$\frac{\sqrt{n}}{\hat{\sigma}_{\mathrm{SRS}}(n)}(\hat{\theta}_{\mathrm{SRS}}(n) - \theta) \Rightarrow N(0,1)$$

as $n \to \infty$, where $\Rightarrow$ denotes convergence in distribution (Billingsley 1995, Chapter 5) and $N(a,b^2)$ is a normal random variable with mean a and variance b^2.

Now let $z_\beta = \Phi^{-1}(\beta)$ be the β-level critical point of a $N(0,1)$ random variable, where Φ is its CDF, and $Q^{-1}(p) = \inf\{x : Q(x) \geq p\}$ for a CDF Q and $0 < p < 1$. Unfolding the above CLT then leads to the SRS β-level UCB for θ as

$$B_{\mathrm{SRS}}(n) = \hat{\theta}_{\mathrm{SRS}}(n) + z_\beta \hat{\sigma}_{\mathrm{SRS}}(n)/\sqrt{n},$$

which satisfies (2), resulting in the decision rule (3) to (asymptotically) fulfill (1).

4 STRATIFIED SAMPLING

We next review how to apply stratified sampling to estimate θ; see Section 4.3 of Glasserman (2004) for further details on SS. Suppose that the simulation producing the output (L,C) also generates an auxiliary variable Z. We will use Z as a *stratification variable* as follows. Partition the support R of Z as $R = \cup_{s=1}^{t} R_s$, where $R_s \cap R_{s'} = \emptyset$ for $s \neq s'$. We call R_s the sth *stratum*, and we assume that $\lambda_s \equiv P(Z \in R_s)$ is known for each $s = 1, 2, \ldots, t$. Using the law of total probability, we then express the failure probability as

$$\theta = \sum_{s=1}^{t} P(L \geq C, Z \in R_s) = \sum_{s=1}^{t} \lambda_s P(L \geq C | Z \in R_s) = \sum_{s=1}^{t} \lambda_s \theta'_s, \tag{4}$$

where $\theta'_s = P(L'_s \geq C'_s)$ and (L'_s, C'_s) has the conditional distribution of (L,C) given that $Z \in R_s$.

Now assuming that we can generate observations of (L'_s, C'_s) for each stratum s, the representation of θ in (4) forms the basis of stratified sampling, which estimates each θ'_s via simulation. Specifically, define SS sampling weights $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_s)$, where each $\gamma_s \in [0,1]$ with $\sum_{s=1}^{t} \gamma_s = 1$. For a total sample size n, define the sample size for stratum s as $n_s = \lfloor \gamma_s n \rfloor$. For simplicity, we assume that $\gamma_s n$ is integer-valued, so $n_s = \gamma_s n$ and $\sum_{s=1}^{t} n_s = n$. Let $(L'_{s,i}, C'_{s,i})$, $i = 1, 2, \ldots, n_s$, be a sample of n_s i.i.d. observations of (L'_s, C'_s), and we estimate θ'_s by

$$\hat{\theta}'_{\mathrm{SS},s}(n_s) = \frac{1}{n_s} \sum_{i=1}^{n_s} I(L'_{s,i} \geq C'_{s,i}). \tag{5}$$

Assuming that samples across strata are independent, we then define the SS estimator of θ as

$$\hat{\theta}_{\mathrm{SS}}(n) = \sum_{s=1}^{t} \lambda_s \hat{\theta}'_{\mathrm{SS},s}(n_s).$$

The variance of $\hat{\theta}_{\mathrm{SS}}(n)$ is σ^2_{SS}/n, where $\sigma^2_{\mathrm{SS}} = \sum_{s=1}^{t} \lambda_s^2 \theta'_s(1-\theta'_s)/\gamma_s$.

The SS estimator satisfies the following CLT:

$$\frac{\sqrt{n}}{\hat{\sigma}_{\mathrm{SS}}(n)}(\hat{\theta}_{\mathrm{SS}}(n) - \theta) \Rightarrow N(0,1),$$

as $n \to \infty$, where $\hat{\sigma}^2_{\mathrm{SS}}(n) = \sum_{s=1}^{t} \lambda_s^2 \hat{\theta}'_{\mathrm{SS},s}(n_s)[1 - \hat{\theta}'_{\mathrm{SS},s}(n_s)]/\gamma_s$, which consistently estimates σ^2_{SS}. We thus obtain the SS β-level UCB

$$B_{\mathrm{SS}}(n) = \hat{\theta}_{\mathrm{SS}}(n) + z_\beta \hat{\sigma}_{\mathrm{SS}}(n)/\sqrt{n}, \tag{6}$$

which satisfies (2), resulting in the decision rule (3) to fulfill (1).

5 COMBINED SS AND LATIN HYPERCUBE SAMPLING

We now combine SS with Latin hypercube sampling (SS+LHS), which Sherry, Gabor, and Hess (2013) and Dube et al. (2014) use in their initial RISMC studies. McKay, Conover, and Beckman (1979) originally developed LHS as a way to efficiently extend stratified sampling to higher dimensions, and it reduces variance by producing negatively correlated outputs. Stein (1987) further analyzes the approach and shows that the LHS estimator of a mean has asymptotic variance that is no larger than its SRS counterpart. Owen (1992) proves a CLT for the LHS estimator of a mean of bounded outputs, which Loh (1996) extends to outputs having a finite absolute third moment. LHS is perhaps the most commonly used VRT in certain fields, such as nuclear engineering; e.g., Helton and Davis (2003) survey works on LHS, with an emphasis on nuclear applications, and cites over 300 references.

For SS+LHS, we need to assume additional structure for the problem. Let H'_s denote the joint CDF of (L'_s, C'_s) for stratum s, and we now require the following:

Assumption 1 For each stratum $s = 1, 2, \ldots, t$, there is a vector-valued function $w_s : \Re^{d_s} \to \Re^2$ such that if $U_j,\ j = 1, 2, \ldots, d_s$, are i.i.d. unif$[0,1]$ random variables, then

$$w_s(Q_{s,1}^{-1}(U_1), Q_{s,2}^{-1}(U_2), \ldots, Q_{s,d_s}^{-1}(U_{d_s})) \sim H'_s, \tag{7}$$

where $Q_{s,j},\ j = 1, 2, \ldots, d_s$, are CDFs.

Under Assumption 1, we can generate $(L'_s, C'_s) \sim H'_s$ using (7) as follows. First transform d_s i.i.d. uniforms $U_j,\ j = 1, 2, \ldots, d_s$, into

$$X_{s,j} = Q_{s,j}^{-1}(U_j) \sim Q_{s,j}, \quad j = 1, 2, \ldots, d_s, \tag{8}$$

which are independent but not necessarily identically distributed. Then feed the $X_{s,j}$ into the function w_s, and its output (L'_s, C'_s) has joint distribution H'_s. For example, this is exactly the setting of the RISMC studies in Sherry, Gabor, and Hess (2013) and Dube et al. (2014), where w_s represents a detailed nuclear-specific computer code modeling the progression of a hypothesized accident in a NPP; see Hess et al. (2009) for a survey of widely used codes in the nuclear industry. In general, a code run involves numerically solving a system of differential equations, which can require enormous computational effort, making variance reduction crucial to reduce the number of runs needed to achieve an acceptable precision.

Before describing the approach to apply LHS, we first explain how to generate n_s i.i.d. outputs of (L'_s, C'_s) under Assumption 1 when using SRS for each stratum s. We begin by generating $U_{s,i,j},\ i = 1, 2, \ldots, n_s$, $j = 1, 2, \ldots, d_s$, as $n_s d_s$ i.i.d. unif$[0,1]$ random numbers, which we arrange in an $n_s \times d_s$ grid:

$$\begin{array}{cccc} U_{s,1,1} & U_{s,1,2} & \cdots & U_{s,1,d_s} \\ U_{s,2,1} & U_{s,2,2} & \cdots & U_{s,2,d_s} \\ \vdots & \vdots & \ddots & \vdots \\ U_{s,n_s,1} & U_{s,n_s,2} & \cdots & U_{s,n_s,d_s} \end{array}. \tag{9}$$

Applying the function w_s from (7) to each row leads to n_s pairs of outputs

$$\begin{array}{ccccccc} (L'_{s,1}, C'_{s,1}) & = & w_s(Q_{s,1}^{-1}(U_{s,1,1}), & Q_{s,2}^{-1}(U_{s,1,2}), & \ldots, & Q_{s,d_s}^{-1}(U_{s,1,d_s})), \\ (L'_{s,2}, C'_{s,2}) & = & w_s(Q_{s,1}^{-1}(U_{s,2,1}), & Q_{s,2}^{-1}(U_{s,2,2}), & \ldots, & Q_{s,d_2}^{-1}(U_{s,2,d_s})), \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ (L'_{s,n_s}, C'_{s,n_s}) & = & w_s(Q_{s,1}^{-1}(U_{s,n_s,1}), & Q_{s,2}^{-1}(U_{s,n_s,2}), & \ldots, & Q_{s,d_s}^{-1}(U_{s,n_s,d_s})). \end{array}$$

For each $i = 1, 2, \ldots, n_s$, the independence of the d_s uniforms in row i of (9) ensures that $(L'_{s,i}, C'_{s,i}) \sim H'_s$ by (7). Moreover, the independence of the rows in (9) leads to $(L'_{s,i}, C'_{s,i})$, $i = 1, 2, \ldots, n_s$, being mutually independent. Thus, $(L'_{s,i}, C'_{s,i})$, $i = 1, 2, \ldots, n_s$, are i.i.d. with distribution H'_s. We further assume that (9) for different strata $s = 1, 2, \ldots, t$, are independent.

Now we explain how to apply LHS to obtain a *dependent* sample of n_s outputs of (L'_s, C'_s) for stratum s. For each $j = 1, 2, \ldots, d_s$, let $\pi_{s,j} = (\pi_{s,j}(1), \pi_{s,j}(2), \ldots, \pi_{s,j}(n_s))$ be a random permutation of $(1, 2, \ldots, n_s)$; i.e., $\pi_{s,j}(i)$ is the number to which i is mapped in permutation $\pi_{s,j}$, and the distribution of $\pi_{s,j}$ assigns probability mass $1/(n_s!)$ to each of the $n_s!$ permutations of $(1, 2, \ldots, n_s)$. Assume that $\pi_{s,j}$, $j = 1, 2, \ldots, d_s$, are d_s independent permutations. Then for each $i = 1, 2, \ldots, n_s$, and $j = 1, 2, \ldots, d_s$, define

$$V_{s,i,j} = \frac{\pi_{s,j}(i) - 1 + U_{s,i,j}}{n_s}, \tag{10}$$

which we arrange in an $n_s \times d_s$ grid

$$\begin{array}{cccc} V_{s,1,1} & V_{s,1,2} & \cdots & V_{s,1,d_s} \\ V_{s,2,1} & V_{s,2,2} & \cdots & V_{s,2,d_s} \\ \vdots & \vdots & \ddots & \vdots \\ V_{s,n_s,1} & V_{s,n_s,2} & \cdots & V_{s,n_s,d_s} \end{array}. \tag{11}$$

It is easy to show that each row i of (11) consists of d_s i.i.d. unif$[0,1]$ random variables. Applying the function w_s from (7) to each row leads to n_s pairs of outputs

$$\begin{array}{rclcccc} (L'_{s,1}, C'_{s,1}) & = & w_s(Q_{s,1}^{-1}(V_{s,1,1}), & Q_{s,2}^{-1}(V_{s,1,2}), & \ldots, & Q_{s,d_s}^{-1}(V_{s,1,d_s})), \\ (L'_{s,2}, C'_{s,2}) & = & w_s(Q_{s,1}^{-1}(V_{s,2,1}), & Q_{s,2}^{-1}(V_{s,2,2}), & \ldots, & Q_{s,d_2}^{-1}(V_{s,2,d_s})), \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ (L'_{s,n_s}, C'_{s,n_s}) & = & w_s(Q_{s,1}^{-1}(V_{s,n_s,1}), & Q_{s,2}^{-1}(V_{s,n_s,2}), & \ldots, & Q_{s,d_s}^{-1}(V_{s,n_s,d_s})). \end{array} \tag{12}$$

Because each row i of (11) consists of d_s i.i.d. unif$[0,1]$ random variables, we have that $(L'_{s,i}, C'_{s,i}) \sim H'_s$ by (7) for each $i = 1, 2, \ldots, n_s$. But all entries in column j of (11) share the same permutation $\pi_{s,j}$, making the n_s pairs $(L'_{s,i}, C'_{s,i})$, $i = 1, 2, \ldots, n_s$, *dependent*. We call the n_s pairs an *LHS sample* of size n_s.

We estimate the failure probability θ'_s for stratum s using SS+LHS (abbreviated SL) by

$$\hat{\theta}'_{\mathrm{SL},s}(n_s) = \frac{1}{n_s} \sum_{i=1}^{n_s} I(L'_{s,i} \geq C'_{s,i}), \tag{13}$$

for $(L'_{s,i}, C'_{s,i})$ from (12). We assume the t LHS samples across strata $s = 1, 2, \ldots, t$, are independent. Then we define the SS+LHS estimator of the failure probability θ based on an overall sample size of $n = \sum_{s=1}^{t} n_s$ as

$$\hat{\theta}_{\mathrm{SL}}(n) = \sum_{s=1}^{t} \lambda_s \hat{\theta}'_{\mathrm{SL},s}(n_s).$$

The CLT of Owen (1992) for bounded LHS outputs applies to the indicator functions in (13), so $(\sqrt{n}/\sigma'_{\mathrm{SL},s})(\hat{\theta}'_{\mathrm{SL},s}(n_s) - \theta'_s) \Rightarrow N(0,1)$ as $n_s \to \infty$, where we provide an expression for $\sigma'^2_{\mathrm{SL},s}$ below. Because the t estimators $\hat{\theta}'_{\mathrm{SL},s}(n_s)$, $s = 1, 2, \ldots, t$, are independent, the overall SS+LHS estimator of θ satisfies a CLT,

$$\frac{\sqrt{n}}{\sigma_{\mathrm{SL}}} (\hat{\theta}_{\mathrm{SL}}(n) - \theta) \Rightarrow N(0,1)$$

as $n \to \infty$, where $\sigma^2_{\mathrm{SL}} = \sum_{s=1}^{t} \lambda_s^2 \sigma'^2_{\mathrm{SL},s} / \gamma_s$.

The dependence of $I(L'_{s,i} \geq C'_{s,i})$, $i = 1, 2, \ldots, n_s$, complicates the exact form of $\sigma'^2_{\mathrm{SL},s}$, which we next provide. Define $X_{s,j}$ as in (8), and for $(L'_s, C'_s) = w_s(X_{s,1}, X_{s,2}, \ldots, X_{s,d_s})$, let $v_s(X_{s,1}, X_{s,2}, \ldots, X_{s,d_s}) = I(L'_s \geq C'_s)$, whose mean is θ'_s. Also, let $v_{s,j}(X_{s,j}) = E[v_s(X_{s,1}, X_{s,2}, \ldots, X_{s,d_s})|X_{s,j}] - \theta'_s$ be the jth *main effect* of v_s, which has mean 0. Then Owen (1992) (see also Stein 1987) gives an additive approximation v'_s to v_s as

$$v'_s(X_{s,1}, X_{s,2}, \ldots, X_{s,d_s}) = \theta'_s + \sum_{j=1}^{d_s} v_{s,j}(X_{s,j}). \tag{14}$$

This is the best additive approximation to v_s in the sense that if v''_s is another additive approximation to v_s, then

$$E[(v''_s(X_{s,1}, X_{s,2}, \ldots, X_{s,d_s}) - v_s(X_{s,1}, X_{s,2}, \ldots, X_{s,d_s}))^2] \geq E[(v'_s(X_{s,1}, X_{s,2}, \ldots, X_{s,d_s}) - v_s(X_{s,1}, X_{s,2}, \ldots, X_{s,d_s}))^2].$$

Now define $\varepsilon_s = v_s(X_{s,1}, X_{s,2}, \ldots, X_{s,d_s}) - v'_s(X_{s,1}, X_{s,2}, \ldots, X_{s,d_s})$ as the residual of the additive approximation v'_s. Then the asymptotic variance in the CLT for the SS+LHS estimator $\hat{\theta}'_{\mathrm{SL},s}(n_s)$ for stratum s is $\sigma'^2_{\mathrm{SL},s} = \mathrm{Var}[\varepsilon_s]$. In contrast, for the estimator $\hat{\theta}'_{\mathrm{SS},s}(n_s)$ of θ_s for stratum s using SS-alone (i.e., without LHS) from (5), the CLT asymptotic variance is

$$\theta'_s(1-\theta'_s) = \mathrm{Var}[\varepsilon_s] + \sum_{j=1}^{d_s} \mathrm{Var}[v_{s,j}(X_{s,j})];$$

see Owen (1992) and Stein (1987). Hence, LHS removes the variability of the additive part of $v_s = v'_s + \varepsilon_s$. In general, for an arbitrary output function v_s, LHS can greatly reduce variance when v_s is nearly an additive function of its inputs.

Consistently estimating $\sigma'^2_{\mathrm{SL},s}$ is nontrivial, so instead we consider using *replicated LHS (rLHS)*, as developed by Iman and Conover (1980) and Stein (1987). For each stratum s, rather than generating a single LHS sample of size n_s as in (12), we instead generate $r \geq 2$ *independent* LHS samples, each of size $m_s = n_s/r$, which we assume is integer-valued. (Stein 1987 suggests choosing r so that m_s/d_s is "large," and for our numerical results in Section 7, we set $r = 10$.) Specifically, for $k = 1, 2, \ldots, r$, let $(L^{(k)}_{s,i}, C^{(k)}_{s,i})$, $i = 1, 2, \ldots, m_s$, be an LHS sample of size m_s as in (12) except with m_s replacing n_s in (10) and in each random permutation $\pi_{s,j}$ of $(1, 2, \ldots, m_s)$. Assume that the r LHS samples for $k = 1, 2, \ldots, r$, are mutually independent. For each $k = 1, 2, \ldots, r$, let

$$\hat{\theta}^{(k)}_{\mathrm{SL},s}(m_s) = \frac{1}{m_s} \sum_{i=1}^{m_s} I(L^{(k)}_{s,i} \geq C^{(k)}_{s,i})$$

be an estimator of θ'_s from replicate k, and define an estimator of θ using each replicate k across strata as

$$\hat{\theta}^{(k)}_{\mathrm{SL}}(m) = \sum_{s=1}^{t} \lambda_s \hat{\theta}^{(k)}_{\mathrm{SL},s}(m_s)$$

for $m = \sum_{s=1}^{t} m_s = n/r$. Now $\hat{\theta}^{(k)}_{\mathrm{SL}}(m)$, $k = 1, 2, \ldots, r$, are mutually independent, so we can construct a UCB by computing their sample mean and sample variance. Specifically, define the SS+rLHS (abbreviated SrL) estimator of θ as the sample average

$$\bar{\theta}_{\mathrm{SrL}}(n, r) = \frac{1}{r} \sum_{k=1}^{r} \hat{\theta}^{(k)}_{\mathrm{SL}}(m),$$

where $n = \sum_{s=1}^{t} n_s = \sum_{s=1}^{t} r m_s$, and define the sample variance

$$\hat{\sigma}^2_{\mathrm{SrL}}(n, r) = \frac{1}{r-1} \sum_{k=1}^{r} [\hat{\theta}^{(k)}_{\mathrm{SL}}(m) - \bar{\theta}_{\mathrm{SrL}}(n, r)]^2.$$

Let $\tau_{r-1,\beta}$ be the β-level critical point of a Student t random variable T_{r-1} with $r-1$ degrees of freedom; i.e., $P(T_{r-1} \leq \tau_{r-1,\beta}) = \beta$. Then the SS+rLHS β-level UCB for θ is

$$B_{\mathrm{SrL}}(n,r) = \bar{\theta}_{\mathrm{SrL}}(n,r) + \tau_{r-1,\beta}\,\hat{\sigma}_{\mathrm{SrL}}(n,r)/\sqrt{r}, \tag{15}$$

which satisfies (2) as $n \to \infty$ with $r \geq 2$ fixed. We then use $B_{\mathrm{SrL}}(n,r)$ in the decision rule (3) to fulfill (1).

6 COMBINED SS, LHS, AND CONDITIONAL MONTE CARLO

We now want to combine SS, rLHS and conditional Monte Carlo to further increase statistical efficiency; e.g., see Section V.4 of Asmussen and Glynn (2007) for an overview of CMC. CMC reduces sampling error by analytically integrating out some of the variability. To apply CMC, we assume the following:

Assumption 2 For each stratum $s = 1,2,\ldots,t$, there exist scalar-valued functions $w_{s,1} : \Re^{d_{s,1}} \to \Re$ and $w_{s,2} : \Re^{d_{s,2}} \to \Re$ such that $d_{s,1} + d_{s,2} = d_s$ and the $\Re^2$-valued function w_s in (7) satisfies

$$\begin{aligned} w_s(Q_{s,1}^{-1}(u_1), Q_{s,2}^{-1}(u_2), \ldots, Q_{s,d_s}^{-1}(u_{d_s})) = (&w_{s,1}(Q_{s,1}^{-1}(u_1), Q_{s,2}^{-1}(u_2), \ldots, Q_{d_{s,1}}^{-1}(u_{d_{s,1}})), \\ &w_{s,2}(Q_{d_{s,1}+1}^{-1}(u_{d_{s,1}+1}), Q_{d_{s,1}+2}^{-1}(u_{d_{s,1}+2}), \ldots, Q_{d_{s,1}+d_{s,2}}^{-1}(u_{d_{s,1}+d_{s,2}}))) \end{aligned}$$

for every $(u_1, u_2, \ldots, u_{d_s}) \in [0,1]^{d_s}$.

In other words, we can generate the load L'_s with function $w_{s,1}$ and the capacity C'_s with function $w_{s,2}$:

$$L'_s = w_{s,1}(Q_{s,1}^{-1}(U_1), Q_{s,2}^{-1}(U_2), \ldots, Q_{s,d_{s,1}}^{-1}(U_{d_{s,1}})), \tag{16}$$

$$C'_s = w_{s,2}(Q_{s,d_{s,1}+1}^{-1}(U_{d_{s,1}+1}), Q_{s,d_{s,1}+2}^{-1}(U_{d_{s,1}+2}), \ldots, Q_{s,d_{s,1}+d_{s,2}}^{-1}(U_{d_{s,1}+d_{s,2}})), \tag{17}$$

with $U_1, U_2, \ldots, U_{d_s}$ as i.i.d. unif[0,1] random numbers. Thus, L'_s and C'_s are *independent* because they use disjoint sets of uniforms. In the RISMC application discussed in Section 1, it is reasonable to have independent load and capacity since the load depends on the way in which the hypothesized accident evolves, whereas the variability in manufacturing and material properties determine the capacity. Indeed, the initial RISMC studies of Sherry, Gabor, and Hess (2013) and Dube, Sherry, Gabor, and Hess (2014) assume the load and capacity are independent.

We further require the following:

Assumption 3 For each stratum $s = 1,2,\ldots,t$, the marginal CDF G'_s of C'_s can be computed analytically or numerically.

Because of the independence of L'_s and C'_s under Assumption 2, we have that

$$\theta'_s = E[I(L'_s \geq C'_s)] = E[E[I(L'_s \geq C'_s)|L'_s]] = E[P(C'_s \leq L'_s|L'_s)] = E[G'_s(L'_s)]. \tag{18}$$

This suggests that rather than estimating θ'_s by averaging copies of $I(L'_s \geq C'_s)$, we can instead average copies of $G'_s(L'_s)$, which we can compute by Assumption 3; this is the basic idea of CMC. In more detail, we generate an $m_s \times d_{s,1}$ grid of $V_{s,i,j}$ using LHS as in (11) but with m_s rows and $d_{s,1}$ columns instead of n_s and d_s, respectively. Applying the load function $w_{s,1}$ from (16) to each row of the $V_{s,i,j}$ leads to

$$\begin{array}{rcllll} L'_{s,1} & = & w_{s,1}(Q_{s,1}^{-1}(V_{s,1,1}), & Q_{s,2}^{-1}(V_{s,1,2}), & \ldots, & Q_{s,d_{s,1}}^{-1}(V_{s,1,d_{s,1}})), \\ L'_{s,2} & = & w_{s,1}(Q_{s,1}^{-1}(V_{s,2,1}), & Q_{s,2}^{-1}(V_{s,2,2}), & \ldots, & Q_{s,d_{s,1}}^{-1}(V_{s,2,d_{s,1}})), \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ L'_{s,m_s} & = & w_{s,1}(Q_{s,1}^{-1}(V_{s,m_s,1}), & Q_{s,2}^{-1}(V_{s,m_s,2}), & \ldots, & Q_{s,d_{s,1}}^{-1}(V_{s,m_s,d_{s,1}})). \end{array} \tag{19}$$

Hence, $L'_{s,i}$, $i = 1,2,\ldots,m_s$, is an LHS sample of size m_s of loads, where each $L'_{s,i} \sim F'_s$, the marginal CDF of L'_s. The LHS+CMC estimator $\bar{G}'_{s,m_s} \equiv (1/m_s)\sum_{i=1}^{m_s} G'_s(L'_{s,i})$ of θ'_s is unbiased by (18).

Moreover, a variance decomposition implies

$$\mathrm{Var}[I(L'_s \ge C'_s)] = \mathrm{Var}[E[I(L'_s \ge C'_s)|L'_s]] + E[\mathrm{Var}[I(L'_s \ge C'_s)|L'_s]] \ge \mathrm{Var}[E[I(L'_s \ge C'_s)|L'_s]] = \mathrm{Var}[G'_s(L'_s)]$$

since variance is always nonnegative. Thus, $G'_s(L'_s)$ has no greater variance than $I(L'_s \ge C'_s)$, so we are guaranteed a variance reduction (compared to SS-alone) by applying CMC to estimate θ'_s. Also, Avramidis and Wilson (1996) show that the asymptotic variance of LHS+CMC is less than that for either LHS-alone or CMC-alone. In addition, CMC has the further computational advantage that capacities do not need to be generated. Another benefit of combining CMC with LHS is that an additive approximation for the output $G_s(L'_s)$ can be more accurate than the one in (14) for the indicator output $I(L'_s \ge C'_s)$ for LHS alone. As a consequence, the combination of LHS and CMC may reduce variance significantly more than either by itself, as we will see in our numerical results in Section 7.

To construct a UCB using combined SS, rLHS, and CMC, we replicate (19) $r \ge 2$ times independently, where on each replicate $k = 1, 2, \ldots, r$, we have $L^{(k)}_{s,i}$, $i = 1, 2, \ldots, m_s$, as the LHS sample of size m_s of loads, and the r LHS samples across the r replicates are independent. For each replicate $k = 1, 2, \ldots, r$, let

$$\bar{G}^{(k)}_{s,m_s} = (1/m_s) \sum_{i=1}^{m_s} G'_s(L^{(k)}_{s,i}) \tag{20}$$

for each scenario $s = 1, 2, \ldots, t$, and the SS+LHS+CMC (abbreviated SLC) estimator of θ from replicate k is

$$\hat{\theta}^{(k)}_{\mathrm{SLC}}(m) = \sum_{s=1}^{t} \lambda_s \bar{G}^{(k)}_{s,m_s}.$$

The $\hat{\theta}^{(k)}_{\mathrm{SLC}}(m)$, $k = 1, 2, \ldots, r$, are mutually independent, and we compute their sample average and sample variance as

$$\bar{\theta}_{\mathrm{SrLC}}(n,r) = \frac{1}{r} \sum_{k=1}^{r} \hat{\theta}^{(k)}_{\mathrm{SLC}}(m),$$

$$\hat{\sigma}^2_{\mathrm{SrLC}}(n,r) = \frac{1}{r-1} \sum_{k=1}^{r} [\hat{\theta}^{(k)}_{\mathrm{SLC}}(m) - \bar{\theta}_{\mathrm{SrLC}}(n,r)]^2,$$

where SrLC denotes SS+rLHS+CMC. Then the SS+rLHS+CMC β-level UCB is

$$B_{\mathrm{SrLC}}(n,r) = \bar{\theta}_{\mathrm{SrLC}}(n,r) + \tau_{r-1,\beta}\, \hat{\sigma}_{\mathrm{SrLC}}(n,r)/\sqrt{r}, \tag{21}$$

which satisfies (2) as $n \to \infty$ with $r \ge 2$ fixed. Thus, using $B_{\mathrm{SrLC}}(n,r)$ in the decision rule (3) fulfills (1).

7 NUMERICAL RESULTS

We now present numerical results from running simulation experiments for the setting of (16) and (17) with the following model. We consider a RISMC analysis of a hypothesized station blackout (SBO) in an NPP having the event tree in Figure 1, which is taken from Dube et al. (2014). The load represents the peak cladding temperature (PCT) of the material surrounding the core during the SBO, and the capacity is the temperature at which core damage occurs. In the tree, E_1, E_2, and E_3 denote intermediate events that determine how the SBO progresses; e.g., the lower (resp., upper) branch of E_2 corresponds to a safety relief valve being stuck open (resp., not being stuck open). From engineering analysis and prior experience, the probabilities of the branches of each intermediate event are known, as depicted in the figure. The intermediate events result in $t = 4$ *scenarios*, each corresponding to a path from left to right through the event tree. Let Z be a random variable for the scenario observed, and we use Z as a stratification variable

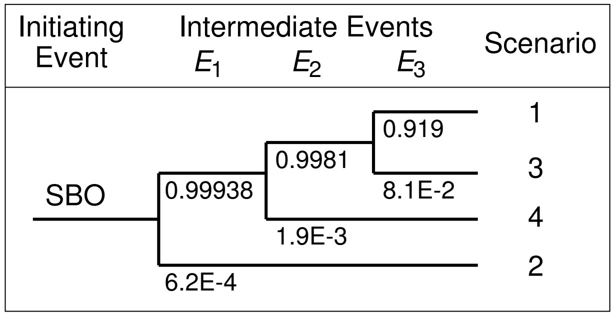

Figure 1: An event tree for a station blackout.

for SS with each scenario as a stratum. We compute the probability λ_s of scenario s by multiplying the branch probabilities of the intermediate events along the path; e.g., $\lambda_4 = 0.99938 \times 0.0019$.

As in Sherry, Gabor, and Hess (2013) and Dube et al. (2014), we will assume that load and capacity are independent in each scenario, as in Assumption 2. Also, while the distribution of the load depends on the scenario observed, the capacity's distribution does not. These are reasonable assumptions since the load depends on how the hypothesized SBO unfolds, whereas the capacity is determined by material properties and manufacturing variability. Sherry, Gabor, and Hess (2013) and Dube et al. (2014) further assume the capacity's distribution is triangular with support $[a,b] = [1800, 2600]$ and mode $c = 2200$, and we work with the same. Thus, the marginal CDF G of the capacity C is

$$G(x) = \begin{cases} (x-a)^2/[(b-a)(c-a)] & \text{for } a \le x \le c, \\ 1-2(b-x)^2/[(b-a)(c-a)] & \text{for } c < x \le b, \end{cases} \tag{22}$$

with $G(x) = 0$ for $x < a$ and $G(x) = 1$ for $x > b$, and we use $G'_s = G$ in computing (20) for CMC.

While Sherry, Gabor, and Hess (2013) and Dube et al. (2014) use detailed nuclear-specific computer codes to generate random loads for each stratum, we instead assume that loads follow lognormal distributions. We chose lognormal distributions because histograms of loads, e.g., in Sherry, Gabor, and Hess (2013), often exhibit heavy right tails. To generate an observation of the lognormal load L'_s for stratum s as in (16), we exponentiated the sum of $d_{s,1}$ independent normals. Specifically, for $s = 1,2,\ldots,t$, and $j = 1,2,\ldots,d_{s,1}$, let $X_{s,j} \sim N(\mu_{s,j}, \sigma^2_{s,j})$ in (8), and set $L'_s = \exp(\sum_{j=1}^{d_{s,1}} X_{s,j})$ for the function $w_{s,1}$ in (16), where $X_{s,j}$, $j = 1,2,\ldots,d_{s,1}$, are independent. Hence, L'_s has a lognormal distribution with parameters $\mu_s = \sum_{j=1}^{d_{s,1}} \mu_{s,j}$ and $\sigma^2_s = \sum_{j=1}^{d_{s,1}} \sigma^2_{s,j}$; i.e., $E[L'_s] = \exp(\mu_s + \sigma^2_s/2)$ and $\mathrm{Var}[L'_s] = \exp(2\mu_s + \sigma^2_s)[\exp(\sigma^2_s) - 1]$. In our experiments, we set $d_{s,1} = 10$ for each stratum $s = 1,2,\ldots,t$. In stratum 1, the $d_{1,1}$ independent normals $X_{1,j}$, $j = 1,2,\ldots,d_{1,1}$, have monotonically increasing means $\mu_{1,j} = 7.5j/[d_{1,1}(d_{1,1}+1)/2]$ and equal variances (EV) $\sigma^2_{1,j} = (0.02)^2/d_{1,1}$. In stratum 2, the $d_{2,1}$ independent normals have equal means (EM) $\mu_{2,j} = 7.6/d_{2,1}$ and monotonically increasing variances $\sigma^2_{2,j} = (0.03)^2 j/[d_{2,1}(d_{2,1}+1)/2]$ for $j = 1,2,\ldots,d_{2,1}$. In stratum 3, the $d_{3,1}$ independent normals have monotonically decreasing means $\mu_{3,j} = 7.7(d_{3,1}-j+1)/[d_{3,1}(d_{3,1}+1)/2]$ and EV $\sigma^2_{3,j} = (0.04)^2/d_{3,1}$ for $j = 1,2,\ldots,d_{3,1}$. In stratum 4, the $d_{4,1}$ independent normals have EM $\mu_{4,j} = 7.8/d_{4,1}$ and monotonically decreasing variances $\sigma^2_{4,j} = (0.05)^2(d_{4,1}-j+1)/[d_{4,1}(d_{4,1}+1)/2]$ for $j = 1,2,\ldots,d_{4,1}$. We thus have $\mu_1 = 7.5$, $\mu_2 = 7.6$, $\mu_3 = 7.7$, $\mu_4 = 7.8$, $\sigma^2_1 = (0.02)^2$, $\sigma^2_2 = (0.03)^2$, $\sigma^2_3 = (0.04)^2$, and $\sigma^2_4 = (0.05)^2$. We then numerically computed the failure probability as $\theta = 0.0465$ using MATLAB (2015). In our requirement (1), we set $\theta_0 = 0.05$ and $\beta = 0.95$, so $\theta < \theta_0$ for our particular parameter values. In our experiments, we suppose that we do not know the value of θ and instead estimate it via simulation, and we apply the decision rule (3) to try to determine with 95% confidence if $\theta < \theta_0$.

Our experiments varied the total sample size $n = 4^q \times 100$ for $q = 1,2,3,4$. We applied the methods SS alone; combined SS and CMC (SS+CMC), which we did not describe in the paper; SS+rLHS; and SS+rLHS+CMC. For each method, we constructed a UCB, which is given by (6), (15), and (21) for

SS-alone, SS+rLHS and SS+rLHS+CMC, respectively. We set the SS sampling weights $\gamma_s = 0.25$ for each stratum $s = 1,2,3,4$. For rLHS, we used $r = 10$ replicated LHS samples within a stratum, so for each stratum s, we generate an LHS sample of size $m_s = \gamma_s n/r$ for each replicate. For SS-alone and SS+CMC, we do not use replications, so $r = 1$.

Table 1 contains the results from running 10^4 independent experiments for each method and total sample size. The column labeled "AHW" gives the average half-width across the 10^4 experiments, where the half-width is the difference between the UCB and the point estimate of θ. For a UCB $B(n)$ and overall sample size n, the *coverage* is the probability $P(\theta < B(n))$. We noted throughout the paper that for each method, its UCB satisfies (2), so the coverage converges to the nominal level $\beta = 0.95$ as $n \to \infty$. But for fixed n, the coverage may differ from β. We estimate the coverage as the fraction of the 10^4 experiments in which $\theta < B(n)$. The *probability of correct decision (PCD)* is estimated as the fraction of the 10^4 experiments that the decision rule in (3) correctly determined that $\theta = 0.0465 < 0.05 = \theta_0$. The column "Sample Var" gives the sample variance of the point estimator of θ across the 10^4 experiments. For each particular method x, the last column presents the *variance-reduction factor* (VRF), which for a given overall sample size n is the ratio of the sample variance for SS-alone over the sample variance for method x.

Table 1: We compare SS, SS+CMC, SS+rLHS, SS+rLHS+CMC for different total sample sizes n in terms of average half-width (AHW), coverage, probability of correct decision (PCD), sample variance, and variance-reduction factor (VRF).

Method	n	r	AHW	Coverage	PCD	Sample Var	VRF SS/x
SS	400	1	$9.44E-03$	0.868	0.336	$4.21E-05$	1.00
	1600	1	$5.08E-03$	0.882	0.387	$1.00E-05$	1.00
	6400	1	$2.60E-03$	0.919	0.701	$2.54E-06$	1.00
	25600	1	$1.31E-03$	0.935	0.995	$6.37E-07$	1.00
SS+CMC	400	1	$2.58E-03$	0.949	0.727	$2.42E-06$	17.39
	1600	1	$1.29E-03$	0.949	0.999	$6.08E-07$	16.47
	6400	1	$6.45E-04$	0.949	1.000	$1.57E-07$	16.15
	25600	1	$3.22E-04$	0.948	1.000	$3.84E-08$	16.59
SS+rLHS	400	10	$7.58E-03$	0.782	0.612	$2.83E-05$	1.49
	1600	10	$4.39E-03$	0.785	0.402	$6.68E-06$	1.50
	6400	10	$2.21E-03$	0.927	0.791	$1.53E-06$	1.66
	25600	10	$1.09E-03$	0.935	0.999	$3.80E-07$	1.68
SS+rLHS+CMC	400	10	$8.56E-04$	0.937	1.000	$2.38E-07$	176.82
	1600	10	$3.42E-04$	0.938	1.000	$3.68E-08$	271.86
	6400	10	$1.63E-04$	0.943	1.000	$8.36E-09$	303.15
	25600	10	$8.08E-05$	0.948	1.000	$2.06E-09$	309.68

We make the following observations about the results in Table 1. Each combined method SS+CMC, SS+rLHS and SS+rLHS+CMC has smaller sample variances and smaller AHWs than SS alone for the same n. The value of the VRF for a combined method corresponds to the approximate factor by which the sample size for SS-alone would need to be increased to obtain an AHW that is comparable to that for the combined method. For example, since the VRF for SS+rLHS is 1.68 for $n = 25600$, using SS-alone would require about a 68% larger sample size to have roughly the same AHW as SS+rLHS. While this is attractive, the VRFs for SS+CMC and SS+rLHS+CMC are much larger, with VRF for SS+CMC around 16 and the VRF for SS+rLHS+CMC over 300 for large n. Thus, SS-alone would require orders of magnitude larger sample sizes than these combined CMC methods to achieve a comparable AHW.

To explain the apparent reason that SS+rLHS+CMC so vastly outperforms SS, SS+CMC and SS+rLHS in our experiments, recall that we previously noted in Section 5 that LHS can greatly reduce variance when the output function v_s is nearly an additive function of its inputs $X_{s,j}$; see the paragraph containing (14). When CMC is not applied, as in (5) and (13), the output function $I(L'_s \geq C'_s)$ is highly non-additive; thus, SS+rLHS does not provide an immense improvement over SS-alone. But employing CMC corresponds to the output function $G(L'_s)$ (see (20) and (22)), so an additive approximation as in (14) seems to be much more accurate, which leads to SS+rLHS+CMC having tremendously smaller variance than SS+CMC and SS+rLHS.

For each method, coverage approaches $\beta = 0.95$ as n grows large, demonstrating the asymptotic validity of our UCBs. Including CMC with other methods leads to improved coverage for each n; thus, the asymptotics for the CLTs based on combining CMC with other approaches seem to require smaller sample sizes to hold.

As n grows large, the PCD for each method approaches 1, as expected. But for a fixed n, all of the combined methods have higher PCD than SS-alone, demonstrating the benefits of the smaller variances with combinations of VRTs.

8 CONCLUDING REMARKS

We presented various methods for estimating a failure probability $\theta = P(L \geq C)$, where L represents a load and C denotes the capacity, both random. The approaches we considered are simple random sampling; stratified sampling; combined SS and replicated Latin hypercube sampling; and combined SS, rLHS and conditional Monte Carlo. For each we showed how to construct an asymptotically valid upper confidence bound for θ.

We also presented numerical results for the RISMC framework for safety analyses of nuclear power plants. Using a stylized version of a problem considered in Dube et al. (2014), we showed that the combined methods can greatly outperform SS-alone, in terms of variance, coverage and probability of correct decision. Sherry, Gabor, and Hess (2013) and Dube et al. (2014) carry out an analysis with SS+LHS using nuclear-specific computer codes, but these papers do not consider the construction of UCBs, as needed for (1) to account for statistical variability. Nor did these studies apply CMC, which our experiments showed can further greatly reduce variance. We plan on testing our combined methods using nuclear-specific computer codes.

Our UCBs using LHS employ replicated LHS with $r \geq 2$ independent LHS samples, which leads to a slight loss in statistical efficiency compared to single-sample LHS (i.e., without replicating). Owen (1992) develops single-sample estimators of the LHS asymptotic variance, and we plan to further adapt this idea for our combined VRTs using LHS.

ACKNOWLEDGEMENTS

This work has been supported in part by the National Science Foundation under Grants No. CMMI-0926949, CMMI-1200065, and DMS-1331010. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation. The author thanks Dr. David Grabaskas for introducing him to the RISMC problem.

REFERENCES

Asmussen, S., and P. Glynn. 2007. *Stochastic Simulation: Algorithms and Analysis*. New York: Springer.

Avramidis, A. N., and J. R. Wilson. 1996. "Integrated Variance Reduction Strategies for Simulation". *Operations Research* 44:327–346.

Billingsley, P. 1995. *Probability and Measure*. Third ed. New York: John Wiley & Sons.

Dube, D. A., R. R. Sherry, J. R. Gabor, and S. M. Hess. 2014. "Application of risk informed safety margin characterization to extended power uprate analysis". *Reliability Engineering & System Safety* 129:19–28.

Glasserman, P. 2004. *Monte Carlo Methods in Financial Engineering*. New York: Springer.

Grossi, P., and H. Kunreuther. (Eds.) 2005. *Catastrophe Modeling: A New Approach to Managing Risk*. New York: Springer.

Helton, J. C., and F. J. Davis. 2003. "Latin Hypercube Sampling and the Propagation of Uncertainty in Analyses of Complex Systems". *Reliability Engineering & System Safety* 81:23–69.

Hess, S., J. Gaertner, J. Gabor, L. Shanley, L. Enos-Sylla, S. Prasad, and S. Hollingsworth. 2009. "Framework for Risk-Informed Safety Margin Characterization". Technical Report 1019206, Electric Power Research Institute, Palo Alto, California.

Iman, R. L., and W. J. Conover. 1980. "Small Sample Sensitivity Analysis Techniques for Computer Models, With an Application to Risk Analysis". *Communications in Statistics, Part A—Theory and Methods* 9:1749–1874. with discussion.

Loh, W.-L. 1996. "On Latin Hypercube Sampling". *Annals of Statistics* 24:2058–2080.

MATLAB 2015. *MATLAB Version R2015a*. Natick, Massachusetts: The MathWorks Inc.

McKay, M. D., W. J. Conover, and R. J. Beckman. 1979. "A Comparison of Three Methods for Selecting Input Variables in the Analysis of Output from a Computer Code". *Technometrics* 21:239–245.

Nuclear Energy Agency Committee on the Safety of Nuclear Installations 2007. "Safety Margins Action Plan (SMAP) – Final Report". Technical Report NEA/CSNI/R(2007)9, Organization for Economic Cooperation and Development.

Owen, A. B. 1992. "A Central Limit Theorem for Latin Hypercube Sampling". *Journal of the Royal Statistical Society B* 54:541–551.

Sherry, R. R., J. R. Gabor, and S. M. Hess. 2013. "Pilot Application of Risk Informed Safety Margin Characterization to a Total Loss of Feedwater Event". *Reliability Engineering and System Safety* 117:65–72.

Stein, M. 1987. "Large Sample Properties of Simulations Using Latin Hypercube Sampling". *Technometrics* 29:143–151. correction 32:367.

U.S. Nuclear Regulatory Commission 1995. "Voltage-Based Repair Criteria For Westinghouse Steam Generator Tubes Affected By Outside Diameter Stress Corrosion Cracking". Nuclear Regulatory Commission Generic Letter NRC GL 95-05, U.S. Nuclear Regulatory Commission, Washington, DC.

U.S. Nuclear Regulatory Commission 2011. "Applying Statistics". U.S. Nuclear Regulatory Commission Report NUREG-1475, Rev 1, U.S. Nuclear Regulatory Commission, Washington, DC.

Wunderlich, W. O. 2005. *Hydraulic Structures: Probabilistic Approaches to Maintenance*. American Society of Civil Engineers.

AUTHOR BIOGRAPHIES

MARVIN K. NAKAYAMA is a professor in the Department of Computer Science at the New Jersey Institute of Technology. He received his Ph.D. in operations research from Stanford University. He won second prize in the 1992 George E. Nicholson Student Paper Competition sponsored by INFORMS and is a recipient of a CAREER Award from the National Science Foundation. A paper he co-authored won the Best Theoretical Paper Award for the 2014 Winter Simulation Conference. He is the Simulation Area Editor for the *INFORMS Journal on Computing* and an associate editor for *ACM Transactions on Modeling and Computer Simulation*. His research interests include simulation, modeling, statistics, risk analysis, and energy. His email address is marvin@njit.edu.