Reliability Engineering and System Safety 165 (2017) 376–394

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

# Efficient Monte Carlo methods for estimating failure probabilities

Andres Alban[a], Hardik A. Darji[a,b], Atsuki Imamura[c], Marvin K. Nakayama[c,*]

[a] Mathematical Sciences Dept., New Jersey Institute of Technology, Newark, NJ 07102, USA
[b] Mechanical Engineering Dept., New Jersey Institute of Technology, Newark, NJ 07102, USA
[c] Computer Science Dept., New Jersey Institute of Technology, Newark, NJ 07102, USA

ARTICLE INFO

*Keywords:*
Probabilistic safety assessment
Risk analysis
Structural reliability
Uncertainty
Monte Carlo
Variance reduction
Confidence intervals
Nuclear regulation
Risk-informed safety-margin characterization

ABSTRACT

We develop efficient Monte Carlo methods for estimating the failure probability of a system. An example of the problem comes from an approach for probabilistic safety assessment of nuclear power plants known as risk-informed safety-margin characterization, but it also arises in other contexts, e.g., structural reliability, catastrophe modeling, and finance. We estimate the failure probability using different combinations of simulation methodologies, including stratified sampling (SS), (replicated) Latin hypercube sampling (LHS), and conditional Monte Carlo (CMC). We prove theorems establishing that the combination SS+LHS (resp., SS+CMC+LHS) has smaller asymptotic variance than SS (resp., SS+LHS). We also devise asymptotically valid (as the overall sample size grows large) upper confidence bounds for the failure probability for the methods considered. The confidence bounds may be employed to perform an asymptotically valid probabilistic safety assessment. We present numerical results demonstrating that the combination SS+CMC+LHS can result in substantial variance reductions compared to stratified sampling alone.

## 1. Introduction

Consider a stochastic model of the behavior of a system (e.g., a structure, such as a building or ship). The system's uncertainties may include random loads and capacities, environmental conditions, material properties, etc. The system fails under specified conditions—e.g., a critical subset of components fails—and the goal is to determine if the failure probability $\theta$ is acceptably small. As the system's complexity renders computing $\theta$ as intractable, we instead apply Monte Carlo simulation to estimate $\theta$. To account for the sampling error of the simulation estimates, a confidence interval for $\theta$ is also needed. Running the simulation model may be expensive, so we want to reduce the sampling error. In this paper, we combine different variance-reduction techniques (VRTs) to estimate $\theta$.

A motivating example comes from a framework for probabilistic safety assessments (PSAs) of nuclear power plants (NPPs) known as *risk-informed safety-margin characterization* (RISMC), which was proposed by an international effort of the Nuclear Energy Agency Committee on the Safety of Nuclear Installations [1]. The purpose of RISMC is to address recent changes in NPPs. For example, many NPPs in the current U.S. fleet are aging past their original 40-year operating licenses, with owners applying for lengthy extensions [2]. Also, for economic reasons, plants are sometimes run at higher output levels, known as power uprates [3]. These and other factors can lead to degradations in safety margins previously deemed acceptable, and RISMC aims to better understand their impacts.

Current NPP PSAs compare a random load to a *fixed* capacity. For example, the U.S. Nuclear Regulatory Commission (NRC) [4, paragraph 50.46(b)(1)], specifies that the 0.95-quantile of the peak cladding temperature (PCT) during a hypothesized loss-of-coolant accident (LOCA) must not exceed 2200° F. The NRC permits a plant licensee to demonstrate its facility's compliance with federal regulations using a 95/95 analysis with Monte Carlo simulation; see Section 24.9 of [5] and [6]. This entails running a computer code [7] modeling the postulated event multiple times with randomly generated inputs to explore a range of uncertainties [8], using the code's outputs to construct a 95% *upper confidence bound* (UCB) for the 0.95-quantile of the PCT, and verifying that the UCB lies below the fixed threshold 2200° F. The UCB accounts for the statistical error of the Monte Carlo estimate due to sampling variability, and the difference between the fixed threshold and the UCB provides a type of safety margin.

Several papers have developed Monte Carlo methods for performing a 95/95 analysis. For example, [9] and [10] apply an approach of [11] based on order statistics, which is valid when employing *simple random sampling* (SRS). However, SRS can produce unusably noisy estimates of the 0.95-quantile, so [12] incorporates VRTs, including antithetic variates (AV) and Latin hypercube sampling (LHS), to obtain more statistically efficient quantile estimators; see Chapter 4 of [13]

* Corresponding author.
*E-mail addresses:* aa655@njit.edu (A. Alban), had8@njit.edu (H.A. Darji), ai64@njit.edu (A. Imamura), marvin@njit.edu (M.K. Nakayama).

http://dx.doi.org/10.1016/j.ress.2017.04.001
Received 25 July 2016; Received in revised form 16 March 2017; Accepted 11 April 2017
Available online 18 April 2017
0951-8320/ 

and Chapter V of [14] for overviews of these and other VRTs for estimating a mean. In addition, [12] provides a UCB for the quantile using a finite-difference approach of [15] and [16]. Utilizing VRTs in nuclear PSAs is especially important because each simulation run may be computationally expensive, e.g., it may require numerically solving systems of differential equations.

In addition to the condition on the PCT, the NRC further imposes requirements on other *criteria*—the core-wide oxidation (CWO<1%) and maximum local oxidation (MLO<17%)—for a PSA of a hypothesized NPP LOCA; see paragraph 50.46(b) of [4]. (The NRC is currently considering replacing MLO with another criterion, the integral time at temperature [17].) The papers [9] and [10] describe approaches to extend the 95/95 analysis based on an SRS quantile estimator for a single criterion to handle multiple criteria with *fixed* capacities.

RISMC differs from a 95/95 analysis in several important ways. First, RISMC assumes each criterion's capacity is *random* rather than fixed. Moreover, instead of examining a quantile, as in a 95/95 study, RISMC measures a safety margin through the *failure probability* $\theta$ that any random load exceeds its random capacity, where $\theta$ should be smaller than a given threshold $\theta_0$, which may be specified by a regulator. RISMC also decomposes a postulated event into scenarios via an *event tree* [18], as in a probabilistic risk analysis (PRA), with uncertainties in each scenario, and Monte Carlo is employed to estimate each scenario's failure probability. To account for the statistical sampling error of the Monte Carlo estimate of the failure probability $\theta$, one should further provide a UCB (or a two-sided confidence interval) for $\theta$, and check if the UCB lies below $\theta_0$. The pilot RISMC studies in [3] and [19] consider only a *single* criterion, PCT, and assume its capacity follows a triangular distribution. These papers apply a combination of stratified sampling (SS) and LHS, but they do not describe how to build a UCB for $\theta$. (Other issues investigated in [3] and [19] include exploring the impact on $\theta$ from altering the distributions of input random variables and from operational changes, e.g., power uprates. A RISMC evaluation may further want to determine the core-damage frequency $\kappa$, which can be estimated by multiplying an estimator of $\theta$ with the (known) frequency of the postulated event.)

Our paper devises Monte Carlo methods to analyze a broad class of systems (not only for RISMC), with uncertainties encapsulated in a *basic random object*. The basic random object can be a random vector, as in structural reliability, where its entries are called *basic variables* [20, Section 1.5], which may be dependent and can represent, e.g., random loads, environmental factors, and material properties. But the basic random object may be more general, e.g., a stochastic process, as in time-dependent reliability [20, Chapter 6]. For example, the load and capacity on a system may vary randomly over time, which is modeled as a stochastic process. We also specify system failure in a general way, as a given binary-valued function of the basic random object. An example is a *series system*, where the basic random object is a random vector of the loads and capacities for a fixed number $q$ of (dependent) criteria, and the system fails when any criterion's load exceeds its capacity, but our framework allows for many other possibilities.

We consider applying combinations of SS, LHS, and conditional Monte Carlo (CMC) to estimate $\theta$. We formally prove that SS+LHS (resp., SS+CMC+LHS) has smaller asymptotic variance than SS (resp., SS+LHS). We also use replicated LHS (rLHS) [21] to construct UCBs for $\theta$, and we prove the UCBs' asymptotic validity (as the total sample size grows large, with the number of replicates fixed). (Although we focus on providing UCBs, our methods can be easily modified to produce a lower confidence bound or two-sided confidence interval.) The combination of SS, CMC, and LHS can be especially effective in reducing variance, as we show through numerical experiments.

We also give insight into why SS+CMC+LHS can work so well. As shown in [22, Section 10.3], LHS substantially reduces variance when the response whose expectation we are estimating is well approximated by an additive function of the input random variables. As we are estimating a probability, the response without CMC is an indicator, which has a poor additive approximation. Thus, LHS by itself may not reduce variance much. In contrast, the conditioning of CMC produces a “smoother” response (in general, no longer binary-valued) with a better additive fit. Hence, LHS can yield much smaller variance when combined with CMC; [23] observes similar synergies between CMC and LHS.

In addition to [3] and [19], other related work includes [24], which provides UCBs for the *single-criterion* (i.e., $q$=1) failure probability when applying combinations of SS, CMC, and rLHS, but the paper does not give proofs of the UCBs' asymptotic validity nor for the reduction in asymptotic variance. The paper [25] estimates the single-criterion failure probability using SS and CMC, but leaves out LHS. Also, [23] provides a theoretical analysis of integrated VRTs (CMC, control variates, and correlation-induction schemes, including LHS and AV) in a general setting, but it does not include SS, which plays a key role in the RISMC framework, nor does the paper provide UCBs, as we do here. The paper [26] combines CMC with AV for estimating the failure probability in structural reliability, but it does not incorporate SS. As explained in [23], AV and LHS can be viewed as special cases of correlation-induction methods, but AV is often outclassed by LHS, especially in combination with CMC.

The methodologies developed in our current paper not only apply for nuclear PSAs, but also can be employed in a wide spectrum of other fields. Civil engineers need to compute the failure probability $\theta$ of a structure (e.g., a building, bridge, or dam) [20]. Insurance companies work with catastrophe models to determine the likelihood of infrastructure failures when subjected to hurricanes, floods, and earthquakes [27]. The Basel Accords [28] specify capital requirements (i.e., capacities) for financial institutions to ensure that they can absorb reasonable losses (i.e., loads). Other examples arise in the assessment of safety and reliability of nuclear weapons [29–31] and the disposal of radioactive waste [32–34].

Compared to SRS, LHS tries to sample more evenly from the sample space, which can produce less-variable estimators of measures of central tendency, e.g., the mean. Combining LHS with SS and CMC can lead to further substantial improvements, as our numerical results show, but perhaps not enough when estimating rare-event probabilities, smaller than say $10^{-4}$. In such cases, other Monte Carlo methods may be more effective. In [35] and [36], the authors develop parametric approximations to the failure probability in terms of less-rare events, which are easier to estimate accurately via Monte Carlo. (In contrast, rather than devising an approximation, the methods in our paper work with the exact model.) The paper [37] employs the method of [35] to estimate the failure probability of a ship hull girder, modeled as a series system. In [38] the authors consider a method they call subset simulation, also known as splitting [14, Section VI.9]; the idea is to contain the rare failure event in a sequence of successively larger, less-rare events, and the failure probability is estimated as a product of estimates of conditional probabilities based on successive events. In [39] the author combines a type of CMC called directional simulation with importance sampling.

To summarize, the main contributions of the current paper are as follows:

- We develop combinations of SS, CMC, and LHS to efficiently estimate a failure probability $\theta$. Our framework allows for a failure to be defined quite generally as a function of a basic random object, which may be a random vector or something more general, e.g., a stochastic process or random field. The combination of the three VRTs can produce substantial variance reduction, as our numerical results in Section 7 indicate. We also establish theory (Theorems 1 and 3) and provide insight (Sections 5.1 and 6.1) into why this is so.
- When applying SS+rLHS or SS+CMC+rLHS, we derive UCBs for $\theta$, which are crucial to account for the statistical error arising from