Proceedings of the 2019 Winter Simulation Conference
N. Mustafee, K.-H.G. Bae, S. Lazarova-Molnar, M. Rabe, C. Szabo, P. Haas, and Y.-J. Son, eds.

RANDOM PERTURBATION TO QUANTIFY INPUT UNCERTAINTY

Huajie Qian

Department of Industrial Engineering and Operations Research
Columbia University
500 W. 120th Street
New York, NY 10027, USA

ABSTRACT

We consider the problem of estimating the output variance in simulation analysis that is contributed from the statistical errors in fitting the input models, the latter often known as the input uncertainty. This variance contribution can be written in terms of the sensitivity estimate of the output and the variance of the input distributions or parameters, via the delta method. We study the direct use of this representation in obtaining efficient estimators for the input-contributed variance, by using finite-difference and random perturbation to approximate the gradient, focusing especially in the nonparametric case. In particular, we analyze a particular type of random perturbation motivated from resampling. We illustrate the optimal simulation allocation and the simulation effort complexity of this scheme, and show some supporting numerical results.

1 INTRODUCTION

In stochastic simulation, input errors arise when the input models used to drive the simulation runs are misspecified or noisily estimated from data. Input errors can propagate to the outputs and lead to an incorrect conclusion, which calls for a need to quantify the impact of these errors.

A common way in quantifying input uncertainty is to estimate *input variance*, i.e., the variance of the output that is contributed from the input noise. Established approaches roughly fall into two categories. The first is to use bootstrap resampling. This involves a two-layer nested simulation where one first draws resamples from the input data and then uses the resampled input models to drive simulation replications. The input variance can be approximated using the bootstrapped variance. In the Bayesian contexts, this also leads to related approaches utilizing posterior sampling. The second line of approach uses delta-method approximation, which involves estimating the sensitivity or gradient of the outputs with respect to the input distributions or parameters, and combining with the estimation variance of these input quantities.

In this paper, we consider the delta-method approximation to estimate input variance. We consider especially the nonparametric case, which to our best knowledge has not been studied. Formally, let $X_1,\dots,X_n$ be i.i.d. data from the single true input model F, $\hat{F}$ be the empirical distribution formed from the data, and $\psi(\cdot)$ be the expected performance measure of interest under some input model. The input variance can then be expressed as $\mathrm{Var}(\psi(\hat{F}))$, for which an asymptotically accurate nonparametric estimator is

$$\mathrm{Var}(\psi(\hat{F})) \approx \frac{1}{n}\mathrm{Var}_{\hat{F}}(IF(X;\hat{F})) := \frac{1}{n^2}\sum_{i=1}^{n} IF(X_i;\hat{F})^2 \tag{1}$$

where each $IF(\cdot;\hat{F})$ is the influence function at the empirical input model which is defined as the directional derivative $IF(x;\hat{F}) := \lim_{\varepsilon\to 0+}\frac{1}{\varepsilon}(\psi((1-\varepsilon)\hat{F}+\varepsilon\delta_x)-\psi(\hat{F}))$ where δ_x denotes the delta measure at x.

The nonparametric case considered here features a growing problem dimension compared to the parametric counterpart, leading to a fast-growing computational burden in estimating the input variance.

Note that the number of influence function components in (1) grows linearly in the data size n. If one computes (1) by estimating each individual component $IF(X_i;\hat{F})$ separately, then the required simulation effort has to scale at least linearly in n because the estimation of each component costs at least a constant number of simulation runs. The heavy computational burden in obtaining accurate input variance estimate has also been seen in approaches using bootstrap resampling where both layers need adequate sample sizes and lead to a multiplicatively large overall required simulation size.

The main contribution of the paper is a computationally efficient approach for estimating input variance using delta-method approximations like (1).

2 THE RANDOM PERTURBATION APPROACH

In the basic form, our approach to estimate gradient (influence function) is to use random perturbation, namely by first generating a random vector in the space of the parameters, and suitably projecting to each dimension to obtain a finite-difference estimate for all components of the gradient. The hope is that the required simulation burden can be reduced through the simultaneous estimation for all components, an idea similar to simultaneous perturbation or smoothing (e.g., Flaxman et al. 2005) used in zero-th order stochastic optimization. Specifically, we estimate each $IF(X_i;\hat{F})$ in (1) by

$$\widehat{IF}(X_i;\hat{F}) = \frac{1}{R}\sum_{r=1}^{R}\frac{1}{h}\hat{\psi}_r((1-h)\hat{F}+hG)S_i(G) \text{ for all } i=1,\ldots,n \tag{2}$$

where G is a discrete distribution supported on the data $\{X_1,\ldots,X_n\}$ that is randomized, each $S_i(G)$ is some "score function" of the perturbation such that $E[\psi((1-h)\hat{F}+hG)S_i(G)/h|\hat{F}] \approx IF(X_i;\hat{F})$ for all $i=1,\ldots,n$, and each $\hat{\psi}_r((1-h)\hat{F}+hG)$ is an independent unbaised simulation replication driven by the input $(1-h)\hat{F}+hG$ for some $h\in(0,1]$. We show that, when the random perturbation $G=(G_1,\ldots,G_n)$ has exchangeable weights, the score function takes the convenient form $S_i(G)=\frac{n-1}{n\text{Var}(G_1)}(G_i-\frac{1}{n})$.

3 PERTURBATION MOTIVATED FROM RESAMPLING

We consider a special class of random perturbations that are motivated from resampling of the input data, and show the superior computational efficiency as well as some other desirable properties of the resulting estimator. Consider a random vector $(s_1,\ldots,s_n)$ following the multinomial distribution that counts s uniform draws ($\sum_{i=1}^n s_i = s$) from n objects. We set the i-th component of perturbation to be $G_i = s_i/s$, and h to be 1 so that $(1-h)\hat{F}+hG = G$. Such a perturbation G is identical to the empirical distribution of a resample of size s uniformly drawn from the data with replacement.

Note that the gradient estimator shown in (2) involves only a single random perturbation G. To implement this, the natural approach is to average the simultaneous estimator (2) over multiple i.i.d. perturbations, arriving at a nested scheme like in bootstrap resampling. However, our main results show that such a nested scheme is not necessary, and our required computational effort is much lighter. More precisely, we show

1. Given a total simulation budget N, the asymptotically optimal allocation that minimizes the estimation error of the input variance is to generate N random perturbations and to run a single simulation replication for each perturbation
2. The required simulation burden N to consistently estimate the input variance is such that $N/\sqrt{n}\to\infty$ if s is chosen such that $s\to\infty$ as $n\to\infty$.

REFERENCES

Flaxman, A. D., A. T. Kalai, A. T. Kalai, and H. B. McMahan. 2005. "Online Convex Optimization in the Bandit Setting: Gradient Descent Without a Gradient". In *Proceedings of the* 16^{th} *Annual ACM-SIAM Symposium on Discrete Algorithms*, 385–394. Philadelphia, Pennsylvania: Society for Industrial and Applied Mathematics.