Proceedings of the 2016 Winter Simulation Conference
T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, eds.

INPUT-OUTPUT UNCERTAINTY COMPARISONS FOR OPTIMIZATION VIA SIMULATION

Eunhye Song

Department of Industrial Engineering & Management Sciences
Northwestern University
2145 Sheridan Road
Evanson, Illinois 60208, USA

ABSTRACT

When an optimization via simulation (OvS) procedure designed for known input distributions is applied to a problem with input uncertainty (IU), it typically does not provide the target statistical guarantee. In this paper, we focus on a discrete OvS problem where all systems share the same input distribution estimated from the common input data (CID). We define the CID effect as the joint impact of IU on the outputs of the systems caused by common input distributions. Our input-output uncertainty comparison (IOU-C) procedure leverages the CID effect to provide the joint confidence intervals (CIs) for the difference between each system's mean performance and the best of the rest incorporating both input and output uncertainty. Under mild conditions, IOU comparisons provide the target statistical guarantee as the input sample size and the simulation effort increase.

1 INTRODUCTION

In many simulation applications, real-world input distributions are estimated based on a limited number of observations, which is the source of IU. As the simulation output is a functional of the estimated input distribution, a discrete OvS procedure designed to choose the system with the largest (or smallest) output mean with a target probability guarantee given the true input distributions may not guarantee the same under IU. However, even in the presence of IU we may still be able to provide the target probability guarantee, if the systems in consideration are affected similarly by IU. Moreover, clearly inferior solutions may still be excluded.

We focus on the case when all systems are simulated with the same input models estimated from the same data. By estimating the joint distribution of the CID effect for all systems' simulation outputs, we create the IOU-C procedure that provides simultaneous CIs for the difference between each system's mean and the best of the rest under both input and stochastic uncertainty with asymptotic probability guarantee as the real-world sample size and the simulation effort increase. This is an extension of multiple comparisons with the best (MCB) procedures to the case with IU. IOU-C procedure also provides a subset of systems likely to be the best with the given probability guarantee. Good empirical performance of IOU-C procedure shows that comparisons under IU may be easier than estimating the actual effect of IU.

2 FRAMEWORK

Our focus is on the case where the real-world input distributions have known distribution families with unknown parameters, which are estimated by maximum-likelihood estimators (MLEs). All k systems in contention share the same real-world input distribution $F(\cdot \mid \theta^c)$ with distribution family F and true parameter vector θ^c, and $\widehat{\theta}$ is the MLE of θ^c based on m real-world observations. Under some regularity conditions, $\sqrt{m}(\widehat{\theta} - \theta^c) \xrightarrow{D} \mathrm{N}(\mathbf{0}_{|\theta^c|}, \Sigma(\theta^c))$ as $m \to \infty$, where $\Sigma(\theta^c)$ is the asymptotic variance-covariance matrix of

$\widehat{\theta}$. Output from system i simulated using $F(\cdot \mid \widehat{\theta})$ is $Y_i(\widehat{\theta}) = \eta_i(\widehat{\theta}) + \varepsilon_i(\widehat{\theta})$, where $\eta_i(\widehat{\theta}) \equiv \mathrm{E}[Y_i(\widehat{\theta})|\widehat{\theta}]$ and $\varepsilon_i(\widehat{\theta})$ is distributed with mean 0 and variance $\sigma_i^2(\widehat{\theta})$.

When θ^c is known, Chang and Hsu (1992) show that we can obtain $1-\alpha$ MCB CIs by finding $w_{i\ell}, i \neq \ell$, such that $\Pr\{\bar{Y}_i(\theta^c) - \bar{Y}_\ell(\theta^c) - (\eta_i(\theta^c) - \eta_\ell(\theta^c)) \geq -w_{i\ell}, \forall \ell \neq i\} \geq 1-\alpha$ for $1 \leq i \leq k$. The joint distribution of $\{\varepsilon_i(\theta^c)\}_{i=1}^k$ is the key to find such $w_{i\ell}, i \neq \ell$. However, when $\widehat{\theta} \neq \theta^c$, we also need the joint distribution of $\{\eta_i(\widehat{\theta}) - \eta_\ell(\widehat{\theta})\}_{i \neq \ell}$. IOU-C procedure splits $w_{i\ell}$ into two parts, i.e., $w_{i\ell} = w_{i\ell}^{(1)} + w_{i\ell}^{(2)}$, where $\Pr\{\eta_i(\widehat{\theta}) - \eta_\ell(\widehat{\theta}) - (\eta_i(\theta^c) - \eta_\ell(\theta^c)) \geq -w_{i\ell}^{(1)}, \forall \ell \neq i\} \geq 1-\alpha_1$ and $\Pr\{\bar{\varepsilon}_i(\widehat{\theta}) - \bar{\varepsilon}_\ell(\widehat{\theta}) \geq -w_{i\ell}^{(2)}, \forall \ell \neq i\} \geq 1-\alpha_2$ for $1 \leq i \leq k$, and $1-\alpha = (1-\alpha_1)(1-\alpha_2)$. Finding $w_{i\ell}^{(2)}, i \neq \ell$, is the same as in a traditional MCB procedure that assumes only stochastic uncertainty exists. Finding the interval lengths due to CID effects, $w_{i\ell}^{(1)}, i \neq \ell$, is the main contribution of IOU-C.

3 INTERVAL LENGTHS DUE TO CID EFFECTS

Assuming $\eta_i(\cdot)$ is a smooth function of $\widehat{\theta}$, the first-order Taylor series approximation gives $\eta_i(\widehat{\theta}) \approx \eta_i(\theta^c) + \beta_i^\top(\widehat{\theta} - \theta^c)$, where $\beta_i = \nabla \eta_i(\theta^c)$. Hence, $\beta_i^\top(\widehat{\theta} - \theta^c)$ captures the CID effect on system i. Similarly, $\{(\beta_i - \beta_\ell)^\top(\widehat{\theta} - \theta^c)\}_{\ell \neq i}$ captures the difference in the CID effects on the output of systems i and $\ell \neq i$, which is asymptotically normally distributed with mean $\mathbf{0}_{k-1}$, and $\mathrm{Var}((\beta_i - \beta_\ell)^\top(\widehat{\theta} - \theta^c)) = (\beta_i - \beta_\ell)^\top \Sigma(\theta^c)(\beta_i - \beta_\ell)/m$ and $\mathrm{Cov}((\beta_i - \beta_\ell)^\top(\widehat{\theta} - \theta^c), (\beta_i - \beta_{\ell'})^\top(\widehat{\theta} - \theta^c)) = (\beta_i - \beta_\ell)^\top \Sigma(\theta^c)(\beta_i - \beta_{\ell'})/m$. Therefore, estimating the joint distribution of $\{\eta_i(\widehat{\theta}) - \eta_\ell(\widehat{\theta})\}_{\ell \neq i}$ reduces to estimating $\beta_i, \forall i$, and $\Sigma(\theta^c)$. The latter can be approximated by $\Sigma(\widehat{\theta})$.

IOU-C procedure estimates β_i as follows: i) bootstrap $\widehat{\theta}^{(1)}, \widehat{\theta}^{(2)}, \ldots, \widehat{\theta}^{(B)}$ from $\mathrm{N}(\widehat{\theta}, \Sigma(\widehat{\theta})/m)$; ii) simulate $Y_i(\widehat{\theta}^{(b)})$ using $F(\cdot \mid \widehat{\theta}^{(b)})$ for $1 \leq b \leq B$; iii) compute $\widehat{\beta}_i$ by fitting a linear regression of $Y_i(\widehat{\theta}^{(b)})$ on $\widehat{\theta}^{(b)} - \widehat{\theta}$ for $1 \leq b \leq B$. Unfortunately, $\mathrm{E}[\widehat{\beta}_i] \neq \beta_i$ in general, unless η_i is a linear function of $\widehat{\theta}$. Under some smoothness conditions on η_i, however, $\widehat{\beta}_i \xrightarrow{p} \beta_i$ as $m \to \infty$ and $B = m^\gamma, 1 < \gamma < 2$. The condition on B is to achieve consistency by balancing the variance and the bias of $\widehat{\beta}_i$ as $m \to \infty$.

Plugging $\widehat{\beta}_1, \widehat{\beta}_2, \ldots, \widehat{\beta}_k$ into the distribution of $\{(\beta_i - \beta_\ell)^\top(\widehat{\theta} - \theta^c)\}_{\ell \neq i}$, we can obtain $w_{i\ell}^{(1)}, \ell \neq i$. IOU-C finds $w_{i\ell}^{(1)}, \ell \neq i$ that still provides $1-\alpha_1$ asymptotic probability guarantee incorporating the estimation error in the regression by solving the following optimization problem for each $(i,\ell), i \neq \ell$:

$$
\begin{aligned}
l_{i\ell} = \min \quad & (\beta_i - \beta_\ell)^\top(\widehat{\theta} - \theta^c) \\
\text{subject to} \quad & \mathscr{B}_i \in \mathrm{CR}_1(1-\alpha_{11}), \\
& (\widehat{\theta} - \theta^c) \in \mathrm{CR}_2(1-\alpha_{12}),
\end{aligned}
$$

where $\mathscr{B}_i^\top = \{(\beta_i - \beta_1)^\top, (\beta_i - \beta_2)^\top, \ldots, (\beta_i - \beta_{i-1})^\top, (\beta_i - \beta_{i+1})^\top, \ldots, (\beta_i - \beta_k)^\top\}$, $\mathrm{CR}_1(1-\alpha_{11}) \subset \mathbb{R}^{p(k-1)}$ is a $1-\alpha_{11}$ confidence region of $\mathscr{B}_i$ obtained from the regressions, and $\mathrm{CR}_2(1-\alpha_{12}) \subset \mathbb{R}^p$ includes $\widehat{\theta} - \theta^c$ with probability $1-\alpha_{12}$, which can be obtained from the estimated asymptotic distribution of $\widehat{\theta}$. Choosing α_{11} and α_{12} such that $(1-\alpha_{11})(1-\alpha_{12}) = 1-\alpha_1$, $w_{i\ell}^{(1)} = -l_{i\ell}, i \neq \ell$, provide the interval lengths due to CID effects with the desired $1-\alpha_1$ asymptotic probability guarantee.

Empirical results show that IOU-C provides a probability guarantee $> 1-\alpha$, whereas an MCB procedure that assumes $\widehat{\theta} = \theta^c$ fails to. Also, the true best system is included in the subset of best systems with probability $> 1-\alpha$, whereas the MCB procedure excludes it from the subset with probability $\gg \alpha$.

REFERENCES

Chang, J. Y., and J. C. Hsu. 1992. "Optimal designs for multiple comparisons with the best". *Journal of Statistical Planning and Inference* 30 (1): 45 – 62.