European Journal of Operational Research 271 (2018) 733–745

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Innovative Applications of O.R.

Self-adjusting the tolerance level in a fully sequential feasibility check procedure

Mi Lim Lee[a], Chuljin Park[b,*], Dong Uk Park[b]

[a] College of Business Administration, Hongik University, Seoul 04066, South Korea
[b] Department of Industrial Engineering, Hanyang University, Seoul 04763, South Korea

ARTICLE INFO

Article history:
Received 13 December 2016
Accepted 21 May 2018
Available online 25 May 2018

Keywords:
Simulation
Stochastic constraints
Ranking and selection
Fully sequential algorithms
Multiple performance measures

ABSTRACT

We consider the problem of determining the feasibility of systems when the performance measures in stochastic constraints need to be evaluated via simulation. We develop a new procedure, namely the adaptive feasibility check procedure. Specifically, the procedure uses an existing feasibility check procedure iteratively as its subroutine with a decreasing sequence of tolerance levels. Our procedure is designed to return the set of strictly feasible systems with at least a prespecified probability. The validity and efficiency of the procedure are investigated through both analytical and experimental results. The procedure is also tested using numerical examples.

1. Introduction

Ranking and selection (R&S) procedures have been actively used for finding the best system among a finite number of simulated systems with some statistical guarantees. In this case, the best system represents the one with the best expected primary performance measure when the measure can be estimated by stochastic simulation. General approaches to R&S appear in a few different forms, such as the fully sequential indifference zone (IZ) frameworks (Kim & Nelson, 2006), the optimal computing budget allocation (OCBA) frameworks (Chen, Lin, Yücesan, & Chick, 2000), and the Bayesian frameworks (Chick, 2006; Frazier & Kazachkov, 2011).

Recently, scholars and practitioners have been interested in R&S problems in the presence of stochastic constraints. Hunter and Pasupathy (2013), Lee, Pujowidianto, Li, Chen, and Yap (2012), and Pasupathy, Hunter, Pujowidianto, Lee, and Chen (2015) provide OCBA frameworks that allocate a finite sampling budget to maximize the probability of correctly selecting the best feasible system when stochastic constraints exist. Gao and Chen (2017a) suggest a feasibility determination procedure for multiple stochastic constraints based on the OCBA framework and Gao and Chen (2017b) combine the procedure with an optimization via simulation algorithm to find the best feasible system among a finite number of systems. In addition, the normality assumption of the R&S has been relaxed through the OCBA framework based on the large deviation theory (Szechtman & Yücesan, 2008) and a Bayesian approach (Szechtman & Yücesan, 2016).

Using fully sequential IZ frameworks, Batur and Kim (2010) suggest feasibility check procedures (FCPs) for finding a set of feasible or near feasible systems regarding multiple stochastic constraints. The procedures are used as a subroutine for selecting the best system that satisfies constraints on one or more secondary performance measures. Andradóttir and Kim (2010) and Healey, Andradóttir, and Kim (2013) propose statistically valid procedures that select the best feasible system with at least a prespecified probability in the presence of a single stochastic constraint. The procedures have been improved in Healey, Andradóttir, and Kim (2014) for selecting the best feasible system under multiple stochastic constraints. In this paper, we focus on the fully sequential IZ framework for constrained R&S.

The fully sequential FCPs in Batur and Kim (2010) have been also used to find a set of feasible systems during or after running optimization algorithms when the number of systems is large. Ahmed and Alkhamis (2009) find the optimal number of doctors, lab technicians, and nurses to operate an emergency medical center by adopting a two-phase approach. In this approach, the first phase includes the FCP in Batur and Kim (2010) to find a set of feasible or near feasible systems and the second phase includes an optimization algorithm proposed in Alkhamis and Ahmed (2004). Tsai and Fu (2014) combine the FCP with a genetic algorithm to solve a discrete optimization via simulation problem with a single stochastic constraint. In addition, Tsai and Liu (2015) and Tsai and Chen (2016) propose simulation optimization frameworks using the FCP to solve some inventory management problems.

* Corresponding author.
E-mail address: parkcj@hanyang.ac.kr (C. Park).

https://doi.org/10.1016/j.ejor.2018.05.045
0377-2217/

Although the FCP is useful itself or within constrained R&S frameworks and optimization algorithms, the procedure requires users to select a proper tolerance level, which is very important in decision-making. If the tolerance level is too small, computational costs for feasibility checks are unnecessarily high. On the other hand, if the tolerance level is too large, the quality of feasibility decisions could be compromised, and the user must allow undesired and unpredictable decisions in which a truly infeasible system is declared as feasible or vice versa with high chances. Declaring a truly infeasible system as feasible may lead to a critical situation when the user considers a hard constraint which must be satisfied. From the viewpoint of optimization, since the constraints almost always hit their limits to optimize the objective, the feasibility decisions of near feasible systems make a difference in stochastically constrained optimization problems (Park & Kim, 2015).

Nevertheless, selecting a proper tolerance level is difficult and time-consuming. In practice, mean and variance configurations are generally unknown, and thus the user needs to run the FCP several times with different values of the tolerance levels. As the number of constraints increases, the number of replications for running the FCP significantly increases and therefore, finding proper tolerance levels with multiple constraints becomes more inconvenient and time-consuming. Even though it is possible to figure out proper values of tolerance levels, it is difficult to statistically guarantee that the FCP returns a set of strictly feasible systems with a prespecified probability. In this paper, we propose a new FCP that is self-adjusting the tolerance level for each system and each constraint. The new procedure is designed to provide a set of feasible systems regarding multiple stochastic constraints while guaranteeing a predetermined probability of a correct decision. See Lee, Park, Park, and Park (2017) for a preliminary work on the new procedure applied to find the optimal number of medical staffs in an emergency department.

This paper is organized as follows: In Section 2, we explain our problem and introduce the existing FCPs as background. Section 3 provides a generic description of the new FCP and its statistical guarantees with proofs. Efficiency and effectiveness of the new procedure are examined analytically and empirically in Section 4. Section 5 concludes the paper.

2. Background

In this section, we formulate our problem, provide notations and assumptions, and introduce existing procedures that are used for determining a set of feasible or near feasible systems with statistical guarantees.

2.1. Problem formulation

We consider k systems whose performance measures can be observed through stochastic simulation. Let Θ denote a set of all systems (i.e., $\Theta = \{1, 2, \ldots, k\}$) and $Y_{i\ell j}$ for $j = 1, 2, \ldots$ denote the jth simulation observation associated with the ℓth performance measure of system i. For any given system i and performance measure ℓ, $y_{i\ell}$ denotes the expectation of $Y_{i\ell j}$ (i.e., $y_{i\ell} = \mathbf{E}[Y_{i\ell j}]$) and $\sigma^2_{i\ell}$ denotes the variance of $Y_{i\ell j}$ (i.e., $\sigma^2_{i\ell} = \mathbf{Var}[Y_{i\ell j}]$). Then our problem is determining the set of strictly feasible systems:

$$\Upsilon := \{i \in \Theta \mid y_{i\ell} < q_\ell, \ell = 1, 2, \ldots, s\}, \tag{1}$$

where q_ℓ, $\ell = 1, 2, \ldots, s$, are threshold constants associated with the ℓth performance measure.

We make the following assumption throughout the paper.

Assumption 1. For each $i = 1, 2, \ldots, k$,

$$\begin{bmatrix} Y_{i1j} \\ Y_{i2j} \\ \vdots \\ Y_{isj} \end{bmatrix} \overset{iid}{\sim} MN_s\left(\begin{bmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{is} \end{bmatrix}, \Sigma_i\right),$$

where $\overset{iid}{\sim}$ denotes independent and identically distributed, MN_s denotes the $s-$dimensional multivariate normal distribution, and Σ_i is the $s \times s$ covariance matrix of the vector $(Y_{i1j}, Y_{i2j}, \ldots, Y_{isj})$.

Note that simulation observations are often assumed to be normally distributed in R&S problems. In practice, the normality assumption can be justified if $Y_{i\ell j}$ values are obtained by either within-replication averages or batch means. It is well-known that using common random numbers (CRN) often improves the efficiency of statistical comparison procedures, but it is not beneficial for FCPs, as mentioned in Batur and Kim (2010). Nevertheless, we consider the case of CRN as well as the independent case because the proposed FCP in this paper can be combined with procedures for selecting the best system such as the fully sequential procedure in Kim and Nelson (2001).

Handling boundary systems is one of the important issues in stochastically constrained optimization via simulation. Park and Kim (2015) develop a new penalty method, namely Penalty Function with Memory, that can theoretically handle such systems with probability one if we are allowed to use an infinite number of observations. Nevertheless, if only finite observations are available, it is extremely difficult to guarantee at least a prespecified probability of correctly determining feasibility of boundary systems because of estimation errors. The purpose of this paper is to design a procedure to find the set of feasible systems regarding stochastic constraints with a finite number of observations, and thus we consider the following assumption regarding boundary systems throughout the paper.

Assumption 2. For each $i = 1, 2, \ldots, k$ and $\ell = 1, 2, \ldots, s$, there is no system i and constraint ℓ such that $y_{i\ell} = \mathbf{E}[Y_{i\ell j}] = q_\ell$.

2.2. Existing feasibility check procedures

In this section, we introduce two fully sequential procedures, denoted by $\mathcal{F}^I_B$ and $\mathcal{F}^I_A$, of Batur and Kim (2010). These two procedures are originally designed not to find a set of strictly feasible systems (i.e., the set Υ of (1)), but to find a set of feasible or near feasible systems. Specifically, $\mathcal{F}^I_B$ and $\mathcal{F}^I_A$ adopt a user-specified parameter, which is a positive real number, namely the tolerance level for each constraint. Let ϵ_ℓ denote the tolerance level corresponding to constraint ℓ, then we can define the following three sets as in Batur and Kim (2010):

$$D \equiv \{i \in \Theta \mid y_{i\ell} \le q_\ell - \epsilon_\ell, \text{ for } \ell = 1, 2, \ldots, s\};$$
$$A \equiv \{i \in \Theta \mid y_{i\ell} < q_\ell + \epsilon_\ell, \text{ for } \ell = 1, 2, \ldots, s\} \setminus D; \text{ and}$$
$$U \equiv \cup_{\ell=1}^{s} \{i \in \Theta \mid q_\ell + \epsilon_\ell \le y_{i\ell}\}.$$

The systems in the sets D, A, and U are called desirable, acceptable, and unacceptable systems, respectively. Any desirable system is clearly feasible regarding constraints and thus $D \subseteq \Upsilon$. Any unacceptable system is clearly infeasible regarding constraints and thus $U \subseteq (\Theta \setminus \Upsilon)$. The set of acceptable systems could include both feasible and infeasible systems. That is, an acceptable system can be declared as feasible or infeasible by the existing FCPs regardless of the true feasibility of the system. Fig. 1 demonstrates regions of the desirable, acceptable, and unacceptable systems with a single stochastic constraint. Under existence of the tolerance levels, $\mathcal{F}^I_B$ and $\mathcal{F}^I_A$ are guaranteed to return a set denoted by F, which includes the set D and is included in the set $D \cup A$ (i.e., $D \subseteq F \subseteq (D \cup A)$), within a user-specified confidence level.