

A two-stage stochastic programming approach for care providers shift scheduling problems

Hajar Shirneshan¹; Ahmad Sadegheih^{2*}; H. Hosseini-Nasab³; Mohammad Mehdi Lotfi¹

1 Yazd university

2 Faculty of Industrial Engineering , Yazd University

3 Department of Industrial Engineering, Yazd University, Yazd, Iran

<https://doi.org/10.22105/jarie.2022.349970.1488>

Abstract

Due to the importance of the health field, the problem of determining the shift scheduling of care providers has been addressed in many studies, and various methods have been proposed to solve it. Given the uncertainty in patients' demand, it is an important issue as how to assign care providers to different shifts. One area facing this uncertainty is the provision of services to cancer patients. A stochastic programming model is developed to account for patient demand uncertainty in this study. The sample average approximation method is presented to determine an optimal schedule by minimizing care providers' normal and overtime costs with different contracts and skills. Then the appropriate sample size is determined based on Monte Carlo and Latin Hypercube methods. In the following, the lower and upper bounds of the optimal solution are calculated. As the numerical results of the study show, the convergence of the lower and upper bounds of the optimal solution is equal to 189247.3 dollars and is achieved with a difference of 0.1%. The Monte Carlo simulation method is used to validate the care provider program in the next stage. As shown, in the worst case, the value of the objective function is equal to 197480 dollars.

Keywords: Healthcare, Shift scheduling, Uncertainty, Stochastic programming, Sample average approximation

1. Introduction

Due to the increase of chronic diseases such as Covid-19, the costs of healthcare systems are increasing dramatically [1]. The statistics issued by the World Health Organization indicate that personnel planning will be an important priority in the field of health in the next decade. Due to the high cost of staffing, which accounts for about 40% of the total costs, medical centers must reduce their expenditures. In this regard, proper human resource planning will help significantly. A practical issue in healthcare is the planning of service shifts. This subset of employee scheduling varies greatly depending on the type of job and work regulations. The Nurse Rostering Problem¹

¹ NRP

or Nurse Scheduling Problem² aims to create a schedule and assign available service providers to hospital shifts under various constraints over a period [2]. The significance of this issue for improving service quality, staff satisfaction, health conditions, and reducing hospital costs has encouraged researchers to study it. Solving this problem results in a schedule that specifies how many people are required for different skills and when they can offer their service on a given planning horizon. The program must comply with labor laws, employee preferences, employee availability, labor demand, workload and demand, employment contracts, and ergonomic and technical constraints. The high variety of problems in modeling, making assumptions, and solving methods have increased the attractiveness of those problems. In the field of home care for cancer patients, shift planning is of particular importance due to the existence of different contracts for service providers, the high cost of specialized services in the field, and the high uncertainties in patients' conditions.

This program determines how to assign service providers with different skills and contracts to work shifts. In the Iranian Health Control Center, care providers' contracts with different skills are set in three modes: full-time, part-time, and hourly. In addition, the cost of service is increased by 20 to 40 percent per hour as a contract change from full-time to part-time or hourly. Therefore, planning for the proper assignment of service providers can significantly save system costs. As the case is, most medical centers use manual planning, and the Iranian Health Control Center is no exception [3].

To address the issue of planning service providers in the Iranian Health Control Center, the present study focuses on providing specialized services to cancer patients in the event of demand uncertainty. Thus, a double-stage stochastic model is presented for programming. In the first step, a mechanism is devised to assign care providers to work shifts. The amount of overtime required for each skill on each shift is determined in the second stage.

2. Literature review

Medical staff planning has been a topic of research since the 1950s. According to Ernst et al. [4], it is not easy to make a schedule that can satisfy employees' needs. The task of medical staff planning is often complicated by staffing requirements as well as government and hospital regulations. Planners should consider the conditions and the number of the patients, the expertise, work experience, and preferences of the medical staff, the hospital policies, and the rules and regulations set by the government [4]. Considering the significance of medical staff scheduling in healthcare, more studies on this issue have been published over the last two decades. Klinz et al. [5] proposed two mathematical models to minimize the total number of work shifts and nurses' general unhappiness. Topaloglu et al. [6] presented a fuzzy goal programming model for NSPs to measure the uncertainty through the objective evaluation of hospital regulations and nurse preferences. Topaloglu and Selim [7] introduced a multi-objective integer program for NSPs to produce an equitable schedule for nurses and satisfy hospital management objectives. Landa-silva and Le [8] presented a multi-objective approach to cope with real-world uncertainties in NSPs. To do so, they proposed an evolutionary algorithm to achieve high-quality non-dominated schedules. Ohki et al. [9] established a cooperative genetic algorithm to re-optimize nurse schedules. Zhang et al. [10] presented an optimization algorithm that was hybrid and swarm-based. It combined a variable neighborhood search and a genetic algorithm to cope with a highly-constrained nurse scheduling problem in modern hospitals. Maenhout et al. [11] studied the nurse allocation issue

² NSP

and used the column generation method to deal with it. Santos et al. [12] introduced cutting as a concept in integer programming to solve the corresponding problems innovatively. Ingels and Maenhout [13] considered the effects of defining and including reserve duties in rosters of medium-term shifts for the personnel. They used a three-stage method that imitated the process of workforce management to measure the robustness achieved. After the personnel roster was designed, the events that unexpectedly occurred would be simulated, and an optimization model would determine the adjustments required for the balance of the supply and demand. Bagheri et al. [14] introduced a stochastic mathematical model for an NSP in a heart surgery center to minimize the regular and overtime assignment costs. They assumed that patients' demand and length of stay would be uncertain. So, they used a sample average approximation method to solve the model. In another study, Punnakitikashem et al. [15] sought to minimize the overload of nurses through an integer MP model of a stochastic type. They dealt with the staffing cost as a hard budget constraint in the model. Moreover, they used the Benders' decomposition and Lagrangian relaxation methods to obtain non-dominated solutions. The resulting model was implemented in two medical and surgical wards at the Northeast Texas Hospital. Chen et al. [16] studied an integrated problem of allocating a medical staff and scheduling a general staff under uncertain conditions. They solved the problem by employing a double-stage algorithm to determine a medical staff with the smallest possible size and make the best schedule for it. Ang et al. [17] introduced a decision support system based on a goal programming method for NSPs. They examined the workload distribution, shift equity, and staff satisfaction. They also pursued minimizing the nurse-patient ratio (NPR) calculated based on the number of patients allocated to each nurse. Hamid et al. [18] devised a mathematical model with multiple objectives to schedule a nursing staff, which took the decision-making styles of nurses into account. The objectives addressed in that study were the minimization of the total cost of staffing, minimization of the average index of the incompatibility in the decision-making styles of the nurses assigned to the same shift days, and maximization of the overall satisfaction of nurses with their shifts. Moreover, three meta-heuristics were developed to solve the problem, including the multi-objective Keshtel algorithm, non-dominated sorting genetic algorithm II, and multi-objective tabu search. Hassani et al.[19] developed a sustainable approach with a robust scenario-based optimization method. They proposed the differential evolution(DE) algorithm to solve the problem and compared the performance against the genetic algorithm. The results show that the DE algorithm has good performance. Kheiri et al.[20] studied the multi-stage nurse rostering formulation. They proposed a sequence-based selection hyper-heuristic that uses a statistical Markov model and incorporates an algorithm for building feasible initial solutions. Empirical results and analysis show that the suggested approach has significant potential for difficult problem instances. A brief classification of the models reviewed in the literature is presented in Table 1.

Table 1. A brief review of the literature

Author	Year	Objective	Uncertainty	Approach
Klinz et al.[5]	2006	Minimizing the total number of work shifts and the general unhappiness of all nurses	-	Heuristic
Topaloglu et al. [21]	2007	Minimizing nurse assignment costs	Fuzzy	Exact
Topaloglu & Selim [Y]	2010	Minimizing deviations from nurse preferences and hospital regulations	Fuzzy	Exact
Landa-Silva [8]	2008	Satisfaction of nurse preferences and work regulations	-	Meta-heuristic
Ohki et al. [9]	2012	Minimizing the penalty function to evaluate shift schedules	-	Meta-heuristic

Zhang et al. [22]	2011	Maximizing the quality of objectives concerning the importance of constraints	-	Meta-heuristic
Maenhout et al. [11]	2013	Minimizing the penalty associated with different types of nurses	-	Exact
Santos et al. [12]	2016	Minimizing the penalty of assignment	-	Heuristic
Ingels et al. [13]	2015	Minimizing the allocation penalty and changing the nurse schedule	-	Exact & Simulation
Bagheri et al. [14]	2016	Minimizing the normal and overtime hours of nurses	Stochastic	Sample Average Approximation
Punnakitikashem et al. [15]	2013	Minimizing the excess workload on nurses and the cost of staffing	Stochastic	Benders & Lagrangian
Chen et al. [16]	2016	Minimizing the penalty of the soft constraints of nurses' preferences	-	Exact
Ang et al. [17]	2018	Minimizing the average and maximal deviations from the target ratios of nurse to patient	-	Exact
Hamid et al. [18]	2020	Minimizing the total cost of staffing and the sum of incompatibility among nurses and maximizing the satisfaction of nurses with their assigned shifts	-	Meta-heuristic
Hassani et al. [19]	2021	Minimizing the total cost of allocating shifts to nurses, reserve nurses required, overtime and underemployed costs of a particular type of shift, cost of mismatching the nurse preferences with the roster	robust scenario-based optimization	Meta-heuristic
Kheiri et al. [20]	2021	Minimizing violation of eight soft constraints	-	hyper-heuristic with statistical Markov model

According to a comprehensive literature review, the issue of different contracts for care providers has not been addressed. But in real-world shift scheduling, different sources of uncertainties need to be addressed to provide a high-quality schedule. In this study, the subject of the uncertainty of patient demand is considered. To fill this research gap, in this study, the uncertainty of patients' demand and the types of service providers' contracts and skills are used as a basis to develop a two-stage stochastic programming model. In the following, the model is solved with the sample average approximation method, and the parameters of the solution method are adjusted. In the end, the obtained planning validity is shown using the simulation method.

The rest of this article is as follows. Section 3 presents the proposed optimization model and describes its structure. The solution approach is introduced in Section 4, and detailed descriptions are provided for the SAA method too. Section 5 presents the statistical experiments. Finally, the concluding remarks are made in Section 6.

3. Problem definition

Hospitals and health centers should provide the necessary services to patients in common and critical situations such as Covid-19. In recent years, due to the decrease in available care providers and the increase in diseases, the tendency for cooperation between health center managers and

researchers to properly plan appropriate services for patients has increased. Therefore, one of the most important issues is the proper distribution of care providers between work shifts. On the other hand, it is not possible to accurately determine the demand for each skill in many cases. Therefore, a two-stage stochastic planning model is proposed to achieve high-quality planning. Some of the assumptions of the problem are as follows:

1. All care providers have identical skills
2. Demand behavior is the random variable based on a specific distribution function
3. Each care provider is only assigned one shift each day
4. Each care provider has a specific contract

After solving the proposed model, the work plan obtained for a five-day horizon can be as follows (Table 1):

Table 1-An example of care providers' schedule

	Saturday	Sunday	Monday	Tuesday	Wednesday
Nurse 1	Morning	Afternoon	Morning	Afternoon	Afternoon
General practitioner 1	Afternoon	Morning	-	Morning	Afternoon
Specialist physician 1	Afternoon	Afternoon	Afternoon	-	Morning

In the proposed mathematical model, care providers are assigned to certain shifts, and the number of overtime hours required in possible conditions is determined. The required duration of each skill per day and each shift (de_{smd}) is stochastic. Care providers have three professions: nurses, general practitioners, and specialist physicians. Contracts are also available in three types, full-time, part-time, and hourly, containing 8 hours, 4 hours, and 2 hours per shift, respectively. The stochastic demand model for the problem of scheduling care providers can be formulated with the notations as follows:

Sets

- S: Set of skills (xx: nurse, xy: general practitioner, xz: specialist physician)
M: Set of shifts
D: Set of days
N: Set of contracts (full time, part-time, hourly)
 ξ : Set of scenarios ($\xi = 1, 2, \dots, B$)
 I_{xx} : Set of nurses
 I_{xy} : Set of general practitioners
 I_{xz} : Set of specialist physicians

Parameters

- aa_{ij} : 1, if nurse i is under contract j
 ab_{ij} : 1, if general practitioner i is under contract j
 ac_{ij} : 1, if specialist physician i is under contract j
 h_j : Number of the hours of service by contract j per shift

- hh_j : Number of contract hours j per month
 de_{smd} : Number of the hours required of skill s per shift m per day d
 ca_j : Cost of the nurse service with contract j per hour
 cb_j : Cost of the general practitioner with contract j per hour
 cc_j : Cost of the specialist physician with contract j per hour
 c_i : Additional service cost per hour for skill i ($i=xx, xy, xz$)
 e : Minimum number of the shifts for a full-time care provider

Variables

- xx_{imd} : One, if nurse i is set for shift m on day d ; otherwise, 0
 xy_{imd} : One, if general practitioner i is assigned to shift m on day d ; otherwise, 0
 xz_{imd} : One, if specialist i is assigned to shift m on day d ; otherwise, 0
 p_{imd} : Number of the additional hours required for skill i on shift m per day d

$$\begin{aligned}
 \text{Min } Z = & \sum_{d=1}^D \sum_{m=1}^M \sum_{i=1}^{I_{xx}} \sum_{j=1}^N h_j ca_j aa_{ij} xx_{imd} + \sum_{d=1}^D \sum_{m=1}^M \sum_{i=1}^{I_{xy}} \sum_{j=1}^N h_j cb_j ab_{ij} xy_{imd} \\
 & + \sum_{d=1}^D \sum_{m=1}^M \sum_{i=1}^{I_{xz}} \sum_{j=1}^N h_j cc_j ac_{ij} xz_{imd} + \sum_{\xi \in B} \sum_{d=1}^D \sum_{m=1}^M \sum_{i \in S} \phi(\xi) c_i p_{imd}^{\xi}
 \end{aligned} \tag{1}$$

$$\sum_{m=1}^M xx_{imd} \leq 1 \quad i = 1, \dots, I_{xx}, d = 1, \dots, D \tag{2}$$

$$\sum_{m=1}^M xy_{imd} \leq 1 \quad i = 1, \dots, I_{xy}, d = 1, \dots, D \tag{3}$$

$$\sum_{m=1}^M xz_{imd} \leq 1 \quad i = 1, \dots, I_{xz}, d = 1, \dots, D \tag{4}$$

$$\sum_{d=1}^D \sum_{m=1}^M aa_{ij} xx_{imd} \geq e \quad j = 1, i = 1, \dots, I_{xx} \tag{5}$$

$$\sum_{d=1}^D \sum_{m=1}^M ab_{ij} xy_{imd} \geq e \quad j = 1, i = 1, \dots, I_{xy} \tag{6}$$

$$p_{(xx)md}^{\xi} = \left(de_{(xx)md}^{\xi} - \sum_{j=1}^N \sum_{i=1}^{I_{xx}} h_j aa_{ij} xx_{imd} \right) f_{(xx)md}^{\xi} \quad m = 1, 2, d = 1, \dots, D \tag{7}$$

$$de_{(xx)md} \leq \sum_{j=1}^N \sum_{i=1}^{I_{xx}} h_j aa_{ij} xx_{imd} + M f_{(xx)md}^{\xi} \quad m = 1, 2, d = 1, \dots, D \tag{8}$$

$$de_{(xx)md}^{\xi} \geq \sum_{j=1}^N \sum_{i=1}^{I_{xx}} h_j aa_{ij} xx_{imd} - (1 - f_{(xx)md}^{\xi}) M \quad m = 1, 2, d = 1, \dots, D \tag{9}$$

$$p_{(xy)md}^{\xi} = \left(de_{(xy)md}^{\xi} - \sum_{j=1}^N \sum_{i=1}^{I_{xy}} h_j ab_{ij} xy_{imd} \right) f_{(xy)md}^{\xi} \quad m = 1,2, d = 1, \dots, D \quad (10)$$

$$de_{(xy)md}^{\xi} \leq \sum_{j=1}^N \sum_{i=1}^{I_{xy}} h_j ab_{ij} xy_{imd} + M f_{(xy)md}^{\xi} \quad m = 1,2, d = 1, \dots, D \quad (11)$$

$$de_{(xy)md}^{\xi} \geq \sum_{j=1}^N \sum_{i=1}^{I_{xy}} h_j ab_{ij} xy_{imd} - (1 - f_{(xy)md}^{\xi}) M \quad m = 1,2, d = 1, \dots, D \quad (12)$$

$$p_{(xz)md}^{\xi} = \left(de_{(xz)md}^{\xi} - \sum_{j=1}^N \sum_{i=1}^{I_{xz}} h_j ab_{ij} xz_{imd} \right) f_{(xz)md}^{\xi} \quad m = 1,2, d = 1, \dots, D \quad (13)$$

$$de_{(xz)md}^{\xi} \leq \sum_{j=1}^N \sum_{i=1}^{I_{xz}} h_j ab_{ij} xz_{imd} + M f_{(xz)md}^{\xi} \quad m = 1,2, d = 1, \dots, D \quad (14)$$

$$de_{(xz)md}^{\xi} \geq \sum_{j=1}^N \sum_{i=1}^{I_{xz}} h_j ab_{ij} xz_{imd} - (1 - f_{(xz)md}^{\xi}) M \quad m = 1,2, d = 1, \dots, D \quad (15)$$

$$xx_{imd} \in \{0,1\} \quad i = 1, \dots, I_{xx}, m = 1,2, d = 1, \dots, D \quad (16)$$

$$xy_{imd} \in \{0,1\} \quad i = 1, \dots, I_{xy}, m = 1,2, d = 1, \dots, D \quad (17)$$

$$xz_{imd} \in \{0,1\} \quad i = 1, \dots, I_{xz}, m = 1,2, d = 1, \dots, D \quad (18)$$

$$p_{imd} \geq 0 \quad i \in S \quad i = xx, xy, xz, m = 1,2, d = 1, \dots, D \quad (19)$$

The objective function minimizes regular work hours and overtime hours. In this regard, $\phi(\xi)$ is the probability of scenario $\xi = 1,2, \dots, B$ and $\sum_{\xi \in B} \phi(\xi) = 1$. Constraints (2), (3) and (4) ensure that each nurse, general practitioner and specialist physician is assigned to maximally one shift a day. Constraints (5) and (6) ensure that the full-time nurses and general practitioners must work on at least e shifts. Constraints (7), (8) and (9) specify the number of the overtime hours of nursing skill per shift. Constraints (10), (11) and (12) specify the number of the overtime hours for specialist physicians per shift. Constraints (13), (14) and (15) specify the number of overtime hours of nursing skill per shift. Eventually, constraints (16) to (19) define the model's variables.

This research assumes that the required number of hours of skill s on shift m per day (de_{smd}) has a discrete uniform distribution in the interval (a, b) . An exact solution can be obtained for small-size problems, but, as the size of the problem increases, the solution time increases too. This study solves the problem with the sample average approximation (SAA) algorithm. A recourse action model is applied to formulate the model of solving the problem with that algorithm. Section 4 delineates the basic features of the new model [23].

Programming with the stochastic integer recourse model

Stochastic programming models have appeared as extensions of optimization problems with random parameters. Consider the optimization problem below[23]:

$$\begin{aligned} \min \quad & cx \\ \text{s. t.} \quad & Ax = b \\ & Tx = h \\ & x \in X \end{aligned}$$

where $X \subset \mathbb{R}^n$ indicates the non-negativity of the decision variable x and possibly the integrality constraints on it. In addition to m_1 deterministic constraints of $Ax = b$, there is a set of m constraints of $Tx = h$, where the parameters T and h depend on information and become available only after a decision is made on x . A class of stochastic programming models, known as recourse models, is obtained by allowing additional or recourse decisions after realizing the random variables T and h . So, recourse models are dynamic; the stages model the time discretely on the basis of the existing data. If the uncertainty is all dissolved simultaneously, it can be captured by a recourse model two stages, present and future. Considering a first-stage decision x , the infeasibility of $h - Tx$ for every possible (q, T, h) is compensated with minimum costs, while second-stage decisions are made as an optimal solution of the second-stage problem. This specifies the minimal recourse costs as a function of the first-stage decision x , and the realization of ξ is denoted by $v(x, \xi)$. Its expectation, $Q(x) = \mathbb{E}_\xi[v(x, \xi)]$, yields the expected recourse costs associated with the first-stage decision x . Thus, the two-stage recourse model is:

$$\begin{aligned} \min \quad & cx + Q(x) \\ \text{s. t.} \quad & Ax = b \\ & x \in X \end{aligned}$$

Where the objective function $cx + Q(x)$ specifies the total expected costs of decision x [23]. The stochastic demand model of scheduling care providers can be expressed as the following recourse model.

$$\begin{aligned} \text{Min } Z = & \sum_{d=1}^D \sum_{m=1}^M \sum_{i=1}^{I_{xx}} \sum_{j=1}^N h_j ca_j aa_{ij} xx_{imd} + \sum_{d=1}^D \sum_{m=1}^M \sum_{i=1}^{I_{xy}} \sum_{j=1}^N h_j cb_j ab_{ij} xy_{imd} \\ & + \sum_{d=1}^D \sum_{m=1}^M \sum_{i=1}^{I_{xz}} \sum_{j=1}^N h_j cc_j ac_{ij} xz_{imd} + E[Q(x, \xi)] \\ \text{s. t.} \quad & (1) - (5) \end{aligned}$$

where $E[Q(x, \xi)]$ is the recourse action function, and

$$\begin{aligned} Q(x, \xi) = \min \quad & \sum_{\xi \in B} \sum_{d=1}^D \sum_{m=1}^M \sum_{i \in S} \phi(\xi) c_i p_{imd}^\xi \\ \text{s. t.} \quad & (6) - (16) \end{aligned}$$

The $\xi \in B$ vector contains numerous scenarios. So, to obtain $E[Q(x, \xi)]$, lots of similar integer linear programs (ILPs)[24] must be solved, where is a difficult calculation task. Since it is hard to provide an exact solution to the proposed model, the next section proposes an approximation.

4. Sample Average Approximation (SAA)

There are several solution methods, such as SAA, to solve stochastic models. The SAA method is a Monte Carlo simulation-based method that solves stochastic programming problems by generating random samples and approximating the expected function values through the average functions of the corresponding samples. The stop criterion determines how long the algorithm will last. Over the years, various authors have used the idea of sample average approximation to solve stochastic programs. For example, it was employed to solve stochastic knapsack problems [25], stochastic routing problems [26], supply chain problems [27] and investment problems [28]. Due to the high applicability of the SAA method, it has been selected to solve the model in this study. The method is delineated below.

Suppose M is the number of replications, N is the number of scenarios in the sample problem, and N' denotes the sample size used to estimate $C^T \hat{x} + E[Q(\hat{X}, \xi)]$ for a given feasible solution \hat{X} . So, the SAA method can be described as follows [26]:

1. Repeat the following steps for $m = 1, \dots, M$:

- a) Generate $\xi^1, \xi^2, \dots, \xi^N$ as an N random sample
- b) Solve the problem by SAA method and take \hat{X}_N^m as a solution vector and \hat{Z}_N^m as an optimal objective value.
- c) Generate $\xi^1, \xi^2, \dots, \xi^{N'}$ as an independent random sample. Evaluate $\hat{g}_{N'}(\hat{X}_N^m)$ and $S_{\hat{g}_{N'}(\hat{X}_N^m)}^2$ as follows:

$$\begin{aligned} \hat{g}_{N'}(\hat{X}_N^m) &= \sum_{d=1}^D \sum_{m=1}^M \sum_{i=1}^{I_{xx}} \sum_{j=1}^N h_j c a_j a a_{ij} x x_{imd} + \sum_{d=1}^D \sum_{m=1}^M \sum_{i=1}^{I_{xy}} \sum_{j=1}^N h_j c b_j a b_{ij} x y_{imd} \\ &\quad + \sum_{d=1}^D \sum_{m=1}^M \sum_{i=1}^{I_{xz}} \sum_{j=1}^N h_j c c_j a c_{ij} x z_{imd} + \frac{1}{N'} \sum_{n=1}^{N'} \sum_{d=1}^D \sum_{m=1}^M \sum_{i \in S} \phi(\xi) c_i p_{imd}^\xi \\ S_{\hat{g}_{N'}(\hat{X}_N^m)}^2 &= \frac{1}{N'(N'-1)} \sum_{n=1}^{N'} \left[\sum_{d=1}^D \sum_{m=1}^M \sum_{i=1}^{I_{xx}} \sum_{j=1}^N h_j c a_j a a_{ij} x x_{imd} \right. \\ &\quad + \sum_{d=1}^D \sum_{m=1}^M \sum_{i=1}^{I_{xy}} \sum_{j=1}^N h_j c b_j a b_{ij} x y_{imd} + \sum_{d=1}^D \sum_{m=1}^M \sum_{i=1}^{I_{xz}} \sum_{j=1}^N h_j c c_j a c_{ij} x z_{imd} \\ &\quad \left. + \sum_{d=1}^D \sum_{m=1}^M \sum_{i \in S} \phi(\xi) c_i p_{imd} - \hat{g}_{N'}(\bar{X}) \right] \end{aligned}$$

2. Evaluate \bar{Z}_N^M and $S_{\bar{Z}_N^M}^2$.

$$\bar{Z}_N^M = \frac{1}{M} \sum_{m=1}^M \hat{Z}_N^m, S_{\bar{Z}_N^M}^2 = \frac{1}{M(M-1)} \sum_{m=1}^M [\hat{Z}_N^m - \bar{Z}_N^M]^2$$

The following formula serves to calculate the confidence interval for the optimality gap:

$$\hat{g}_{N'}(\hat{X}_N^m) - \bar{Z}_N^M + Z_\alpha \left\{ S_{\hat{g}_{N'}(\hat{X}_N^m)}^2 + S_{\bar{Z}_N^M}^2 \right\}^{0.5}$$

Here is $Z_\alpha = \Phi^{-1}(1 - \alpha)$, in which $\Phi(Z)$ stands for the cumulative pattern of the standard normal distribution.

3. In the case of each solution \hat{X}_N^m , the parameter $m = 1, \dots, M$ determines the optimality gap with $\hat{g}_{N'}(\hat{X}_N^m) - \bar{Z}_N^M$ along with an estimated variance of $S_{\hat{g}_{N'}(\hat{X}_N^m)}^2 + S_{\bar{Z}_N^M}^2$. One of the M candidate solutions is selected based on the least estimated objective value.

In the algorithm, \bar{Z}_N^M and $\hat{g}_{N'}(\hat{X}_N^m)$ are the lower and upper bounds of the optimal value, respectively [29]. The parameter \bar{Z}_N^M shows an unbiased estimator of the optimal objective function $E(\hat{Z}_N)$. Here, $\bar{Z}_N^M = E(\hat{Z}_N)$ and $E(\hat{Z}_N) \leq Z^*$. Moreover, $\hat{g}_{N'}(\hat{X}_N^m)$ presents an unbiased estimator of the objective value $E(\hat{Z}_N)$, but $E(\hat{g}_{N'}(\hat{X}_N^m)) \geq Z^*$.

5. Numerical results

5.1. Case study

This section reports a case study planned at the Iranian Health Control Center. There were nine nurses, seven general practitioners, and three specialist physicians with 24 working days divided into two shifts, morning and afternoon. The corresponding data were obtained from the Iranian Health Control Center to evaluate the distribution of the demand for each skill each day and each shift. The demand had a uniform distribution in the intervals (24, 36), (12, 18), and (5, 9) per hour for the nurse, general practitioner, and specialist physician in each shift, respectively.

Ta shows the cost per contract hour for different skills. The cost of each additional hour of service for a nurse, general practitioner, and specialist physician is 90, 160, and 240 dollars, respectively.

Table 3. The wage of each skill per hour (\$)

	Nurse	General practitioner	Specialist physician
Full time	50	60	110
Part-time	60	110	150
Hourly	70	150	200

The proposed approach was put to practice in Python, employing a GUROBI optimization solver (<http://www.gurobi.com/>) on a mac book pro with an 8-core CPU and an 8-GB RAM.

5.2. Numerical results

This section is dedicated to the experimental results achieved from implementing our approach in the Iranian Health Control Center. In this approach, the N quantities of 1, 20, 50, and 100 and the M quantities of 10 and $N' = 20000$ for the SAA algorithm presented in Section 3 and the Monte Carlo method were used to generate random numbers. Table 2-7 show the results.

Column $\hat{g}_{N'}(\bar{X})$ shows the value of the objective function based on N' independent random sample, and column $S_{\hat{g}_{N'}(\bar{X})}^2$ indicates the value of variance. Column \hat{Z}_N^m specifies the optimal value of the objective function considering the N scenario. The gap column shows $\hat{g}_{N'}(\bar{X}) - \bar{Z}_N^M$, and *Var* column refers to $S_{\hat{g}_{N'}(\bar{X})}^2 + S_{\bar{Z}_N^M}^2$.

As the value of N was increased from 1 to 20, 50, and 100, the variance of the estimates changed from 222039.63 to 56973.116, 10280.988 and 3449.325; there was a decrease of 99.97% in total. Moreover, with the increase of N, the mean values of $\hat{g}_{N'}(\bar{X})$ and \bar{Z}_N^M began to converge to make an optimal solution. The mean of $\sum_{i=1}^M \hat{g}_{N'}^i(\bar{X}) - \bar{Z}_N^M$ decreased from 1097677 to 25116 for N values of 1 to 100. The total reduction was 97.8%. These values showed a convergence for an optimal solution.

Table 2. Simple Monte Carlo method: N = 1, M = 10 and N' = 20000

m	$\hat{g}_{N'}(\bar{X})$	$S_{\hat{g}_{N'}(\bar{X})}^2$	\hat{Z}_N^m	Gap	Var
1	198972.3	286.3	184279	13881.35	222060.9
2	195957.3	263.64	185240	10866.42	222038.2
3	195868.2	297.41	185570	10777.28	222072
4	197264	243.38	188200	12173.06	222018
5	196737.7	253.38	183060	11646.75	222028
6	195290.8	259.71	184470	10199.94	222034.3
7	193692.3	246.14	185130	8601.37	222020.7
8	195807.1	268.63	186560	10716.22	222043.2
9	196799.7	274.34	183450	11708.76	222048.9
10	194287.3	257.48	184950	9196.35	222032.1
$\bar{Z}_N^M = 185090.9$ $S_{\bar{Z}_N^M}^2 = 221774.59$					

Table 3. Simple Monte Carlo method: N = 20, M = 10 and N' = 20000

	$\hat{g}_{N'}(\bar{X})$	$S_{\hat{g}_{N'}(\bar{X})}^2$	\hat{Z}_N^m	Gap	Var
1	189620.6	260.45	188152.5	953.4	56977.13
2	189600.5	257.26	189520	933.273	56973.94
3	189591.9	255.13	188311.5	924.7	56971.81
4	189544	260.26	187329.5	876.76	56976.94
5	189592	258.66	188913.5	924.83	56975.34
6	189613.1	259.02	189344.5	945.89	56975.7
7	189583.7	258.92	187789.5	916.47	56975.6

8	189544.2	253.49	188928	877.02	56970.17
9	189689.9	252.14	189571.5	1022.66	56968.82
10	189642.1	249.03	188811.5	974.85	56965.71
$\bar{Z}_N^M = 188667.2$ $S_{\bar{Z}_N^M}^2 = 56716.68$					

Table 4. Simple Monte Carlo method: N = 50, M = 10 and N' = 20000

	$\hat{g}_{N'}(\bar{X})$	$S_{\hat{g}_{N'}(\bar{X})}^2$	\hat{Z}_N^m	Gap	Var
1	189523.9	258.39	189129.4	515.29	10281.55
2	189554.4	257.36	189046.4	545.8	10280.52
3	189502.4	263.52	188939.8	493.75	10286.68
4	189510.6	254.97	188681.8	501.98	10278.13
5	189529.3	257.63	188575	520.65	10280.79
6	189529.2	259.21	189338.8	520.61	10282.37
7	189543.1	254.34	188856.2	534.51	10277.5
8	189519.1	259.17	189368	510.43	10282.33
9	189488.9	258.57	188677.8	480.26	10281.73
10	189546.5	255.12	189473	537.86	10278.28
$\bar{Z}_N^M = 189008.62$ $S_{\bar{Z}_N^M}^2 = 10023.16$					

Table 5. Simple Monte Carlo method: N = 100, M = 10 and N' = 20000

	$\hat{g}_{N'}(\bar{X})$	$S_{\hat{g}_{N'}(\bar{X})}^2$	\hat{Z}_N^m	Gap	Var
1	189520.5	261.47	189356.8	243.11	3454.56
2	189500.7	255.04	189063.1	223.32	3448.13
3	189524.1	252.5	189333.3	246.72	3445.59
4	189565.9	254.14	189285.6	288.46	3447.23
5	189500.6	258.71	189063.4	223.16	3451.8
6	189551.9	251.47	188997.8	274.48	3444.56
7	189526.2	256.26	189471.2	248.76	3449.35
8	189527	257.76	189517.6	249.65	3450.85
9	189543.4	258.78	189388.6	266.04	3451.87
10	189525.2	256.22	189296.5	247.8	3449.31
$\bar{Z}_N^M = 189277.39$ $S_{\bar{Z}_N^M}^2 = 3193.09$					

The next step generated uncertain parameters with the Latin Hypercube Sampling Scheme [30] to reduce the SAA method's variance. This scheme is a stratified random sampling method by which samples are selected from many variables so that the sample for each variable has the highest degree of classification. As shown in Table 6, with the Latin Hypercube Sampling Scheme, the variance at the lowest point (N = 100) in the Monte Carlo simulation decreased from

3449.325 to 141.531, which means a reduction of 96%. Figure 1 and Figureshow the box plot for the Monte Carlo sampling method and the Latin Hypercube Sampling Scheme, respectively.

Table 6. LHS method (N = 100, M = 10 and N' = 20000)

	$\hat{g}_{N'}(\bar{X})$	$S_{\hat{g}_{N'}(\bar{X})}^2$	\hat{Z}_N^m	Gap	Var
1	189582.3	14.17	189340.8	273.01	143.57
2	189577.4	14.36	189308.1	268.12	143.76
3	189582.2	14.07	189332.8	272.93	143.47
4	189580.7	14.09	189281.3	271.4	143.49
5	189578.6	13.89	189365.4	269.33	143.29
6	189576.2	14.01	189288.6	266.88	143.41
7	189578.7	14.31	189247.3	269.45	143.71
8	189582.9	14.3	189275	273.61	143.7
9	189579	13.96	189319.3	269.68	143.36
10	189577.5	14.15	189334.2	268.18	143.55
$\bar{Z}_N^M = 189309.28$					
$S_{\bar{Z}_N^M}^2 = 129.4$					

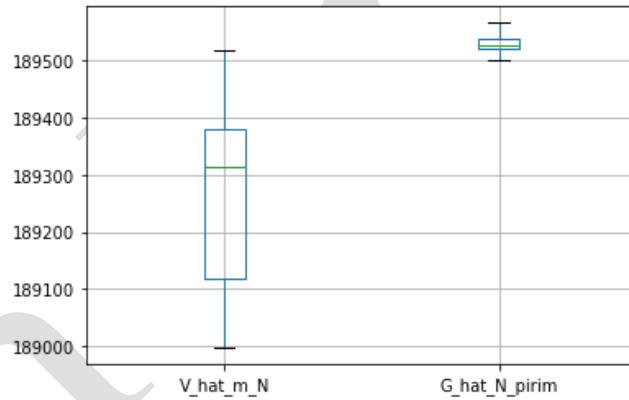


Figure 1. Simple Monte Carlo method (N = 100, M = 10 and N' = 20000)

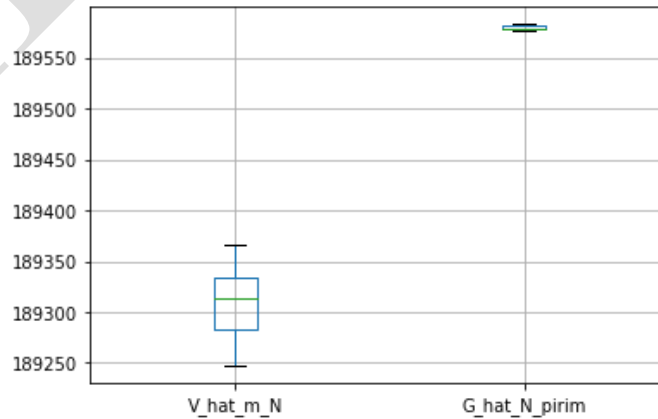


Figure 2. LHS method (N = 100, M = 10 and N' = 20000)

Any significant difference between the means of the Gap of Monte Carlo and the Latin Hypercube Sampling methods was checked with the Mann-Whitney test. As the test was performed with α of 0.05, the null hypothesis was rejected (p -value = 0.013). Therefore, a significant difference was detected between the means of those two methods. Since the Latin Hypercube Sampling method reduced the variance of Var to 96%, the least value of the objective function of this method was selected. Based on the results, 189247.3 was the lowest value of the objective function and, thus, was selected as the best answer. Table 7-Table 9 show the shift plans of nurses, general practitioners and specialist physicians determined based on the best response obtained. In these tables, morning or afternoon shifts are assigned to each care provider; otherwise, that provider would not be busy. The letters M and A in the tables stand for the morning and the afternoon shifts, respectively.

Morning shift



Afternoon Shift



Table 7-Nurses' work schedule

	days																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	A	M	M	M	A	A	A	A	A	M	A	A	A	A	A	A	A	A	M	M	A	M	M	A
2	M	M	M	M	M	A	A	A	M	M	M	A	M	A	A	M	A	M	M	A	M	A	M	A
3	M	M	A	M	M	M	M	M	M	A	M	M	A	M	M	M	A	M	A	M	M	M	A	A
4	A	A	M	A	A	M	M	M	M	A	A	A	A	M	M	A	M	A	M	A	A	A	M	M
5	A	A	A	A	A	A	M	A	A	A	M	M	M	A	M	M	M	A	A	M	A	M	A	M
6	M	A	A	A	M	A	A	M	A	A	A	M	M	M	A	A	M	M	A	A	M	M	A	M
7	A	A	A	A	M	M	M	A	A	M	A	M	M	A	A	A	M	A	A	A	M	A	A	M
8	M	A	A	A	A	M	A	M	A	M	M	M	M	M	M	M	M	M	A	M	A	A	A	M
9	M	A	M	A	M	A	M	A	A	M	M	A	A	A	M	A	M	M	M	M	A	A	M	M

Table 8-General practitioners' work schedule

	days																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	A	M	M	A	M	A	A	A	A	A	A	A	A	A	A	A	A	M	M	M	A	A	M	A
2	M	A	A	M	A	M	M	M	M	M	M	M	M	M	M	M	M	A	A	A	M	M	M	M
3	A	M	A	A	A	A	M	A	A	M	M	A	A	M	A	A	M	A	M	A	A	M	A	A
4	-	-	-	-	-	-	-	-	-	A	-	-	M	A	M	M	A	-	-	M	-	A	A	M
5	-	A	M	-	-	M	A	M	M	-	-	-	-	-	-	-	-	-	A	-	M	-	-	-
6	M	-	-	M	M	-	-	-	-	-	A	M	-	-	-	-	-	M	-	-	-	-	-	-
7	M	A	M	M	M	M	A	M	M	A	A	M	M	A	M	M	A	M	A	M	M	A	A	M

Table 9-Specialist physicians' work schedule

	days																								
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
1	M	M	M	M	-	M	-	M	M	-	M	M	M	-	M	A	M	M	M	M	M	-	-	M	M
2	A	A	A	A	M	A	M	A	A	M	A	A	A	M	A	-	A	A	A	A	M	A	A	A	A
3	-	-	-	-	A	-	A	-	-	A	-	-	-	A	-	M	-	-	-	-	A	M	-	-	-

After the service planning was conducted, the Monte Carlo simulation method validated the model. As many as 1000 problems were generated randomly, and calculations were performed to obtain the values of their objective functions. The histogram in Figure 2 presents the obtained values. As the results revealed, the average value of the total cost was 191521, which shows a 1% difference from the best solution obtained through solving the model by the SAA method. The worst value of the objective function obtained from the simulation method was 1974480, which shows a 4% difference from the best result obtained by solving the SAA model model.

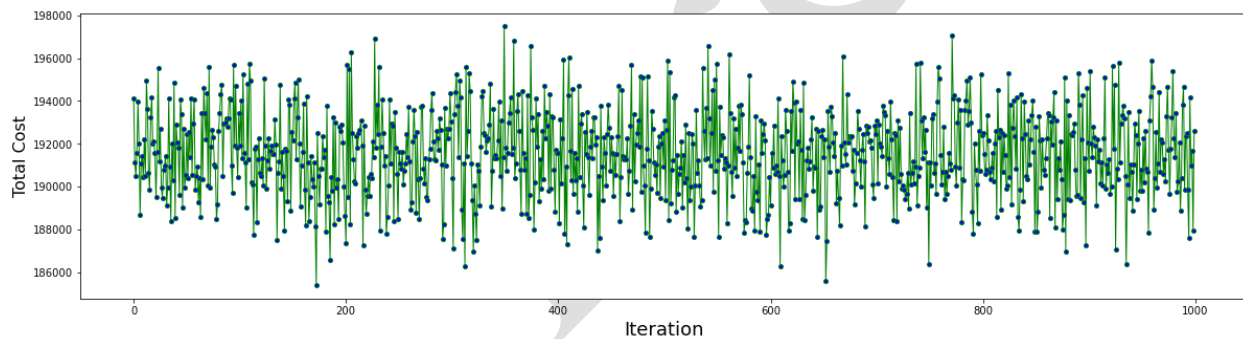


Figure 2. Simulation results

5.3. Managerial implications

In addition to the significant cost reduction resulting from more efficient shift scheduling, the daily use of shift schedules has important managerial implications for the workload of home care, hospital administrators and care providers. This schedule can free the care providers to deal with other tasks that require closer interactions with patients. Moreover, setting a shift schedule makes it possible to hold training courses and update the care providers in their free time. Another advantage of this planning is to provide a robust program against changes in patient demand. The continuous use of the planning can be beneficial for patients. If they ever face a shortage of care providers, the necessary predictions have already been made in a schedule. To provide better plans in this case, continual cooperation between the Iranian Health Control Center and universities seems necessary. Hospitals and other home care centers can also do planning with more constraints, if required.

6. Conclusion

In real-world shift scheduling, care providers cooperate with various practitioners, and work by different contracts in home care centers under different sources of uncertainty, such as patient demand. Among the patients for whom demand uncertainty may occur are cancer patients, who experience unpredictable conditions during their illness. This study addresses the issue, and the demand distribution is estimated based on the available data. To this end, a two-stage stochastic programming model is presented, and the problem is solved with the SAA method. Based on the results of implementing the Latin Hypercube Sampling Scheme, the upper and lower bounds of the optimal solution begin to converge when the sample size (N) is increased. This also leads to the reduction of the variance of the solutions. The solution method has proved to be efficient by using the Monte Carlo simulation method. In addition, numerical results have been used for the shift planning of care providers.

There are several recommendations to make for future research in nurse scheduling. The specific skills needed by individual patients can be considered a basis for assigning relevant care providers. Moreover, human factors involved in care providing seem interesting topics to study. Other methods can also be tried to solve stochastic programming models. Finally, models may be developed in other areas such as fire stations and emergency centers where shift planning is needed.

References

- [1] Pascal. Zurn, Carmen. Dolea, Barbara. Stilwell, World Health Organization. Department of Human Resources for Health., and International Council of Nurses., *Nurse retention and recruitment : developing a motivated workforce*. International Council of Nurses, 2005.
- [2] S. H. Katholieke, H. Vives, P. de Causmaecker, K. U. Leuven, R. Lagatie, and S. Haspelslagh, "Negotiation Protocols for Distributed Nurse Rostering," 2009. [Online]. Available: <https://www.researchgate.net/publication/239927115>
- [3] P. de Causmaecker and G. vanden Berghe, "A categorisation of nurse rostering problems," in *Journal of Scheduling*, Feb. 2011, vol. 14, no. 1, pp. 3–16. doi: 10.1007/s10951-010-0211-z.
- [4] A. T. Ernst, H. Jiang, M. Krishnamoorthy, and D. Sier, "Staff scheduling and rostering: A review of applications, methods and models," in *European Journal of Operational Research*, Feb. 2004, vol. 153, no. 1, pp. 3–27. doi: 10.1016/S0377-2217(03)00095-X.
- [5] B. Klinz, U. Pferschy, and J. Schauer, "ILP Models for a Nurse Scheduling Problem," in *Operations Research Proceedings 2006*, Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 319–324. doi: 10.1007/978-3-540-69995-8_52.
- [6] H. G. Okuno, M. Ali, S. Topaloglu, and H. Selim, "Nurse Scheduling Using Fuzzy Multiple Objective Programming," 2007. doi: 10.1016/j.fss.2009.10.003.
- [7] S. Topaloglu and H. Selim, "Nurse scheduling using fuzzy modeling approach," *Fuzzy Sets and Systems*, vol. 161, no. 11, pp. 1543–1563, Jun. 2010, doi: 10.1016/j.fss.2009.10.003.
- [8] D. Landa-silva and K. N. Le, "A Simple Evolutionary Algorithm with Self-adaptation for Multi-objective Nurse Scheduling," in *Adaptive and Multilevel Metaheuristics*, Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 133–155. doi: 10.1007/978-3-540-79438-7_7.

- [9] M. Ohki, "Effective operators using parallel processing for nurse scheduling by cooperative genetic algorithm," 2012. Accessed: Feb. 13, 2022. [Online]. Available: <https://www.inderscienceonline.com/doi/abs/10.1504/IJDMMM.2012.045136?journalCode=ijdmmm>
- [10] A. A. el Adoly, M. Gheith, and M. Nashat Fors, "A new formulation and solution for the nurse scheduling problem: A case study in Egypt," *Alexandria Engineering Journal*, vol. 57, no. 4, pp. 2289–2298, Dec. 2018, doi: 10.1016/j.aej.2017.09.007.
- [11] B. Maenhout and M. Vanhoucke, "An integrated nurse staffing and scheduling analysis for longer-term nursing staff allocation problems," *Omega (United Kingdom)*, vol. 41, no. 2, pp. 485–499, Apr. 2013, doi: 10.1016/j.omega.2012.01.002.
- [12] H. G. Santos, T. A. M. Toffolo, R. A. M. Gomes, and S. Ribas, "Integer programming techniques for the nurse rostering problem," *Annals of Operations Research*, vol. 239, no. 1, pp. 225–251, Apr. 2016, doi: 10.1007/s10479-014-1594-6.
- [13] J. Ingels and B. Maenhout, "The impact of reserve duties on the robustness of a personnel shift roster: An empirical investigation," *Computers and Operations Research*, vol. 61, pp. 153–169, Sep. 2015, doi: 10.1016/j.cor.2015.03.010.
- [14] M. Bagheri, A. Gholinejad Devin, and A. Izanloo, "An application of stochastic programming method for nurse scheduling problem in real word hospital," *Computers and Industrial Engineering*, vol. 96, pp. 192–200, 2016, doi: 10.1016/j.cie.2016.02.023.
- [15] P. Punnakitikashem, J. M. Rosenberber, and D. F. Buckley-Behan, "A stochastic programming approach for integrated nurse staffing and assignment," *IIE Transactions (Institute of Industrial Engineers)*, vol. 45, no. 10, pp. 1059–1076, Oct. 01, 2013. doi: 10.1080/0740817X.2012.763002.
- [16] P. S. Chen, Y. J. Lin, and N. C. Peng, "A two-stage method to determine the allocation and scheduling of medical staff in uncertain environments," *Computers and Industrial Engineering*, vol. 99, pp. 174–188, Sep. 2016, doi: 10.1016/j.cie.2016.07.018.
- [17] B. Y. Ang, S. W. S. Lam, Y. Pasupathy, and M. E. H. Ong, "Nurse workforce scheduling in the emergency department: A sequential decision support system considering multiple objectives," *Journal of Nursing Management*, vol. 26, no. 4, pp. 432–441, May 2018, doi: 10.1111/jonm.12560.
- [18] M. Hamid, R. Tavakkoli-Moghaddam, F. Golpaygani, and B. Vahedi-Nouri, "A multi-objective model for a nurse scheduling problem by emphasizing human factors," *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, vol. 234, no. 2, pp. 179–199, Feb. 2020, doi: 10.1177/0954411919889560.
- [19] M. R. Hassani and J. Behnamian, "A scenario-based robust optimization with a pessimistic approach for nurse rostering problem," *Journal of Combinatorial Optimization*, vol. 41, no. 1, pp. 143–169, Jan. 2021, doi: 10.1007/s10878-020-00667-0.
- [20] A. Kheiri, A. Gretsista, E. Keedwell, G. Lulli, M. G. Epitropakis, and E. K. Burke, "A hyper-heuristic approach based upon a hidden Markov model for the multi-stage nurse rostering problem," *Computers and Operations Research*, vol. 130, Jun. 2021, doi: 10.1016/j.cor.2021.105221.
- [21] H. G. Okuno, M. Ali, S. Topaloglu, and H. Selim, "Nurse Scheduling Using Fuzzy Multiple Objective Programming," 2007.
- [22] Z. Zhang, Z. Hao, and H. Huang, "Hybrid Swarm-Based Optimization Algorithm of GA & VNS for Nurse Scheduling Problem," in *LNCS*, vol. 7030, 2011, pp. 375–382. doi: 10.1007/978-3-642-25255-6_48.

- [23] L. Stougie and M. H. van der Vlerk, "Approximation in Stochastic Integer Programming," 2005. doi: DOI 10.18452/8294.
- [24] T. v Mikosch, S. I. Resnick, S. M. Robinson, J. R. Birge, and F. Louveaux, "Springer Series in Operations Research and Financial Engineering Introduction to Stochastic Programming." [Online]. Available: <http://www.springer.com/series/3182>
- [25] A. J. Kleywegt, A. Shapiro, T. Homem-de-mello, and S. J. Optim, "THE SAMPLE AVERAGE APPROXIMATION METHOD FOR STOCHASTIC DISCRETE OPTIMIZATION *," 2001. doi: <https://doi.org/10.1137/S1052623499363220>.
- [26] B. Verweij, S. Ahmed, A. Kleywegt, G. Nemhauser, and A. Shapiro, "The Sample Average Approximation Method Applied to Stochastic Routing Problems: A Computational Study," 2001. doi: <https://doi.org/10.1023/A:1021814225969>.
- [27] T. Santoso, S. Ahmed, M. Goetschalckx, and A. Shapiro, "A stochastic programming approach for supply chain network design under uncertainty," *European Journal of Operational Research*, vol. 167, no. 1, pp. 96–115, Nov. 2005, doi: 10.1016/j.ejor.2004.01.046.
- [28] B. K. Pagnoncelli, S. Ahmed, and A. Shapiro, "Sample average approximation method for chance constrained programming: Theory and applications," *Journal of Optimization Theory and Applications*, vol. 142, no. 2, pp. 399–416, Jul. 2009, doi: 10.1007/s10957-009-9523-6.
- [29] W.-K. Mak, D. P. Morton, and R. K. Wood, "Monte Carlo bounding techniques for determining solution quality in stochastic programs," 1999.
- [30] M. D. Mckay, R. J. Beckman, and W. J. Conover, "American Society for Quality A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code," 1979. doi: <https://doi.org/10.1080/00401706.1979.10489755>.