IMPROVING CUSTOMER WAITING TIME FOR MEDICINE-RETRIEVAL CENTER

Mateo Restrepo Sierra

Mathematical Engineering Universidad EAFIT Medellín, Colombia

Juan Sebastián Cárdenas-Rodríguez

Mathematical Engineering Universidad EAFIT Medellín, Colombia

David Plazas Escudero

Mathematical Engineering Universidad EAFIT Medellín, Colombia

ABSTRACT

Neuromedica is a Colombian pharmacy which provides treatment for people with neurological diseases. ´ Recently, Neuromédica started attending patients from other pharmacy which led to a significant increase in the waiting time. In this pharmacy, people are classified and attended due to certain priorities. The data, given by Neuromédica, is analyzed using boxplots, Kruskal-Wallis and Kolmogorov-Smirnov tests with Python's library Scipy. The objective of this work is to determine the number of assistants and queue logistic such that the waiting time has a significant reduction, with the purpose to provide a satisfactory level of service. A discrete-event simulation model was created and implemented in Python. A heuristic approach to minimize the waiting time is used. Additionally, a sensitivity analysis is made on the assumed distributions.

Keywords: Complex system modeling, queue theory, discrete-event simulation, simulation-optimization, statistical analysis of input-output data, white and black-box validation, conceptual modeling and python.

1 INTRODUCTION

Health care is one of the few systems that involve the great majority of the population, therefore there is an increasing demand on all the provided services by health institutions. This is the case of the system analyzed in this work: medication retrieval. These kinds of systems have complex interactions and often evolve into long queues with large waiting times, which creates dissatisfaction among the patients. Thus, health institutions have to face this issue and strive for solutions.

Among the different approaches to treat this problem, Discrete-Event Simulation (DES) emerges as a possible solution. Different authors, such as [\(Thorwarth and Arisha 2009\)](#page-17-0) and [\(Jun, Jacobson, and Swisher](#page-17-1) [1999\)](#page-17-1), provide a review of DES applied to health care. For the specific case of pharmacies, [\(Reynolds,](#page-17-2) [Vasilakis, McLeod, Barber, Mounsey, Newton, Jacklin, and Franklin 2011\)](#page-17-2) developed a DES model to analyze the automatization of some processes. More references can be found in the references within the mentioned work. Some more references in applications of queue theory in health care (and some in pharmacy) can be found in the excellent review provided by [\(Lakshmi and Iyer 2013\)](#page-17-3).

Following the ideas of the above-presented authors, we propose a DES model for the medication retrieval process in the Colombian context, focusing on queue logistics. The main justification for this methodology is that it provides a tool for experimentation without having to use the real system, it takes less time and does not bring risks [\(Reynolds, Vasilakis, McLeod, Barber, Mounsey, Newton, Jacklin, and](#page-17-2) [Franklin 2011\)](#page-17-2).

2 CONCEPTUAL MODEL

2.1 Problem Definition

Neuromédica is a health services institution (IPS for its abbreviation in spanish), specialized in health care for adult population focused on neurological disorders; this is a specialized center for Grupo [SURA.](https://www.epssura.com/) Their Almacentro complex is in charge of the distribution of medicine for their patients. Recently, they started receiving patients from DEMPOS (another medicine supplier in Colombia) and they are worried about the significant increase in the waiting time of their costumers.

In this complex, patients arrive and take a turn accordingly to their service type and priority, then they enter a queue until they are called into a specific counter; generally, they are called once and the medicine is provided instantly, but sometimes they are recalled a second time. In total, there are 15 counters. Each operator has some service types assigned with certain priorities, but they can also call patients outside these rules. It is important to remark that a patient enters the system exactly when he takes the turn.

In Figure [1,](#page-1-0) the general activity scheme is presented; it depicts the flow of a patient throughout the system; and in Figure [2](#page-2-0) we present a minimalist representation of the system at hand.

Figure 1: Activity Cycle Diagram

Restrepo, Cardenas, and Plazas ´

Figure 2: System diagram.

The complex attends the services presented in Table [1,](#page-2-1) each service has associated an ID; on the other hand, each module attends specific IDs with certain priorities, as shown in Table [2.](#page-3-0) For example, "Admissions 1" attends, first, service 4 which represents "Appointment confirmation & PAC, MP and policy SURA"; then it attends service 5 "Procedure after appointment"; and finally, it calls service 9 "Preferential appointment confirmation". It is important to mention that each module can attend outside this policies, for instance, if there are not patients waiting in line with ID 4, 5 or 9, "Admissions 1" can call patients with other IDs. Moreover, the attendants can voluntarily call, at any moment, any patient waiting in line.

Module	\mathbf{IDs}
Admissions 1	4,5,9
Admissions 2	4,9,5
Pharmacy 1	3,7,8,10
Pharmacy 2	3,8,7,10
Pharmacy 3	7,3,8
Pharmacy 4	7,3,8
Pharmacy 5	3,7,8,10
Pharmacy 6	3,8,7,10
Pharmacy 7	3,7,8,10
Pharmacy Retrieval	10
Pharmacy 9	3,8,7,10
Pharmacy 10	3,7,8,10
Pharmacy 11	3,7,8,10
Pharmacy 12	3,7,8,10
Pharmacy 13	3,8,7,10
Pharmacy 14	3,7,8,10

Restrepo, Cardenas, and Plazas ´

Table 2: Policy for each module

2.2 Modelling Objectives

2.2.1 Model Purpose

To determine the priorities of each module in order to reduce the queue time of patients.

2.2.2 Specific Objectives

- To select the most appropriate data to be considered in the model.
- To apply statistical tests to filtered data in order to contrast different hypothesis.
- To validate and verify the implemented model using the available data.
- To explore multiple feasible strategies for queue priorities in the modules.

2.2.3 General Project Objectives

- *Time-Scale*: a final report must be available in 5 weeks.
- *Flexibility*: high, since our implementation in Python allows us to make modifications easily.
- *Ease-of-use*: no, due to its implementation in Python.

We decided to implement the model in Python since it allows us to develop a more realistic model, fitting the available data to a wider set of distributions and obtaining a better fit. On the other hand, the Python model is more adaptative for our further experimentation and optimization; although, this kind of implementation makes the model not usable by everyone. Moreover, Python is open-source.

2.3 Model Inputs & Outputs

2.3.1 Experimental Factors/Model Inputs

- Number of modules with priority by hours.
- Queue logistics.
- Type of service.

2.3.2 Outputs

To determine achievement of objectives:

- Waiting time for each type of patient.
- Average time in system.
- Reduce the percentage of patients waiting.

To determine reasons for failure to meet objectives:

- Average number of patients waiting by service type per hour.
- Percentage of people waiting is not reduce.
- Leisure time.

2.3.3 Parameters

- Inter-arrival times for each type of patient and hour of the day and day of the week.
- Duration of the attention by type of service and attendant.
- Probability of quitting.
- Type of services.
- Inter-service times for each attendant.

2.4 Model Content

2.4.1 Model Scope

Table 3: Model scope.

Table 4: Model level of detail

2.4.3 Assumptions

- There will always be medicine available.
- There are no skippers.
- Attendants can call other priorities if they are available.
- Attendants cannot voluntarily call patients.

2.4.4 Simplifications

- The patients will not be called a second time, only once.
- Quitters are treated as a single constant probability.
- As we will discuss later in Section [7,](#page-11-0) we only consider the data provided for the last week of August.

3 BACKGROUND

In the research done by [\(Shimshak, Damico, and Burden 1981\)](#page-17-4), they propose several queueing models to evaluate the effectiveness of them and improve waiting times, considering the theoretical background discussed in [\(Cobham 1954\)](#page-17-5), where he claims the queue models can be classified in two types: singlechannel and multiple-channel systems. It is important to remark that these results constitute a theoretical and preliminary background in the queue theory with priorities.

Returning to the ideas proposed in [\(Shimshak, Damico, and Burden 1981\)](#page-17-4), they use three different queueing schemes:

- 1. Multiple server model queue with no priorities.
- 2. Priority discipline queueing model with two priorities.
- 3. Priority discipline queueing model with preemptive service.

The main idea of said work is to analyze different the described alternatives to determine the effective work distribution of New Jersey Hospital pharmacy. Finally, a more recent application was introduced by [\(Ndukwe, Omale, and Opanuga 2011\)](#page-17-6), where they apply queueing disciplines to improve the waiting times of outpatients in a Nigerian hospital, reducing the average waiting time from 167min to only 55.1min.

4 MODEL DATA

The data was provided directly by Neuromédica, it covers all records of August 2019. The data includes for each patient:

- Date.
- The arrival time.
- Waiting time.
- Time of service's start.
- Service duration.
- Time of service's end.
- Attendant.
- Counter.
- Type of service.
- Completion.

Furthermore, they provided another document that contains the types of services with their ID and their corresponding counters (see Tables [1](#page-2-1) and [2\)](#page-3-0). The names of the different attendants including their assigned services were provided as well.

It is important to highlight that the patient's ID was not granted by Neuromedica for security reasons; ´ thus, it was not possible to extend the system with SURA's authorization procedure.

All data in this work is handled in minutes.

4.1 Arrivals

Using the dataset available, the inter-arrival times were initially calculated and then filtered for each day of the week by hours. The resulting data were grouped by days of the week, for instance, we separated the data for each Monday of August at 6 am, 7 am, etc.

As we will discuss later in Section [7,](#page-11-0) the model did not pass the validation process and we decided to reduce the data to one week: last week of August, since it is the only complete week of the month, due to Holidays and different schedules in the Complex. This decision reduces the used data from 27k to 6.7k. Therefore all the results and figures that are going to be presented use this amount of data.

In Figures [3a](#page-8-0) and [3b,](#page-8-0) we present the autocorrelation and scatter plot respectively. These plots show that there's no significant autocorrelation in the data since the autocorrelation plot is not strictly above the x-axis and the scatter plot shows no obvious behavior.

Figure 3: Data for Monday 26 at 6 a.m.

For the implementation, the distributions of the inter-arrival times are required. Therefore, for each day of the week and each hour of the day, it was checked to see if they come from the same distribution, hence a homogeneity test, e.g. for each Monday of the month, at 6 am, we compared if the inter-arrival times come from the same distribution. The homogeneity was tested using a Kruskal-Wallis test [\(Wackerly,](#page-17-7) Muñoz, Humbertotr, et al. 2010, pp. 765-767).

In the same way, a Kolmogorov-Smirnov test [\(Banks, J., B.L., and D. 2005,](#page-17-8) pp. 230-231) was applied to determine the best distribution that fits the inter-arrival time data, for each hour in the month. The test was performed with all the available theoretical distributions available in the stats module in the [SciPy](https://docs.scipy.org/doc/scipy/reference/stats.html) package.

In the case of the rejection of homogeneity by the Kruskal-Wallis test, we selected the distribution using boxplots, selecting the one that better characterizes the general behavior. In the other case, we selected the best fit distribution for the first day. See Appendix [A](#page-17-9) for the detailed fitting.

For instance, we present the box plots of the data for Thursdays at 7 a.m. in Figure [4a.](#page-9-0) The Kruskal-Wallis test for this data concluded that there is not sufficient evidence to reject the null hypothesis, i.e. the data comes from the same distribution. Therefore the data can be considered homogeneous. The best-fit distribution is shown in Figure [4b.](#page-9-0)

Figure 4: Boxplots for inter-arrival times.

Figure 5: Distribution fitting for Thursdays at 7 a.m.

5 IMPLEMENTATION

We implemented the simulation model using the programming language Python. Python was chosen over a simulation software as this problem does not have many components but each of these components has many backend behaviors that are not shown, hence simulation software such as Simul8 would not be as useful. Furthermore, Python is open-source hence it could be used and tested by anyone who needed it.

For this implementation, we did not use a simulation package, instead, we programmed each of the components such as queues, event handling and more. We decided to take this approach, instead of using a simulation library as SimPy, for three main reasons. Firstly, to manage all the priorities that each module has and use it accordingly. Secondly, this approach is easy to run several times the simulation without using multi-processing packages. Lastly, so we can have more control over the system and the debugging process.

This implementation is easily modified, as it is object-oriented. Hence, changing behavior of any of the components would be done by changing the methods of each object and obtaining in this manner the desired change. Furthermore, visualizing the simulation can be done through outputs in the terminal but, a graphical user interface was not implemented.

luisinavasquez finished attending Customer 378 at 9:01. Customer 405 of type Farmacia General arrived at 9:01 luisinavasquez started attending Customer 403 at 9:01. leidydiaz started attending Customer 387 at 9:01. jenosita30 finished attending Customer 399 at 9:01. sceballos finished attending Customer 397 at 9:01. jenosita30 started attending Customer 405 at 9:01. angelaimitola finished attending Customer 400 at 9:01. leidydiaz finished attending Customer 387 at 9:01. sceballos started attending Customer 401 at 9:01. dianabetancur started attending Customer 372 at 9:01. angelaimitola started attending Customer 402 at 9:01. luisinavasquez finished attending Customer 403 at 9:02. Customer 406 of type Preferencial Farmacia arrived at 9:02 sezaocho finished attending Customer 386 at 9:02. dianabetancur finished attending Customer 372 at 9:02. Customer 407 of type Preferencial Farmacia arrived at 9:02 Customer 408 of type Preferencial Farmacia arrived at 9:03 vmonsalve started attending Customer 389 at 9:03. Customer 409 of type Farmacia General arrived at 9:03 dvalderrama started attending Customer 395 at 9:03. angelaimitola finished attending Customer 402 at 9:03. angelaimitola started attending Customer 408 at 9:03. Customer 410 of type Farmacia PAC, MP y Poliza Sura arrived at 9:03 sceballos finished attending Customer 401 at 9:03. Customer 411 of type Farmacia General arrived at 9:03 dvalderrama finished attending Customer 395 at 9:03. Customer 412 of type Farmacia General arrived at 9:03 vmonsalve finished attending Customer 389 at 9:04. Customer 413 of type Farmacia General arrived at 9:04 sceballos started attending Customer 409 at 9:04.

Figure 6: Examples of output of the simulation.

On the contrary, it is easily used as all of the important aspects of the simulation are packed in different methods of the main class. In this manner, extracting information, testing the model and extracting the outputs are just done by calling those respective functions.

6 RESULTS

We refer to [\(Robinson 2004,](#page-17-10) Ch. 9). The system in consideration has a fixed schedule since they open every business day at 6 a.m. and closes at 7 p.m. (Saturdays from 7 a.m. to 1 p.m.), therefore it can be considered as a terminating simulation model. Consequently, the output of the simulation is transient, hence the distribution of the output is constantly changing. For instance, the number of patients is not the same for each day.

The initial condition is assumed as zero because the system does not start with patients inside of it, although, in the complex, patients start arriving before the opening hour and start accumulating. Therefore, it is clear that the distributions in the first hour for all days are biased and this issue cannot be addressed with the available data.

How many times do we have to run the model in order to obtain trustworthy results? When the model is being simulated, it is desired to obtain the highest achievable accuracy, but this often requires a lot of simulation runs and, therefore, computational resources and time. Given that these objectives are conflictive, the idea is to find a balance between them. To achieve this goal, we use the strategy proposed in [\(Byrne 2013\)](#page-17-11) and [\(Currie and Cheng 2016\)](#page-17-12).

This methodology is based on confidence intervals; we select a maximum allowed deviation δ , in terms of a proportion of the mean. The simulation has to be previously run n_0 times, to estimate the standard deviation. The appropriate number must satisfy

$$
n \ge \left(\frac{t_{\alpha/2; n_0 - 1} \times \sigma}{\delta}\right)^2 \tag{1}
$$

Calculating this value for our model, using the test described previously we obtain a large value of *n*. The reason for this is because the fit distributions can generate a big amount of outliers, therefore it is difficult to find stability for the means.

Figure [7](#page-11-1) is a time series of the average waiting time of the model for a trial of 5 weeks, which shows clearly that the nature of the output is transient since it changes constantly.

Figure 7: Time series for average waiting time in a week

7 VALIDATION AND VERIFICATION

It is important to remark that some of the procedures that are going to be present in this section are doing twice. Model 1 uses the data of all the month provide by Neuromédica that we are going to refer to it as the model 1 and the other one, i.e model 2 only uses the data of the last of the month.

Although a model cannot be formally validated, we use different forms of validation: black and white validation. White validation is a more empirical procedure, checking that numerous parts of the model operate properly. On the other hand, black-box validation is intended to check the general behavior of the model, comparing it with the real system using confidence intervals [\(Robinson 2004,](#page-17-10) Ch. 12).

Furthermore, the waiting time for each of the type of customers was tested with the waiting time obtained in the simulation by the Kruskal-Wallis test.

7.1 White-Box Validation

We used several white-box validation methodologies since the model in consideration does not have a lot of components. The first test consists of submitting the model to extreme conditions, such as:

- One attendant.
- All modules open.

The results using one attendant in the Model 2 are shown in Figure [8](#page-12-0) which presented the total number of people that are waiting in the system for one week. This validates the model since it shows that the waiting times augment during the day when it is used only one attendant. It can see a repetitive pattern in the graphic, this represents when the system is closed, which provides a general idea that the model is working well.

Restrepo, Cardenas, and Plazas ´

Figure 8: Number of patients for each hour with one attendant

In Figs. [9,](#page-12-1) [10](#page-13-0) it can be seen two different box plots for each of the model compared with the one obtained in the data set. In both, the first box plot determines one in which the test assured that both where homogeneous and the second box plot is non-homogeneous.

It is important to see that in the second model, more outliers are generated. This was done with the intention of simulating more human behavior, as the attendants generally have some erratic behavior which which worsen the system. On the other hand, the second model has a lot more cases in which the data is homogeneous making it preferable to use.

Lastly, it is important to remark that the second model is not homogeneous in the type of client "Farmacia General" which is the one with the biggest waiting times in the real system. Hence, it is known that although the model works better than the previous one, it is not a fully working simulation model.

(a) Homogeneous (Pref Confirmación de citas). (b) Non-Homogeneous (Trámites después de Citas).

Figure 9: Boxplots validation for Model 1.

Figure 10: Boxplots validation for Model 2.

Furthermore, whilst the model was being implemented, multiple runs were performed to check initial conditions, arrivals/attention distributions and general proper behavior of the system.

7.2 Black-Box Validation

As previously stated, the general behavior of the model can be analyzed using the corresponding confidence interval to check the difference of the means. The formula its calculation is [\(Robinson 2004,](#page-17-10) pp. 218):

$$
\bar{X}_{S} - \bar{X}_{R} \pm t_{2n-2,\alpha/2} \sqrt{\frac{S_{S}^{2} + S_{R}^{2}}{n}}
$$
\n(2)

 $\bar{X}_{\rm S}$ = mean of simulated output data

 \bar{X}_R = mean of real system output data

 S_S = standard deviation of simulated output data

 S_R = standard deviation of real system output data

 $n =$ number of observations (this must be the same for the

simulated and real system data)

 $t_{2n-2,\alpha/2}$ = value from Student's *t* distribution with 2*n*−2 degrees of freedom and

a significance level of $\alpha/2$

So the obtained result has more sense, it is going to be used the type of pacient "Confirmacion de citas". This was chosen because this type of pacient is the one that the simulation is capable of characterizing. Hence, all experiments are going to be analyzed through the waiting time of this customer.

With this test, it was found that the interval for the difference of the means is:

$$
[-3.74976, 1.57127]
$$
 (3)

8 SENSITIVITY ANALYSIS AND EXPERIMENTATION

8.1 Sensitivity Analysis

To see the robustness of the model an sensitivity analysis was achieved. This procedure consists of a variation of some of the parameters of the model between -10% to 10% .

The model, as explained before, some parameters where not able to be fitted because the data was too small. In this manner, we chose to see how the parameters of this elected distributions affected the output of the model. The result if this test is shown in Fig. [11.](#page-14-0)

Figure 11: Sensitivity analysis

It is seen that, although the simulation does change while changing the parameters of this distributions, this does not present a major problem as the output of the model does not change drastically. Hence, it is confirmed that the suppositions are not affecting in a damaging manner the simulation.

8.2 Optimization

The priority that each module attend (Table [2\)](#page-3-0) is not only is a key factor to consider in the queue logistics, but also a highly feasible strategy to implement in the real system since it only requires the adjustment on the queueing platform. The idea is to find the best combinations of priorities per (open) module to minimize the average queue time. The optimization scheme we propose takes the initial "solution", i.e. Table [5,](#page-14-1) and apply a heuristic approach based on local search.

Table 5: Sets and priorities.

We use a standard Variable Neighborhood Search (VNS) (Hansen, Mladenović, and Urošević 2006), using classic neighborhood operators:

• In-Set Swap: Swaps two priorities within the same set (Figure [12\)](#page-15-0).

Figure 12: In-set swap operator.

• Best Insertion: Takes a node and attempts to find the best insertion within the set (Figure [13\)](#page-15-1).

 \mathbb{P}^{16} for $[p_1 \ p_2 \ ... \ p_{i-1} \ p_{i+1} \ ... \ p_j \ p_i \ p_{j+1} ... \ p_n]$ set k $[p_1 p_2 ... p_{i\text{-}1} p_i p_{i\text{+}1} ... p_j p_{j\text{+}1} ... p_n]$

Figure 13: Best-insertion in-set operator.

• Inter-Set Swap: Same as in-set, but attempting between two nodes in different sets [\(14\)](#page-15-2).

$$
\begin{array}{c}\n\text{set } k_1[p_1 p_2 \dots p_i \dots p_n] \\
\downarrow \\
\text{set } k_2[q_1 q_2 \dots q_j \dots q_m]\n\end{array}
$$

Figure 14: Inter-set swap operator.

We refer to [\(Paraskevopoulos, Repoussis, Tarantilis, Ioannou, and Prastacos 2008\)](#page-17-14) and (Fosin, Carić, [and Ivanjko 2014\)](#page-17-15) for the local search operators. If the third operator swaps a priority into a set that already had that priority, one is chosen randomly and deleted. Otherwise, we assume that the number of priorities will not change, e.g. the number of appearances in all sets of policy "3" remains constant on each iteration and so forth.

After the optimization is done, an ANOVA test is performed in order to check if the mean of the output data is significantly improved after the optimization (Wackerly, Muñoz, Humbertotr, et al. 2010, pp. 662-665).

The result of this heuristic is [6.](#page-16-0) Furthermore, the evolution of the waiting time for the selected type of patient is shown in the Fig. [15](#page-16-1)

Restrepo, Cardenas, and Plazas ´

Figure 15: Waiting time for selected pacient changing over the iterations of VNS

Set	Amount	Priority
1	1	[5,4,9]
2	1	[10]
3	1	[8, 10, 7, 3]
4	1	[7,3,10,8]
5	1	[8, 10, 3, 7]
6	1	[3,7,8,10]
7	1	[8,3,7,10]
8	1	[7,8,3,10]
9	1	[8,7,3,10]
10	1	[10,7,3,8]
11	1	[3,8,7,10]
12	1	[7,3,8,10]

Table 6: Found table of priorities.

It is important to notice that the best solution found was by using all the resources in all the different ways possible, so each one attend a different and unique priority. This is due to using all the resources to attend all the possible priorities, therefore reducing this waiting time.

9 CONCLUSIONS

It was successfully analyzed the received data, using statistical tests which allowed to fit data to use in the simulation, find auto-correlation and test homogeneity between data. It is important to remark that in the fitting of the data the distributions have a left bias as most of the times are positive and close to zero.

It was successfully implemented the simulation model in Python, which partially represents the Neuromédica pharmacy medicine retriveal system. The model did not fully validate with the tests done, which presents an important limitation of the implemented model; even though, the results given are not useful for the real system, this article proposes a methodology for solving this types of problems.

For further work, a model using continuous simulation could be used as the inter-arrival times are so close to zero that the system could be analyzed supposing a continuous flux of patients. Also, a graphical user interface can be done so it is easily explained without the use of code.

REFERENCES

- Banks, J., C. J., N. B.L., and N. D.. 2005. *Discrete-event System Simulation*. Prentice-Hall international series in industrial and systems engineering. Pearson Prentice Hall.
- Byrne, M. D. 2013. "How many times should a stochastic model be run? An approach based on confidence intervals". In *Proceedings of the 12th International conference on cognitive modeling, Ottawa*.
- Cobham, A. 1954. "Priority assignment in waiting line problems". *Journal of the Operations Research Society of America* 2 (1): 70–76.
- Currie, C. S., and R. C. Cheng. 2016. "A practical introduction to analysis of simulation output data". In *2016 Winter Simulation Conference (WSC)*, 118–132. IEEE.
- Fosin, J., T. Caric, and E. Ivanjko. 2014. "Vehicle routing optimization using multiple local search ´ improvements". *automatika* 55 (2): 124–132.
- Hansen, P., N. Mladenović, and D. Urošević. 2006. "Variable neighborhood search and local branching". *Computers & Operations Research* 33 (10): 3034–3045.
- Jun, J., S. H. Jacobson, and J. R. Swisher. 1999. "Application of discrete-event simulation in health care clinics: A survey". *Journal of the operational research society* 50 (2): 109–123.
- Lakshmi, C., and S. A. Iyer. 2013. "Application of queueing theory in health care: A literature review". *Operations research for health care* 2 (1-2): 25–39.
- Ndukwe, H., S. Omale, and O. Opanuga. 2011. "Reducing queues in a Nigerian hospital pharmacy".
- Paraskevopoulos, D. C., P. P. Repoussis, C. D. Tarantilis, G. Ioannou, and G. P. Prastacos. 2008. "A reactive variable neighborhood tabu search for the heterogeneous fleet vehicle routing problem with time windows". *Journal of Heuristics* 14 (5): 425–455.
- Reynolds, M., C. Vasilakis, M. McLeod, N. Barber, A. Mounsey, S. Newton, A. Jacklin, and B. D. Franklin. 2011. "Using discrete event simulation to design a more efficient hospital pharmacy for outpatients". *Health care management science* 14 (3): 223–236.

Robinson, S. 2004. *Simulation: The Practice of Model Development and Use*. Wiley.

- Shimshak, D. G., D. G. Damico, and H. D. Burden. 1981. "A priority queuing model of a hospital pharmacy unit". *European journal of operational Research* 7 (4): 350–354.
- Thorwarth, M., and A. Arisha. 2009. "Application of discrete-event simulation in health care: a review".
- Wackerly, D. D., R. Muñoz, J. Humbertotr et al. 2010. *Estadística matemática con aplicaciones*. Number 519.5 W3.

A Appendix

A.1 Distribution Fitting

As mentioned in Section [4,](#page-7-0) the fitting of distributions was performed in Python using SciPy. The fitting was performed with the Kolmogorov-Smirnov test against all the available theoretical distributions. In Table [7](#page-17-16) the best fit distributions for each day and each hour are presented. We use the conventions in SciPy. The distribution name matches the online documentation and the parameters have the following convention: [*distribution parameters*, location, scale]. All the parameters are used according to the probability density functions presented in SciPy documentation.

	Farmacia PAC, MP laplace y Poliza Sura		0.94197	[4.05, 2.19116]
	Preferencial Far- macia	exponpow	0.99944	[0.60618, 0.23333, 4.92841]
	Entrega Progra- mada	norm		[1.93889, 3.60495]
Att 9	Farmacia General	johnsonsu	0.88688	$[-5.09161, 1.07842, -0.24934,$ 0.04755]
	Farmacia PAC, MP y Poliza Sura	norm		[4.29545, 4.93202]
	Preferencial Far- macia	norm		[2.11389, 2.31354]
	Entrega Progra- mada	norm		[4.25, 5.26326]
Att 19	Farmacia General	johnsonsb	0.50391	[8.74138, 1.38249, -0.49468, 1834.287611
	Preferencial Far- macia	nakagami	0.91117	[0.37578, 0.23333, 4.78737]
Att 1	Farmacia General	gengamma	0.45301	$[10.30977, 0.35725, -0.08809,$ 0.00339]
	Farmacia PAC, MP y Poliza Sura	norm		[2.55, 1.18782]
	Preferencial Far- macia	norm		[2.55833, 1.37886]
	Entrega Progra- mada	norm		[0.59375, 0.63076]

Table 7: Attendants distribution fitting by hours and service.

Table 8: Inter-arrival times distribution by hours.