UNIVERSITY OF TWENTE.

**Faculty of Electrical Engineering,
Mathematics & Computer Science**

Predicting hospital bed census due to planned surgeries using queueing models

Bram de Jong

MSc. Thesis Applied Mathematics
January 13, 2022

Supervisors:
Prof. Dr. R.J. Boucherie
Dr. ir. W.R.W. Scheinhardt
Dr. K. Proksch

Stochastic Operations Research
Faculty of Electrical Engineering,
Mathematics and Computer Science
University of Twente

Abstract

Nurses are scheduled for work according to a prediction for the number of inpatient patients. The size of the prediction interval plays a key role here. A model has been created to predict the number of patients that are inpatient due to planned surgeries and to analyze its prediction interval. This has been done by using two approaches: firstly by applying an $M/M/\infty$ queueing model and secondly by applying an $M/G/\infty$ queueing model. We have applied the $M/G/\infty$ model to a normal and log-normal service distribution, where an additional update rule is introduced once a patient undergoes surgery. The schedule, or blueprint, of these surgeries has been either deterministically made in advance, stochastically made in advance or the schedule can be altered up until the moment of surgery. For both the $M/M/\infty$ model and the $M/G/\infty$ model, we have simulated both the deterministic and stochastic blueprint for various parameters. Additionally, for the $M/G/\infty$ model, we have used a log-normal service distribution, where we compare the additional update rule to the case where no additional update rule is used. We see that the introduction of the additional update rule does not benefit the quality of the prediction. The size of the prediction interval of the models does not decrease with absolute certainty, but the simulations show that it decreases almost always when time progresses.

Preface

I would like to express my gratitude towards my graduation committee. First of all, I would like to thank prof. dr. Richard Boucherie for his supervision and effort he put into this project. Even though this project was entirely during the COVID-19 pandemic, our weekly meetings and involvement of his research group helped me feel connected to other mathematical projects.

Secondly, I would like to thank dr. ir. Werner Scheinhardt and dr. Katharina Proksch for reading my work and evaluating it.

Even though this is my thesis, I will have not yet completed my master. I still have to do my internship, which I will do abroad in South-Korea. I am very much looking forward to it.

Contents

1 Introduction

In this Section, we first give a short overview of the situation. Then, we state the research goal and our approach on the topic. Finally, we present the structure of the remainder of the report.

1.1 Background

After undergoing surgery, patients arrive in the hospital ward, where they stay inpatient for a certain amount of time. This amount of time is called the "Length of Stay" and is unknown before they get discharged from the hospital. Consider only one hospital department of patients. In this department, there are already patients who have been there for various times. On each of these patients, there is certain information available, e.g. how long they have been inpatient already, what type of surgery they underwent (arm, leg) and who performed surgery on them. In the future, more patients are scheduled for surgery. Of these new patients, there is fewer information available, as you might not yet know their type of surgery or their surgeon.

Based on all the information that is available, a prediction for the number of inpatient patients somewhere in the future is made. The information that is used for this prediction not only influences the prediction for the number of patients, but also influences the corresponding prediction interval. The number of nurses required for these patients is based on this prediction and its corresponding prediction interval. The working schedule of these nurses is created three months prior to date, with the possibility of change until one month prior. Since hospitals want to ensure that every patient is sufficiently cared for, the number of nurses that are scheduled for work is on the high-end of the prediction interval. Usually, this results in a lot of "wasted" work hours for nurses.

This could be improved on by reducing the size of the prediction interval (while maintaining the same level of significance). Using new information that is gained over time plays a key role in order to gain more certainty on this prediction.

1.2 Goals

The goal of this report is to use the new knowledge that becomes apparent over time, to decrease the size of the prediction interval and reduce the amount of wasted work hours. This report focuses on the knowledge that is gained due to the number of patients that are currently present in the ward.

We set the following objective:

> "We aim to make a model that can be updated to incorporate the new information that is gained to predict the number of patients being present in the ward for some fixed point in the future"

With this research objective, we propose the following research questions.:

1. Does the size of the prediction interval decrease with absolute certainty?
2. How significant is the information gain per time epoch when we make the prediction far into the future compared with making the prediction near in the future?
3. Is there a significant difference in modelling the schedule in various ways?

Sections 3 and 4 discuss question 1, while numerical results in Section 5 discuss question 2.

1.3 Approach

We will only focus on patients that are inpatient due to planned surgeries. In practice, acute surgeries should be taken into account as well. Since we will assume that there is an infinite number of beds available, these acute patients can be implemented into our model without interference with the planned patients.

The schedule (which we also refer to as blueprint) plays a key role in modelling this situation. In this report, we consider three types of schedules of the performed surgeries. Each of these three blueprints only determines the number of patients that undergo surgery in a certain time interval. This implies that the type of surgery is never known beforehand. Every type of blueprint is analyzed for both the $M/M/\infty$ model and the $M/G/\infty$ model. The three types are:

1. a deterministic blueprint: this gives us all the information on the number of surgeries performed in the future. This number is fixed throughout time.

2. a stochastic blueprint: this models the situation in which the number of surgeries performed in the future is a random variable, up until the day arrives on which the surgeries takes place. In this report, we will make certain assumptions on the distribution of this type of blueprint.

3. a dynamic blueprint: this starts with an initialization on the number of surgeries performed in the future, similar to the deterministic blueprint. However, in this scenario more surgeries can be scheduled throughout time. Again, we will make certain assumptions on the distribution of the number of added surgeries

In Sections 3 and 4, we will do a more technical analysis on the effects of these different blueprints.

1.4 Structure of the report

In this report, we create a model, based on queueing theory, that gives a prediction for the number of patients that are inpatient at a certain time. When this prediction is made at various times, the impact of gaining additional information at later times on the prediction and its corresponding prediction interval will be analyzed. The surgeries patients undergo will all be planned surgeries, implying that some sort of schedule is made beforehand. This schedule could be also updated. We look at various options for how this schedule is determined.

Section 2 discusses the available literature on this research topic. Regarding the model, Section 3 analyzes the $M/M/\infty$ queue, after which Section 4 analyzes $M/G/\infty$. In the Section on the $M/G/\infty$ queue, we also look at some specific distributions for the Length of Stay.

Furthermore, we show numerical results of this model, using simulations of these queues by generating patients with their Length of Stay coming from various distributions. Section 5 shows how the model is first be applied to specific cases of the $M/M/\infty$ model, after which we will apply it to similar cases on the $M/G/\infty$ model. In Section 6 we summarize and discuss the results of this report. Finally, recommendations for further research are given.

2 Literature review

In this Section, we will look into insights provided by other research. First of all, the analysis of several queueing systems in hospital has been done by others. Bekker and Koeleman analyzed the variability of hospital admissions following a $G/G/\infty$ queue [1]. They present an optimization for the scheduling of arrivals. In [2], a technique for balancing the lost patients and costs of a patient of an $M/PH/c$ queue is presented. A forecasting model which uses a Richards' curve to predict the arrival rate and a Kaplan-Meier estimation for the Length of Stay is used in [3]. This paper presents a data driven prediction model, based on a Poisson arrival process, to predict the bed census in the ward and ICU of a hospital. Other research on predictions and performance measures of hospitals are [4], [5] and [6].

In [7], an hourly bed census prediction as a function of the Master Surgery Schedule and arrival patterns of acute patients is presented. Analysis on the Master Surgery Schedule is done in [8]. We see that the schedules that other papers use/generate are often more complex than the schedules that are used in this report. This implies that no work could be found using this scheduling approach.

The parameters of the queue could also be estimated. In [9], an estimator for the Length of Stay distribution of an $M/G/\infty$ queue is proposed. An estimation based on the queue length process for the Length of Stay distribution is given in [10]. A Bayesian Neural Network approach is used in [11], to determine the posterior distribution of the Length of Stay. Bayesian analysis on paramaters will not be applied in this report.

A model for the nurse staffing is proposed in [12], based on the hourly bed census predictions in [7]. More insights on nurse staffing can be found in [6]. Further analysis on this topic will not be done.

3 The M/M/∞ queue

In this Section, we will make a prediction model for the number of inpatient patients, assuming a Poisson arrival process, an exponentially distributed Length of Stay and an infinite amount of beds. We first analyze the characteristics of this $M/M/\infty$ queue. Then, we will use this analysis to create a model that is able to make the prediction at various times and use the additional information that is gained at those times, for several scheduling systems of the hospital. We will show some numerical results of this model in Section 5.1.

3.1 Analysis on the M/M/∞ queue

Let us consider only one department of patients. At time $t = 0$, there is a certain number of patients already being inpatient (this number could be a random variable). For these current patients, the distribution of the time until discharge is known (with known rate μ_0). New patients will arrive with rate λ and have a discharge rate μ_1 up until a time $t = \tau_1$. After τ_1, patients with discharge rate μ_2 will arrive with rate λ until time $t = \tau_2$ and so on. Let $X_i(t)$ denote the number of occupied beds by patients with discharge rate μ_i at time t and let $P^i_k(t) = \mathbb{P}[X_i(t) = k)]$ the probability that there are k patients of type i at time t. The assumption is that the $X_i(t)$ are independent. Let $Y(t)$ denote the total number of patients that are inpatient.

Let us consider $0 \leq t \leq \tau_1$ and that there are n_0 patients at $t = 0$. The total number of patients at time t is then $X_0(t) + X_1(t)$. $X_0(t)$ is a pure-death process (see Figure 11 in Appendix A), as there are no new patients coming in and $X_1(t)$ is a birth-death process.

To analyze these processes, we will use the Kolmogorov forward equations

[13]:

$$\frac{d}{dt}P^i_{xy}(s;t) = \sum_k P^i_{xk}(s;t)A^i_{ky}(t) \tag{1}$$

where $A^i(t)$ is the transition rate matrix of type i patients. We will only look at starting at time $t = 0$ in state n_0. The Kolmogorov forward equations for the pure-death process are:

$$\frac{d}{dt}P^0_k(t) = (k+1)\mu_0 P^0_{k+1}(t) - k\mu_0 P^0_k(t), k \leq n_0 - 1 \tag{2}$$

with initial condition $P^0_{n_0}(0) = 1$ and

$$\frac{d}{dt}P^0_{n_0}(t) = -n_0\mu_0 P^0_{n_0}(t) \tag{3}$$

The solution is:

$$P^0_k(t) = \binom{n_0}{k} e^{-\mu_0 kt}(1 - e^{-\mu_0 t})^{n_0-k}, k \geq 0 \tag{4}$$

which is a binomial distribution with parameters n_0 and $e^{-\mu_0 t}$. The derivation can be found in Appendix B.

For the birth-death process (see Figure 12 in Appendix A), the Kolmogorov equations are:

$$\frac{d}{dt}P^1_0(t) = \mu_1 P^1_1(t) - \lambda P^1_0(t) \tag{5}$$

$$\frac{d}{dt}P^1_k(t) = \lambda P^1_{k-1}(t) + (k+1)\mu_1 P^1_{k+1}(t) - (\lambda + k\mu_1)P^1_k(t), k > 0 \tag{6}$$

with initial condition $P^1_0(0) = 1$. The solution is [14]:

$$P^1_k(t) = e^{-\rho_1(t)}\frac{\rho_1^k(t)}{k!}, k \geq 0, \tag{7}$$

where $\rho_1(t)$ satisfies

$$\rho_1(\tau_0) = 0 \tag{8}$$

The solution for $\rho_1(t)$ is:

$$\rho_1(t) = \frac{\lambda}{\mu_1}(1 - e^{-\mu_1 t}), t \in (\tau_0, \tau_1] \tag{9}$$

The derivation can be found in Appendix B. Hence, $X_1(t)$ is a Poisson distributed random variable with mean $\rho_1(t)$.

To determine the distribution of $Y(t)$, we will use the probability generating function (PGF). The probability generating function of a random variable X is defined as [15]:

$$G(z) = \mathbb{E}[z^X] = \sum_{x=0}^{\infty} p(x) z^x \tag{10}$$

$X_0(t)$ has binomial distribution, so it has the following PGF:

$$G_{X_0(t)} = (e^{-\mu_0 t}(z-1)+1)^{n_0} \tag{11}$$

$X_1(t)$ is Poisson distributed, so it has the following PGF:

$$G_{X_1(t)} = e^{-\rho_1(t)(1-z)} \tag{12}$$

Multiplying these two gives the PGF of the total number of patients $Y(t)$, $0 \leq t \leq \tau_1$:

$$\begin{aligned} G_{Y(t)} &= G_{X_0(t)} \cdot G_{X_1(t)} \\ &= (e^{-\mu_0 t}(z-1)+1)^{n_0} \cdot e^{-\rho_1(t)(1-z)} \end{aligned} \tag{13}$$

which is an explicit expression for the PGF of $Y(t)$.

If $X_0(0)$ is a Poisson distributed random variable with mean λ_0, $G_{X_0(t)}$ can then be determined as follows:

$$
\begin{aligned}
\mathbb{E}[z^{X_0(t)}] &= \sum_{n=0}^{\infty} P(N_0 = n) \cdot \mathbb{E}[z^{X_0(t))} | N_0 = n] \\
&= \sum_{n=0}^{\infty} \frac{\lambda_0^n}{n!} e^{-\lambda_0} \cdot (e^{-\mu_0 t}(z-1)+1)^n \\
&= e^{-\lambda_0} \sum_{n=0}^{\infty} \frac{(\lambda_0(e^{-\mu_0 t}(z-1)+1))^n}{n!} \\
&= e^{-\lambda_0} \cdot e^{\lambda_0(e^{-\mu_0 t}(z-1)+1)} \\
&= \exp(-\lambda_0(e^{-\mu_0 t}(1-z)))
\end{aligned}
\tag{14}
$$

which is the PGF of a Poisson distributed random variable. This implies that

$$
X_0(t) \sim \text{Poisson}(\lambda_0 e^{-\mu_0 t}) \tag{15}
$$

The PGF of $Y(t)$ is now the product of the PGF's of the two Poisson random variables, which results in the PGF of a Poisson random variable with mean the sum of the means of $X_0(t)$ and $X_1(t)$:

$$
G_{Y(t)} = \exp\left[-(\lambda_0 e^{-\mu_0 t} + \rho_1(t))(1-z)\right] \tag{16}
$$

Now, let us consider $\tau_1 \leq t \leq \tau_2$. The total number of patients now consists of three types of patients. The first two types are now both pure-death processes, as no new patients of those types will arrive. Once again, assume the initial number of patients at $t = 0$ to be Poisson distributed, denoted by N_0 and denote the number of patients of type 1 at $t = \tau_1$ by N_1. Note that $N_1 = X_1(\tau_1)$ and thus N_1 is also Poisson distributed. The PGF of each individual type can now be determined:

$$
G_{X_1(t)} = \exp(-\rho_1(\tau_1)(e^{-\mu_1(t-\tau_1)}(1-z))) \tag{17}
$$

$$G_{X_2(t)} = e^{-\rho_2(t-\tau_1)(1-z)} \tag{18}$$

where the PGF of $X_1(t)$ is shifted in time and N_1 is used and the PGF of $X_2(t)$ is the same as in (12), but also shifted in time (and different service rate). $G_{X_0(t)}$ is the same as in (14). Since these are now all Poisson distributed, the total amount $Y(t)$ will also be Poisson distributed. The same can now be applied to all patient types at all times.

3.2 Recursive approach on patients with a prolonged stay

In the previous Section it was shown that $Y(t)$ will always be Poisson distributed, under the assumption that N_0 is Poisson distributed. After every time epoch, one additional term is added to the mean of this Poisson distribution. We see that for $Y(\tau_1)$ we have:

$$Y(\tau_1) = X_0(\tau_1) + X_1(\tau_1) \tag{19}$$

Starting for type 0 patients, the only information needed is the number of patients inpatient at the previous time epoch. All of these patients have discharge rate μ_0. Now, since both of the types present at τ_1 are Poisson distributed:

$$Y(\tau_1) \sim \text{Poisson}(\lambda_1), \tag{20}$$

where

$$\lambda_1 = \lambda_0 e^{-\mu_0 \tau_1} + \rho_1(\tau_1) \tag{21}$$

Now, after τ_1, we see that $X_1(t)$ has become a pure-death process as well. For $Y(\tau_2)$ we have:

$$Y(\tau_2) = X_0(\tau_2) + X_1(\tau_2) + X_2(\tau_2) \tag{22}$$

Similarly:

$$Y(\tau_2) \sim \text{Poisson}(\lambda_2), \tag{23}$$

where

$$\lambda_2 = \lambda_0 e^{-\mu_0 \tau_2} + \rho_1(\tau_1) e^{-\mu_1(\tau_2 - \tau_1)} + \rho_2(\tau_2) \tag{24}$$

Hence, we see that

$$Y(\tau_j) \sim \text{Poisson}(\lambda_j), \tag{25}$$

where

$$\lambda_j = \lambda_0 e^{-\mu_0 \tau_j} + \sum_{i=1}^{j} \rho_i(\tau_1) e^{-\mu_i(\tau_j - \tau_i)} \tag{26}$$

So far, we have not used the additional information gained each time epoch. However, we now consider patients that stay inpatient longer than a certain amount of time to have the same discharge rate. Let this time be equal to k time intervals (which may differ in length). We then have the following equation for λ_j:

$$\lambda_j = \begin{cases} \lambda_0 e^{-\mu_0 \tau_j} + \sum_{i=1}^{j} \rho_i(\tau_1) e^{-\mu_i(\tau_j - \tau_i)} & j \leq k \\ \lambda_0 e^{-\mu_k \tau_j} + \sum_{i=1}^{j-k-1} \rho_i(\tau_i) e^{-\mu_k(\tau_j - \tau_i)} + \sum_{i=j-k}^{j} \rho_i(\tau_i) e^{-\mu_i(\tau_j - \tau_i)} & j > k \end{cases} \tag{27}$$

3.3 Modelling and updating the M/M/∞ queue

Let time be slotted, with prediction epochs $\tau_0, \tau_1, \ldots, \tau_S$. Our goal is to predict, at decision epoch τ_s, the number of patients present at time τ_S and relate these predictions at epochs τ_s and τ_{s+1}, $s = 0, \ldots, S-1$. We will refer to patients that undergo surgery and arrive in the ward in $(\tau_{i-1}, \tau_i]$

as type i patients and patients in the system at time 0 as type 0 patients. We assume that the distribution of the Length of Stay of type i patients is exponential(μ_i).

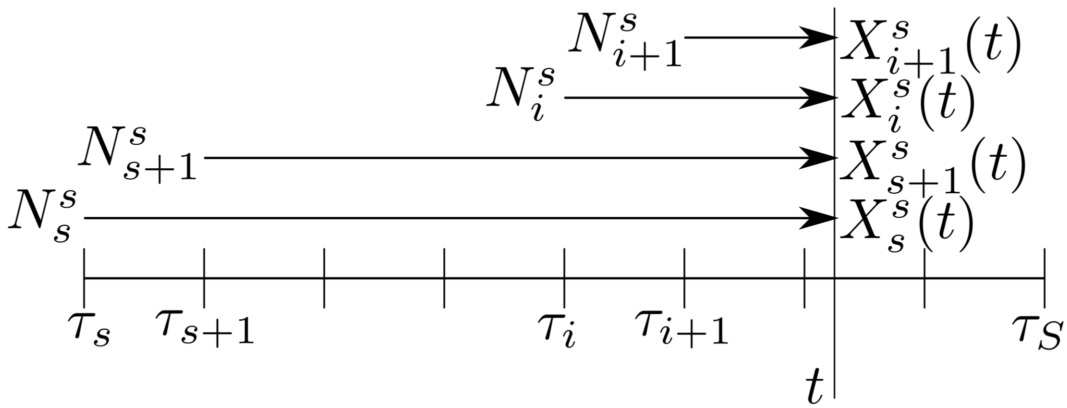

Figure 1: Overview of the situation at prediction epoch τ_s

Consider prediction epoch τ_s (see Figure 1). For each epoch τ_s, we will be interested in the predicted number of patients present at epochs $\tau_s, \tau_{s+1}, \ldots, \tau_S$. Let N_0^0 denote the number of patients in the ward at epoch τ_0. Let N_i^s denote the predicted number of type i patients present in the ward at epoch τ_i for $i = s+1, \ldots, S$ and n_i^s be the realization of N_i^s for $i = 0, \ldots, s-1$. Let N_s^s denote all the patients currently in the ward. Let $X_i^s(t)$ be the (predicted) number of type i patients present at time t, $t \geq \tau_s$. So $N_i^s = X_i^s(\tau_i)$, $i = s, \ldots, S+1$. Let $Y^s(t) = \sum_{i=s}^{S} X_i^s(t)$, $t \geq \tau_s$ be the total (predicted) number of patients that are in the system at time t. The predicted number of patients of type i present in the ward at time τ_i, $i > s$, can be divided in two classes:

- A_i^s - the number of patients that are scheduled before time τ_s to undergo surgery in $(\tau_{i-1}, \tau_i]$

- B_i^s - the number of patients that will be scheduled during $(\tau_s, \tau_{i-1}]$ to undergo surgery in $(\tau_{i-1}, \tau_i]$, but are not scheduled yet

So

$$N_i^s = A_i^s + B_i^s, i > s \tag{28}$$

For $i = s$, we state that $N_s^s = A_s^s$, thus A_s^s also denotes all the patients currently present in the ward.

Consider prediction epoch τ_s and time interval $(\tau_{i-1}, \tau_i]$, $i = s+1, \ldots, S$. As time progresses and additional information is obtained about patients for surgery in $(\tau_{i-1}, \tau_i]$, A_i^s will increase and B_i^s will decrease. We assume that the service of a type i patient starts at τ_i. Let C_i^s denote the number of type i patients that will be scheduled in $(\tau_{s-1}, \tau_s]$ for surgery. We assume that C_i^s is Poisson distributed with mean λ_i^s. We see that

$$A_i^{s+1} = A_i^s + C_i^{s+1} \tag{29}$$

and

$$B_i^s = C_i^{s+1} + B_i^{s+1} \tag{30}$$

Note that if we use the realization of A_i^s, A_i^{s+1} has a shifted Poisson distribution. B_i^s has a Poisson distribution with mean $\sum_{j=s+1}^{S} \lambda_i^j$, since $B_i^S = 0$. We denote this mean by γ_i^s.

When the clock has moved one step forward and we make the prediction one epoch later, some things have changed. First, N_i^{s+1}, $s+1 \leq i \leq S$ has gained more certainty, due to the extra information given by A_i^{s+1}. Secondly, we know more about the service rate of the patients who underwent surgery in $(\tau_s, \tau_{s+1}]$. We update the service rate every time epoch by decreasing the index by 1, if possible. This results in:

$$\rho_i^s(\tau_i) = \frac{\lambda_i^s}{\mu_{i-s}}(1 - e^{-\mu_{i-s}(\tau_i - \tau_{i-1})}) \tag{31}$$

The service rates could be further updated by applying a Bayesian statistical analysis to use the information gained on these patients. This will not be done here, but we will keep distinguishing between different service rates to ensure that applying a statistical analysis is still possible.

Consider making the prediction of the number of patients at different times. $X_i^s(t)$ has the following distribution (see (4)):

$$X_i^s(\tau_S) \sim \text{Bin}\left(N_i^s, e^{-\mu_{i-s}(\tau_S-\tau_i)}\right), s \leq i \leq S \tag{32}$$

Note that $X_S^s(\tau_S)$ equals N_S^s, which is also the case in (32). Assuming all the service and arrival rates are known, we can determine a prediction interval for $Y^s(\tau_S)$, using its expectation and variance, and compare these:

$$\mathbb{E}[Y^s(\tau_S)] = \sum_{i=s}^{S} \mathbb{E}[X_i^s(\tau_S)] \tag{33}$$

We assume that all $X_i^s(t)$ are independent, so the following holds:

$$\text{Var}(Y^s(\tau_S)) = \sum_{i=s}^{S} \text{Var}(X_i^s(\tau_S)) \tag{34}$$

This implies for $Y^s(t)$:

$$\mathbb{E}[Y^s(\tau_S)|N_i^s, i = s \dots S] = \sum_{i=s}^{S} \left[e^{-\mu_{i-s}(\tau_S-\tau_i)} \cdot N_i^s\right] \tag{35}$$

Regarding decision epoch $s+1$:

$$X_i^{s+1}(\tau_S) \sim Bin\left(N_i^{s+1}, e^{-\mu_{i-(s+1)}(\tau_S-\tau_i)}\right), s+1 \leq i \leq S \tag{36}$$

and thus

$$\mathbb{E}[Y^{s+1}(\tau_S)|N_i^{s+1}, i = 0 \ldots S] = \sum_{i=s+1}^{S} \left[e^{-\mu_{i-s+1}(\tau_S - \tau_i)} \cdot N_i^{s+1}\right] \tag{37}$$

From (36), you can already see that the variance of the model is likely to be lower. Since the variance of a binomial distributed random variable equals $np(1-p)$, the closer p gets to 1, the closer the variance goes to 0. This is the case when we update the model. We will now look at the various scheduling systems for the surgeries.

3.3.1 A deterministic blueprint

We first look at the special case where the entire surgery schedule has already been determined from the beginning, implying the number of patients who will undergo surgery is known. We denote the realization of A_i^s by a_i, $s+1 < i \leq S$. Note that $B_i^s = 0$ and that there is no necessity for a prediction epoch index for a_i, since it is a constant. Let A_s^s still denote all the patients that are currently inpatient. Thus we have:

$$N_i^s = a_i \tag{38}$$

We assume now that the length of all time intervals and all service rates are equal, being T and μ respectively. Applying this to (32), the expectation and variance of $X_i^s(\tau_S)$, $i \geq s+1$ become:

$$\mathbb{E}[X_i^s(\tau_S)] = a_i \cdot p_i \tag{39}$$

and

$$\mathrm{Var}(X_i^s(\tau_S)) = a_i \cdot p_i(1 - p_i) \tag{40}$$

where

$$p_i = e^{-\mu(S-i)T} \tag{41}$$

The expectation and variance of $X_i^s(\tau_S)$, $i \leq s$ become:

$$\mathbb{E}[X_i^s(\tau_S)] = x_i^s(\tau_s) \cdot p_s \tag{42}$$

and

$$\mathrm{Var}(X_i^s(\tau_S)) = x_i^s(\tau_s) \cdot p_s(1 - p_s), \tag{43}$$

where $x_i^s(\tau_s)$ is the realization of $X_i^s(\tau_s)$. For $Y^s(\tau_S)$ we get:

$$\mathbb{E}[Y^s(\tau_S)] = \sum_{i=0}^{s}[x_i^s(\tau_s) \cdot p_s] + \sum_{i=s+1}^{S}[a_i \cdot p_i] \tag{44}$$

and

$$\mathrm{Var}(Y^s(\tau_S)) = \sum_{i=0}^{s}[x_i^s(\tau_s) \cdot p_s(1 - p_s)] + \sum_{i=s+1}^{S}[a_i \cdot p_i(1 - p_i)] \tag{45}$$

Note that $\sum_{i=0}^{s} x_i^s(\tau_s) = A_s^s$. We now determine $\mathrm{Var}(Y^{s+1}(\tau_S)) - \mathrm{Var}(Y^s(\tau_S))$:

$$\begin{aligned}
&\mathrm{Var}(Y^{s+1}(\tau_S)) - \mathrm{Var}(Y^s(\tau_S)) \\
&= \sum_{i=0}^{s+1}[x_i^{s+1}(\tau_{s+1}) \cdot p_{s+1}(1 - p_{s+1})] + \sum_{i=s+2}^{S}[a_i \cdot p_i(1 - p_i)] \\
&- \sum_{i=0}^{s}[x_i^s(\tau_s) \cdot p_s(1 - p_s)] - \sum_{i=s+1}^{S}[a_i \cdot p_i(1 - p_i)] \\
&= (p_{s+1}(1 - p_{s+1}) - p_s(1 - p_s)) \sum_{i=0}^{s}\left[(x_i^{s+1}(\tau_{s+1}) - x_i^s(\tau_s))\right] \\
&+ (x_{s+1}^{s+1}(\tau_{s+1}) - a_{s+1})p_{s+1}(1 - p_{s+1})
\end{aligned} \tag{46}$$

Since no new patients will arrive and there are per type i at most a_i patients, $x_i^{s+1}(\tau_{s+1})$ is smaller or equal to both $x_i^s(\tau_s)$ and a_i. However, since $p_{s+1} > p_s$, we do not know whether $p_{s+1}(1 - p_{s+1}) - p_s(1 - p_s)$ is larger or smaller than 0. Therefore, we cannot conclude that the difference in variance is smaller or equal to zero with absolute certainty.

3.3.2 A stochastic blueprint

Let us now consider the case where the number of surgeries is determined beforehand, but this number is a random variable. Note that once again $B_i^s = 0$. We now have:

$$N_i^s = A_i \tag{47}$$

Since this planning remains the same throughout time, there is no need for a prediction epoch index for A_i. We can now use (44) and (45) to determine the expectation and variance of $Y^s(\tau_S)$:

$$\mathbb{E}[Y^s(\tau_S)] = \mathbb{E}\left[\sum_{i=0}^{s}[x_i^s(\tau_s) \cdot p_s] + \sum_{i=s+1}^{S}[A_i \cdot p_i]\right] \tag{48}$$

$$\begin{aligned}\mathrm{Var}(Y^s(\tau_S)) =& \mathbb{E}\left[\sum_{i=0}^{s}[x_i^s(\tau_s) \cdot p_s(1-p_s)] + \sum_{i=s+1}^{S}[A_i \cdot p_i(1-p_i)]\right] \\ &+ \mathrm{Var}\left(\sum_{i=0}^{s}[x_i^s(\tau_s) \cdot p_s] + \sum_{i=s+1}^{S}[A_i \cdot p_i]\right)\end{aligned} \tag{49}$$

Since the planning of type i surgeries are likely to be independent, we now assume that A_i is Poisson distributed with mean δ_i. We also assume that the planning of type i surgeries are independent of the planning of surgeries of other types. This results in:

$$\mathbb{E}[Y^s(\tau_S)] = \sum_{i=0}^{s}[x_i^s(\tau_s) \cdot p_s] + \sum_{i=s+1}^{S}[\delta_i \cdot p_i] \tag{50}$$

$$
\begin{aligned}
\mathrm{Var}(Y^s(\tau_S)) &= \sum_{i=0}^{s}[x_i^s(\tau_s)\cdot p_s(1-p_s)] + \sum_{i=s+1}^{S}[\delta_i\cdot p_i(1-p_i)] \\
&+ \sum_{i=s+1}^{S}\left[p_i^2\cdot\delta_i\right] \qquad (51)\\
&= \sum_{i=0}^{s}[x_i^s(\tau_s)\cdot p_s(1-p_s)] + \sum_{i=s+1}^{S}[\delta_i\cdot p_i]
\end{aligned}
$$

where (51) also follows from the fact that $Y^s(\tau_S)$ now is a sum of a number of binomially distributed random variables and a number of Poisson distributed random variables.

We now determine the difference in variance:

$$
\begin{aligned}
&\mathrm{Var}(Y^{s+1}(\tau_S)) - \mathrm{Var}(Y^s(\tau_S)) \\
=&\sum_{i=0}^{s+1}[x_i^{s+1}(\tau_{s+1})\cdot p_{s+1}(1-p_{s+1})] + \sum_{i=s+2}^{S}[\delta_i\cdot p_i] \\
-&\sum_{i=0}^{s}[x_i^s(\tau_s)\cdot p_s(1-p_s)] - \sum_{i=s+1}^{S}[\delta_i\cdot p_i] \qquad (52)\\
=&\sum_{i=0}^{s}\left[x_i^{s+1}(\tau_{s+1})\cdot p_{s+1}(1-p_{s+1}) - x_i^s(\tau_s)\cdot p_s(1-p_s)\right] \\
+&\left(x_{s+1}^{s+1}(\tau_{s+1}) - \delta_{s+1}\right)p_{s+1}
\end{aligned}
$$

From here, we cannot conclude that the variance will decrease as time progresses.

3.3.3 A dynamic blueprint

Let us now consider the case that new surgeries can be scheduled into the future. Following (29) and (30), this implies that:

$$\begin{aligned} \mathbb{E}[N_i^s] &= \mathbb{E}[A_i^s] + \mathbb{E}[B_i^s] \\ &= A_i^s + \sum_{j=s+1}^{S} \lambda_i^j \end{aligned} \tag{53}$$

and

$$\begin{aligned} \mathbb{E}[N_i^{s+1}] &= \mathbb{E}[A_i^{s+1}] + \mathbb{E}[B_i^{s+1}] \\ &= \mathbb{E}[A_i^s + C_i^{s+1}] + \sum_{j=s+2}^{S} \lambda_i^j \\ &= A_i^s + \lambda_i^{s+1} + \sum_{j=s+2}^{S} \lambda_i^j \\ &= A_i^s + \sum_{j=s+1}^{S} \lambda_i^j \end{aligned} \tag{54}$$

and thus the following holds:

$$\mathbb{E}[N_i^s] = \mathbb{E}[N_i^{s+1}] \tag{55}$$

Since the N_i^s consist of two parts, one of which is a given and the other a Poisson distributed random variable, each $X_i^s(\tau_S)$ (except for $i = 0$), is split up into a binomial distribution and a Poisson distribution (see (4) and (7)). These distributions have a standard, known expectation and variance and thus, the expectation and variance of $Y^s(\tau_S)$ can be determined.

Let us first lose the assumption that all types have equal service rates and all time intervals have equal length for more generality. We then have:

$$\begin{aligned}\mathbb{E}[X_i^s(\tau_S)|A_i^s] &= \mathbb{E}_{B_i^s}[\mathbb{E}[X_i^s(\tau_S|A_i^s, B_i^s]] \\ &= \mathbb{E}_{B_i^s}[(A_i^s + B_i^s)e^{-\mu_{i-s}(\tau_S-\tau_i)}] \\ &= (A_i^s + \gamma^s)e^{-\mu_{i-s}(\tau_S-\tau_i)}\end{aligned} \tag{56}$$

and hence

$$\mathbb{E}[Y^s(\tau_S)|A_i^s, i = s \dots S] = \sum_{i=s}^{S} A_i^s \cdot e^{-\mu_{i-s}(\tau_S-\tau_i)} + \sum_{i=s+1}^{S} \gamma_i^s \cdot e^{-\mu_{i-s}(\tau_S-\tau_i)} \tag{57}$$

For the variance we have:

$$\begin{aligned}\mathrm{Var}(X_i^s|A_i^s, i = s \dots S) &= \mathbb{E}_{B_i^s}[\mathrm{Var}(X_i^s|A_i^s, B_i^s, i = s \dots S)] \\ &+ \mathrm{Var}_{B_i^s}(\mathbb{E}[X_i^s|A_i^s, B_i^s, i = s \dots S]) \\ &= \mathbb{E}_{B_i^s}[(A_i^s + B_i^s)e^{-\mu_{i-s}(\tau_S-\tau_i)}(1 - e^{-\mu_{i-s}(\tau_S-\tau_i)})] \\ &+ \mathrm{Var}_{B_i^s}\left((A_i^s + B_i^s)e^{-\mu_{i-s}(\tau_S-\tau_i)}\right) \\ &= (A_i^s + \gamma_i^s)e^{-\mu_{i-s}(\tau_S-\tau_i)}(1 - e^{-\mu_{i-s}(\tau_S-\tau_i)}) \\ &+ \gamma_i^s \cdot e^{-2\mu_{i-s}(\tau_S-\tau_i)} \\ &= A_i^s \cdot e^{-\mu_{i-s}(\tau_S-\tau_i)}(1 - e^{-\mu_{i-s}(\tau_S-\tau_i)}) \\ &+ \gamma_i^s \cdot e^{-\mu_{i-s}(\tau_S-\tau_i)}\end{aligned} \tag{58}$$

and hence

$$\begin{aligned}\mathrm{Var}(Y^s(\tau_S)|A_i^s, i = s \dots S) = &\sum_{i=s}^{S} A_i^s \cdot e^{-\mu_{i-s}(\tau_S-\tau_i)}(1 - e^{-\mu_{i-s}(\tau_S-\tau_i)}) \\ &+ \sum_{i=s+1}^{S} \gamma_i^s \cdot e^{-\mu_{i-s}(\tau_S-\tau_i)}\end{aligned} \tag{59}$$

We again now assume that the length of all time intervals and all service rates are equal, being T and μ respectively. This gives the following, simplified expressions for the expectation and variance of $Y^s(\tau_S)$.

$$\mathbb{E}[Y^s(\tau_S)|A_i^s, i = s \dots S] = \sum_{i=s}^{S} A_i^s \cdot e^{-\mu(S-i)T} + \sum_{i=s+1}^{S} \gamma_i^s \cdot e^{-\mu(S-i)T} \tag{60}$$

$$\begin{aligned} \mathrm{Var}(Y^s(\tau_S)|A_i^s, i = s \dots S) = &\sum_{i=s}^{S} A_i^s \cdot e^{-\mu(S-i)T}(1 - e^{-\mu(S-i)T}) \\ &+ \sum_{i=s+1}^{S} \gamma_i^s \cdot e^{-\mu(S-i)T} \end{aligned} \tag{61}$$

We can analyze the difference of the variances by updating the model by analyzing 61. For simplicity, we again denote :

$$p_i = e^{-\mu(S-i)T} \tag{62}$$

$$
\begin{aligned}
&\operatorname{Var}(Y^s(\tau_S)|A_i^s, i = s \dots S) \\
&-\operatorname{Var}(Y^{s+1}(\tau_S)|A_i^{s+1}, i = s+1, \dots, S, C_i^{s+1} = c_i^{s+1}) \\
&= \sum_{i=s}^{S} A_i^s \cdot p_i(1-p_i) + \sum_{i=s+1}^{S} \gamma_i^s \cdot p_i \\
&- \sum_{i=s+1}^{S} A_i^{s+1} \cdot p_i(1-p_i) - \sum_{i=s+2}^{S} \gamma_i^{s+1} \cdot p_i \\
&= A_s^s \cdot p_i(1-p_i) + \sum_{i=s+1}^{S} A_i^s \cdot p_i(1-p_i) + \gamma_{s+1}^s \cdot p_{s+1} \\
&+ \sum_{i=s+2}^{S} \gamma_i^s \cdot p_i - \sum_{i=s+1}^{S} A_i^s \cdot p_i(1-p_i) \\
&- \sum_{i=s+1}^{S} c_i^{s+1} \cdot p_i(1-p_i) - \sum_{i=s+2}^{S} \gamma_i^{s+1} p_i \\
&= A_s^s \cdot p_s(1-p_s) + \sum_{i=s+2}^{S} \sum_{j=s+1}^{S} \lambda_i^j \\
&+ \gamma_{s+1}^s p_{s+1} - \sum_{i=s+1}^{S} c_i^{s+1} \cdot p_i(1-p_i) - \sum_{i=s+2}^{S} \sum_{j=s+2}^{S} \lambda_i^j \\
&= A_s^s \cdot p_s(1-p_s) + \gamma_{s+1}^s p_s + \lambda_{s+2}^{s+1} - \sum_{i=s+1}^{S} c_i^{s+1} \cdot p_i(1-p_i)
\end{aligned} \tag{63}
$$

It is clear from (63) that there is no guarantee there will be a decrease in variance going from epoch τ_s to epoch τ_{s+1}, as it depends on the realization c_i^{s+1}.

3.4 Conclusion

We have come up with a model that can determine the expectation and variance of the number of inpatient patients for each of the various scheduling possibilities of the $M/M/\infty$ queue. This model distinguishes between patients that arrive in different time intervals. This ensures that we could use the information gain on these patients.

We see that for the none of the blueprints, there is a guaranteed drop in variance as the prediction epoch progresses.

4 The M/G/∞ queue

In this Section, we will discuss the $M/G/\infty$ model. First, we give some background information on the characteristics of this queue. Then, we do some analysis on a recursive approach of this queue, after which we will use both the background information and analysis to come up with the model.

We will also look at some specific service distributions for this model: the normal distribution and the log-normal distribution. The normal distribution is one step easier to analyze than the log-normal distribution, with the latter being common in healthcare discharge processes [16].

Section 6 analyzes the stochastic blueprint model, where the Length of Stay distribution is log-normal. Here, we compare a special update rule to the regular case.

4.1 Analysis on the M/G/∞ queue

Here, the discharge process has a general distribution. Denote by $Q(t)$ the number of patients inpatient at time t in a regular $M/G/\infty$ birth-death process and W the generic service-time random variable. Then, $Q(t)$ is Poisson distributed with mean [17] [18]:

$$\mathbb{E}[\text{number of patients at time } t] = \int_0^\infty [1 - H(x)]\,(\lambda(t-x)dx \qquad (64)$$

where

$$H(x) = P(\mathrm{W} \leq x) \qquad (65)$$

We consider a homogeneous arrival rate. We can set up the Kolmogorov equations for this model as follows [19]:

$$
\begin{aligned}
\frac{d}{dt}P_0(t) &= -\lambda(1-H(T-t))P_0(t) \\
\frac{d}{dt}P_n(t) &= \lambda(1-H(T-t))P_n(t) + \lambda(1-H(t))P_{n-1}(t), n \geq 1,
\end{aligned}
\tag{66}
$$

where $H(T-t)$ denotes the probability that a patient arriving at time t has completed its service by time T. This horizon T can later be replaced by our horizon τ_S. The initial conditions state that $P_0(0) = 1$ and $P_n(0) = 0$, $n \geq 1$. The initial condition that $P_0(0) = 1$ is not necessarily true, but we can model the patients that are already present at the ward separately, since we assume that there are infinite beds. This way it is also easier to take their residual Length of Stay into account. Additionally, the following also holds:

$$
\sum_{n=0}^{\infty} P_n(t) = 1 \tag{67}
$$

Combining (66) and (67) gives

$$
\begin{aligned}
\frac{d}{dt}P_0(t) &= \lambda(1-H(T-t))P_0(t) \\
\frac{d}{dt}P_n(t) &= \lambda(1-H(T-t))P_n(t) + \lambda(1-H(t))P_{n-1}(t), n \geq 1,
\end{aligned}
\tag{68}
$$

We solve these differential equations by setting up another differential equation using the probability generating function of $Q(t)$, denoted by $P(z,t)$.

$$
\frac{d}{dt}P(z,t) = -\lambda(1-H(T-t))(1-z)P(z,t) \tag{69}
$$

which has solution [19]

$$
P(z,t) = \exp\left(-\int_0^t \lambda(1-H(x))(1-z)dx\right) \tag{70}
$$

From this, we can determine the first two moments of the number of patients.

These moments directly relate to the expectation and variance.

4.2 Recursive approach

First, we approach the hospital without distinguishing between types of patients. Let N^s denote the total number of patients present at τ_s. We set up the following recursion:

$$N^{s+1} = N^s + E^{s+1} - D^{s+1}, 0 \leq s \leq S-1 \tag{71}$$

where E^{s+1} denotes the patients arriving between time epoch τ_s and τ_{s+1} and D^{s+1} denotes the patients leaving between τ_s and τ_{s+1}.

For each patient j, we will determine the probability of them being still present at time t, if they are present at time x. We denote this probability by $p_j(x,t)$. For now, we assume that they start being inpatient at time 0. We see that:

$$p_j(x,t) = \mathbb{P}(W_j \geq t | W_j \geq x) \tag{72}$$

Here, W_j denotes the random variable that is the service time of patient j. We denote the probability density function of the Length of Stay distribution of patient j by $f_j(t)$. Now:

$$p_j(x,t) = \frac{\mathbb{P}(W \geq t)}{\mathbb{P}(W \geq x)} = \frac{1 - \int_0^t f_j(y)dy}{1 - \int_0^x f_j(y)dy} \tag{73}$$

However, for $i > 0$, type i patients arrive in the hospital after time 0. Therefore, we have to shift the times. Let $p_j^i(x,t)$ denote the probability that a type i patient that is present at time x, $x \geq i$, is still present at time t. We see that:

$$p_j^i(x,t) = \frac{1 - \int_0^{t-\tau_i} f_j(y)dy}{1 - \int_0^{x-\tau_i} f_j(y)dy} \tag{74}$$

Let all the knowledge that is available at the time of making the prediction be denoted by $\mathbb{F}$ ($\mathbb{F}$ is a filtration [20]). Every time epoch, $\mathbb{F}$ is updated to contain all new information gained. So at time τ_s, we know the realization of N^s, which we denote by n^s. Additionally, we see that D^{s+1} is the sum of N^s independent Bernoulli trials, each with their own probability $p_j^{s+1}(\tau_s, \tau_{s+1})$. This implies that $N^s - D^{s+1}$ is the sum of N^s independent Bernoulli trials, each with their own probability $1 - p_j(\tau_s, \tau_{s+1})$. We denote these leftover patients by $L^{s+1} = N^s - D^{s+1}$. So we get:

$$N^{s+1} = L^{s+1} + E^{s+1} \tag{75}$$

We assume that the arrival process of patients in $[\tau_{s+1}, \tau_s)$, $s \leq S$, is Poisson distributed with mean λ^s. Given from (64), we then know that E^{s+1} is also Poisson distributed with mean $\lambda^{s+1} \cdot \mathbb{E}[W]$, which we denote by ρ^{s+1}.

We now apply the probability generating function to determine the distribution of N^{s+1}. For L^{s+1}, it holds that its generating function is the product of each individual Bernoulli trial. We define the probability generating function of L^{s+1}, given $N^s = n^s$ as follows:

$$G_{L^{s+1}}(n^s, z) = \mathbb{E}[z^{L^{s+1}} | N^s = n^s] \tag{76}$$

and thus

$$G_{L^{s+1}}(n^s, z) = \prod_{j=1}^{n^s} [q_j^{s+1}(\tau_s, \tau_{s+1}) + p_j^{s+1}(\tau_s, \tau_{s+1}) z], \tag{77}$$

where $[q_j^{s+1}(\tau_s, \tau_{s+1}) = 1 - p_j^{s+1}(\tau_s, \tau_{s+1})$. For N^{s+1} it then holds that:

$$\begin{aligned} G_{N^{s+1}}(n^s, z) &= G_{L^{s+1}}(n^s, z) \cdot G_{E^{s+1}}(z) \\ &= e^{-\lambda^s(1-z)} \cdot \prod_{j=1}^{n^s} [q_j(\tau_s) + p_j(\tau_s) z] \end{aligned} \tag{78}$$

We can get the expectation and variance for N^{s+1} directly from (71), since E^{s+1} and L^{s+1} are independent random variables. We know that E^{s+1} is Poisson distributed with mean ρ^{s+1}, implying its mean and variance are also ρ^{s+1}. The expectation and variance from L^{s+1} is simply the sum of the expectations and variances of each independent Bernoulli trial. We get:

$$\mathbb{E}[L^{s+1}] = \sum_{j=1}^{N^s}[p_j^{s+1}(\tau_s, \tau_{s+1})] \tag{79}$$

and

$$\mathrm{Var}(L^{s+1}) = \sum_{j=1}^{N^s}[p_j^{s+1}(\tau_s, \tau_{s+1}) \cdot q_j^{s+1}(\tau_s, \tau_{s+1})] \tag{80}$$

This implies:

$$\mathbb{E}[N^{s+1}|N^s = N^s] = \sum_{j=1}^{N^s}[p_j^{s+1}(\tau_s, \tau_{s+1})] + \rho^{s+1} \tag{81}$$

and

$$\mathrm{Var}(N^{s+1}|N^s = N^s) = \sum_{j=1}^{N^s}[p_j^{s+1}(\tau_s, \tau_{s+1}) \cdot q_j^{s+1}(\tau_s, \tau_{s+1})] + \rho^{s+1} \tag{82}$$

4.3 Modelling and updating the M/G/∞ queue

Again, let time be slotted, with prediction epochs $\tau_0, \tau_1, \ldots, \tau_S$. Our goal is to predict, at decision epoch τ_s, the number of patients present at time τ_S and relate these predictions at epochs τ_s and τ_{s+1}, $s = 0, \ldots, S-1$. We will refer to patients that undergo surgery and arrive in the ward in $(\tau_{i-1}, \tau_i]$ as type i patients and patients in the system at time 0 as type 0 patients. We assume that the arrival process of type i patients between decision epoch τ_s and τ_{i-1} is a Poisson process with mean λ_i^s (see Figures 1 and 2).

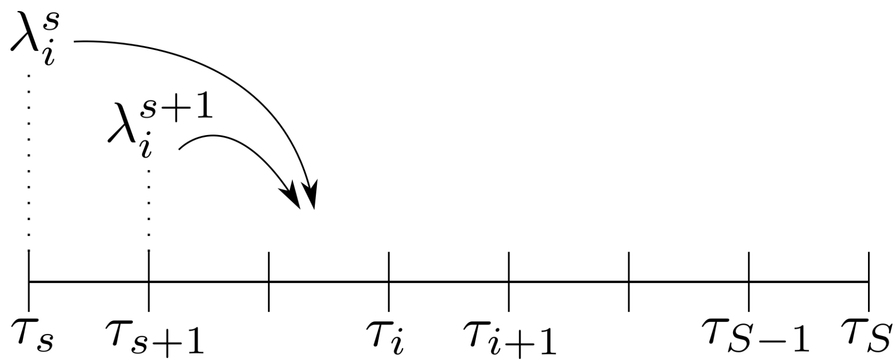

Figure 2: Overview of the arrival process at prediction epoch τ_s

Consider prediction epoch τ_s (see Figure 1). For each decision epoch τ_s, we will be again interested in the predicted number of patients present at epochs $\tau_s, \tau_{s+1}, \ldots, \tau_S$. Let N_0^0 denote the number of patients in the ward at epoch τ_0. Let N_i^s denote the predicted number of type i patients present in the ward at epoch τ_i for $i = s+1, \ldots, S$ and N_i^s for $i = 0, \ldots, s-1$ be the realization of the number of type i patients. Let N_s^s be the patients currently present in the ward. Let $X_i^s(t)$ be the (predicted) number of type i patients present at time t, $t \geq \tau_s$. So $N_i^s = X_i^s(\tau_i)$, $i = s, \ldots, S+1$. Let $Y^s(t) = \sum_{i=s}^{S} X_i^s(t)$, $t \geq \tau_s$ be the total (predicted) number of patients that are in the system at time t. The predicted number of patients of type i present in the ward at time τ_i, $i > s$, can be divided in two classes:

– A_i^s - the number of patients that are scheduled before time τ_s to undergo surgery in $(\tau_{i-1}, \tau_i]$.

– B_i^s - the number of patients that will be scheduled during $(\tau_s, \tau_{i-1}]$ to undergo surgery in $(\tau_{i-1}, \tau_i]$, but are not scheduled yet.

So

$$N_i^s = A_i^s + B_i^s \tag{83}$$

When time progresses and additional information is obtained about patients for surgery in $(\tau_{i-1}, \tau_i]$, A_i^s will increase and B_i^s will decrease. We assume that the service of a type i patient starts at τ_i. Let C_i^s denote the number of type i patients that will be scheduled in $(\tau_{s-1}, \tau_s]$ for surgery. We assume that C_i^s is Poisson distributed with mean λ_i^s. We see again that

$$A_i^{s+1} = A_i^s + C_i^{s+1} \tag{84}$$

and

$$B_i^s = C_i^{s+1} + B_i^{s+1} \tag{85}$$

Note that A_i^s, $i = s+1, \ldots, S$ will become a realization, since this value will be known once making the prediction. We see that B_i^s has a Poisson distribution with mean $\gamma_i^s = \sum_{j=s+1}^{S} \lambda_j^s$. Additionally, note that $B_i^i = 0$, since no more unknown patients will be scheduled. Combining these two implies that N_i^s will have a shifted Poisson distributed random variable, where the shift relies on the realization of A_i^s and the parameter (variance) is γ_i^s.

We are interested in $Y^s(\tau_S)$ and for that we need $X_i^s(\tau_S)$. For each of the N_i^s patients, we can use (74) in order to determine the probability that they are still there at τ_S. Hence, $X_i^s(\tau_S)$ is a sum of N_i^s independent Bernoulli trials, each with their own probability $p_j^i(\tau_i, \tau_S)$. This is called a Poisson binomial distribution [21]. We will now determine the expectation and variance of $Y^s(\tau_S)$ using various approaches for the surgery schedule. Since all $X_i^s(\tau_S)$, $i = 0 \ldots S$ are independent, we have the following:

$$\mathbb{E}[Y^s(\tau_S)] = \sum_{i=0}^{S} \mathbb{E}[X_i^s(\tau_S)] \tag{86}$$

and

$$\mathrm{Var}(Y^s(\tau_S)) = \sum_{i=0}^{S} \mathrm{Var}(X_i^s(\tau_S)) \tag{87}$$

4.3.1 A deterministic blueprint

We first assume that the entire surgery schedule has already been created. This implies that the only unknown is which patients will be undergoing surgery, but the number of patients who will be undergoing surgery is already known. A_i^s is now known and we denote the realization of A_i^s by a_i. Note that there is no necessity to include the epoch at which they are scheduled, since it will be constant. Also note that $B_i^s = 0$ for all $0 \leq i, s \leq S$. Now, $X_i^s(\tau_S)$ still has a Poisson binomial distribution, this time with parameters a_i and $p_j^i(\tau_i, \tau_S)$. We determine the expectation and variance of each $X_i^s(\tau_S)$. We see that:

$$\mathbb{E}[X_i^s(\tau_S)] = \begin{cases} \sum_{j=1}^{x_i^s(\tau_s)} \left[p_j^i(\tau_s, \tau_S)\right], & 0 \leq i \leq s \\ \sum_{j=1}^{a_i} \left[p_j^i(\tau_i, \tau_S)\right], & s < i < S \end{cases} \tag{88}$$

and

$$\mathrm{Var}\left(X_i^s(\tau_S)\right) = \begin{cases} \sum_{j=1}^{x_i^s(\tau_s)} \left[p_j^i(\tau_s, \tau_S)(1 - p_j^i(\tau_s, \tau_S))\right], & 0 \leq i \leq s \\ \sum_{j=1}^{a_i} \left[p_j^i(\tau_i, \tau_S)(1 - p_j^i(\tau_s, \tau_S))\right], & s < i < S \end{cases} \tag{89}$$

This implies:

$$\mathbb{E}[Y^s(\tau_S)] = \sum_{i=0}^{s} \left[\sum_{j=1}^{x_i^s(\tau_s)} \left[p_j^i(\tau_s, \tau_S)\right]\right] + \sum_{i=s+1}^{S} \left[\sum_{j=1}^{a_i} [p_j^i(\tau_i, \tau_S)]\right] \tag{90}$$

and

$$\begin{aligned} \mathrm{Var}(Y^s(\tau_S)) = &\sum_{i=0}^{s} \left[\sum_{j=1}^{x_i^s(\tau_s)} \left[p_j^i(\tau_s, \tau_S)(1 - p_j^i(\tau_s, \tau_S))\right]\right] \\ &+ \sum_{i=s+1}^{S} \left[\sum_{j=1}^{a_i} [p_j^i(\tau_i, \tau_S)(1 - p_j^i(\tau_i, \tau_S))]\right] \end{aligned} \tag{91}$$

We now look at the difference in variance of $Y^s(\tau_S)$ and $Y^{s+1}(\tau_S)$:

$$
\begin{aligned}
&\mathrm{Var}(Y^{s+1}(\tau_S)) - \mathrm{Var}\,(Y^s(\tau_S)) \\
=&\sum_{i=0}^{s+1}\left[\sum_{j=1}^{x_i^{s+1}(\tau_{s+1})}\left[p_j^i(\tau_{s+1},\tau_S)(1-p_j^i(\tau_{s+1},\tau_S))\right]\right] \\
+&\sum_{i=s+2}^{S}\left[\sum_{j=1}^{a_i}[p_j^i(\tau_i,\tau_S)(1-p_j^i(\tau_i,\tau_S))]\right] \\
-&\sum_{i=0}^{s}\left[\sum_{j=1}^{x_i^s(\tau_s)}\left[p_j^i(\tau_s,\tau_S)(1-p_j^i(\tau_s,\tau_S))\right]\right] \\
-&\sum_{i=s+1}^{S}\left[\sum_{j=1}^{a_i}[p_j^i(\tau_i,\tau_S)(1-p_j^i(\tau_i,\tau_S))]\right] \\
=&\sum_{i=0}^{s}\left[\sum_{j=1}^{x_i^{s+1}(\tau_{s+1})}[p_j^i(\tau_{s+1},\tau_S)(1-p_j^i(\tau_{s+1},\tau_S))]\right. \\
-&\left.\sum_{j=1}^{x_i^s(\tau_s)}[p_j^i(\tau_s,\tau_S)(1-p_j^i(\tau_s,\tau_S))]\right] \\
+&\sum_{j=1}^{x_{s+1}^{s+1}(\tau_{s+1})}[p_j^i(\tau_{s+1},\tau_S)(1-p_j^i(\tau_{s+1},\tau_S))] \\
-&\sum_{j=1}^{a_{s+1}}[p_j^i(\tau_{s+1},\tau_S)(1-p_j^i(\tau_{s+1},\tau_S))]
\end{aligned}
\tag{92}
$$

Since $a_{s+1} \geq x_{s+1}^{s+1}(\tau_{s+1})$, it is clear that

$$
\sum_{j=1}^{x_{s+1}^{s+1}(\tau_{s+1})}[p_j^i(\tau_{s+1},\tau_S)(1-p_j^i(\tau_{s+1},\tau_S))]-\sum_{j=1}^{a_{s+1}}[p_j^i(\tau_{s+1},\tau_S)(1-p_j^i(\tau_{s+1},\tau_S))] \leq 0
\tag{93}
$$

However, it is difficult to determine what happens to the first two terms. Using (74), it is clear that

$$p_j^i(\tau_{s+1}) \geq p_j^i(\tau_s) \tag{94}$$

We know that $x_i^s(\tau_s) \geq x_i s + 1(\tau_{s+1})$, $i = 0, \dots s$, but we can not say anything about $p_j^i(\tau_{s+1})(1 - p_j^i(\tau_{s+1}))$ compared with the previous time epoch.

We can now relate (90) and (91) to the results of the $M/M/\infty$ queue in (44) and (45) respectively. For an exponentially distributed Length of Stay with parameter μ_j, the memorylessness property results in:

$$p_j^i(x,t) = \frac{1 - \int_0^t \mu_j e^{-\mu_j y} dy}{1 - \int_0^x \mu_j e^{-\mu_j y} dy} = e^{-\mu_j(t-x)} \tag{95}$$

We then see that the probability of a type i patient staying in the hospital until τ_S becomes $e^{-\mu(\tau_S - \tau_i)}$. We now assume, similar to the $M/M/\infty$ queue, that all patients are independent, identically distributed with rate μ and all time epochs have equal length T. We apply this to (90) and (91):

$$\begin{aligned}
\mathbb{E}[Y^s(\tau_S)] &= \sum_{i=0}^{s} \left[\sum_{j=1}^{x_i^s(\tau_s)} \left[e^{-\mu(S-i)T} \right] \right] + \sum_{i=s+1}^{S} \left[\sum_{j=1}^{a_i} \left[e^{-\mu(S-i)T} \right] \right] \\
&= \sum_{i=0}^{s} \left[x_i^s(\tau_s) \cdot e^{-\mu(S-i)T} \right] + \sum_{i=s+1}^{S} \left[a_i \cdot e^{-\mu(S-i)T} \right] \qquad (96) \\
&= A_s^s \cdot e^{-\mu(S-s)T} + \sum_{i=s+1}^{S} \left[a_i \cdot e^{-\mu(S-i)T} \right]
\end{aligned}$$

$$
\begin{aligned}
\mathrm{Var}[Y^s(\tau_S)] &= \sum_{i=0}^{s}\left[\sum_{j=1}^{x_i^s(\tau_s)}\left[e^{-\mu(S-i)T}(1-e^{-\mu(S-i)T})\right]\right] \\
&+ \sum_{i=s+1}^{S}\left[\sum_{j=1}^{a_i}[e^{-\mu(S-i)T}(1-e^{-\mu(S-i)T})]\right] \\
&= \sum_{i=0}^{s}\left[x_i^s(\tau_s)\cdot e^{-\mu(S-i)T}(1-e^{-\mu(S-i)T})\right] \\
&+ \sum_{i=s+1}^{S}\left[a_i\cdot e^{-\mu(S-i)T}(1-e^{-\mu(S-i)T})\right] \\
&= A_s^s\cdot e^{-\mu(S-s)T}(1-e^{-\mu(S-i)T}) \\
&+ \sum_{i=s+1}^{S}\left[a_i\cdot e^{-\mu(S-i)T}(1-e^{-\mu(S-i)T})\right]
\end{aligned} \tag{97}
$$

where A_s^s again denotes all the patients present at τ_s. We see that these expressions are equal to (44) and (45) respectively.

4.3.2 A stochastic blueprint

Now consider that again there is a fixed blueprint for the hospital, but the blueprint itself is unknown. In other words, we no longer use the realization of A_i. We have

$$N_i^s = A_i \tag{98}$$

Again, there is no need for a prediction epoch index. We now have the following for the expectation and variance:

$$
\mathbb{E}[X_i^s(\tau_S)] = \begin{cases} \sum_{j=1}^{x_i^s(\tau_s)} p_j^i(\tau_s,\tau_S), & i \le s \\ \mathbb{E}\left[\sum_{j=1}^{A_i} p_j^i(\tau_i,\tau_S)\right], & i > s \end{cases} \tag{99}
$$

and

$$\mathrm{Var}(X_i^s(\tau_S)) = \begin{cases} \sum_{j=1}^{x_i^s(\tau_s)} p_j^i(\tau_s,\tau_S)(1-p_j^i(\tau_s,\tau_S)), & i \le s \\ \mathrm{Var}\left(\sum_{j=1}^{A_i} p_j^i(\tau_i,\tau_S)(1-p_j^i(\tau_i,\tau_S))\right), & i > s \end{cases} \tag{100}$$

We now assume that A_i is Poisson distributed with mean δ_i. The expectation of $X_i^s(\tau_S)$ is:

$$\mathbb{E}[X_i^s(\tau_S)] = \begin{cases} \sum_{j=1}^{x_i^s(\tau_s)} p_j^i(\tau_s,\tau_S), & i \le s \\ \sum_{a_i=1}^{\infty} \left[\frac{\delta_i^{a_i} e^{\delta_i}}{a_i!} \sum_{j=1}^{a_i} p_j^i(\tau_i,\tau_S)\right], & i > s \end{cases} \tag{101}$$

The variance of $X_i^s(\tau_S)$ for $i \le s$ is identical to (89):

$$\mathrm{Var}(X_i^s(\tau_S)) = \sum_{j=1}^{x_i^s(\tau_s)} [] p_j^i(\tau_s,\tau_S)(1-p_j^i(\tau_s,\tau_S))], i \le s \tag{102}$$

The variance for $i > s$ is more complicated. We see that:

$$\begin{aligned} \mathrm{Var}(X_i^s(\tau_S)) =& \mathbb{E}\left[\sum_{j=1}^{A_i} [p_j^i(\tau_i,\tau_S)(1-p_j^i(\tau_i,\tau_S))]\right] \\ &+ \mathrm{Var}\left(\sum_{j=1}^{A_i} [p_j^i(\tau_i,\tau_S)]\right) \\ &= \sum_{a_i=1}^{\infty}\left[\frac{\delta_i^{a_i}\cdot e^{-\delta_i}}{a_i!}\cdot \sum_{j=1}^{a_i} p_j^i(\tau_i,\tau_S)(1-p_j^i(\tau_i,\tau_S))\right] \\ &+ \sum_{a_i=1}^{\infty}\left[\frac{\delta_i^{a_i}\cdot e^{-\delta_i}}{a_i!}\cdot \left[\sum_{j=1}^{a_i} p_j^i(\tau_i,\tau_S)\right]^2\right] \\ &- \left(\sum_{a_i=1}^{\infty}\left[\frac{\delta_i^{a_i} e^{\delta_i}}{a_i!}\sum_{j=1}^{a_i} p_j^i(\tau_i,\tau_S)\right]\right)^2, i > s \end{aligned} \tag{103}$$

Looking at (103), the variance for $Y^s(\tau_S)$ becomes an even longer summation. Therefore, we now consider the case where we assume that type i patients have identical Length of Stay distributions. This implies that all $p_j^i(\tau_i, \tau_S)$ are equal, i.e. each patient of the same type has the same distribution with probability density function $f_i(t)$. We now denote $p_j^i(\tau_i, \tau_S)$ by $p_i(\tau_i, \tau_S)$. Using the analysis of a binomially distributed random variable with an initial Poisson distributed parameter (see Section 3.1), we then have:

$$\mathbb{E}[X_i^s(\tau_S)] = \begin{cases} x_i^s(\tau_s) \cdot p_i(\tau_s, \tau_S), & 0 \leq i \leq s \\ \delta_i \cdot p_i(\tau_i, \tau_S), & s < i \leq S \end{cases} \tag{104}$$

and

$$\mathrm{Var}(X_i^s(\tau_S)) = \begin{cases} x_i^s(\tau_s) \cdot p_i(\tau_s, \tau_S)(1 - p_i(\tau_s, \tau_S)), & 0 \leq i \leq s \\ \delta_i \cdot p_i(\tau_i, \tau_S), & s < i \leq S \end{cases} \tag{105}$$

We thus get for $Y^s(\tau_S)$:

$$\mathbb{E}[Y^s(\tau_S)] = \sum_{i=0}^{s} \left[x_i^s(\tau_s) \cdot p_i(\tau_s, \tau_S)\right] + \sum_{i=s+1}^{S} \left[\delta_i \cdot p_i(\tau_i, \tau_S)\right] \tag{106}$$

and

$$\begin{aligned} \mathrm{Var}\left(Y^s(\tau_S)\right) = &\sum_{i=0}^{s} [x_i^s(\tau_s) \cdot p_i(\tau_s, \tau_S)(1 - p_i(\tau_s, \tau_S))] \\ &+ \sum_{i=s+1}^{S} \left[\delta_i \cdot p_i(\tau_i, \tau_S)\right] \end{aligned} \tag{107}$$

We determine the difference in expectation:

$$\begin{aligned}
&\mathbb{E}[Y^s(\tau_S)] - \mathbb{E}[Y^{s+1}(\tau_S)] = \\
&\sum_{i=0}^{s} \left[x_i^s(\tau_s) \cdot p_i(\tau_s, \tau_S) - x_i^{s+1}(\tau_{s+1}) \cdot p_i(\tau_{s+1}, \tau_S)\right] \\
+&\sum_{i=s+1}^{S} \left[(\delta_i - \delta_i) p_i(\tau_i, \tau_S)\right] \\
+&\delta_{s+1} \cdot p_{s+1}(\tau_{s+1}, \tau_S) - x_{s+1}^{s+1}(\tau_{s+1}) \cdot p_{s+1}(\tau_{s+1}, \tau_S)
\end{aligned} \tag{108}$$

and the difference in variance:

$$\begin{aligned}
&\mathrm{Var}(Y^{s+1}(\tau_S)) - \mathrm{Var}(Y^s(\tau_S)) = \\
&\sum_{i=0}^{s} \Bigg[x_i^{s+1}(\tau_{s+1}) \cdot p_i(\tau_{s+1}, \tau_S)(1 - p_i(\tau_{s+1}, \tau_S)) \\
-&x_i^s(\tau_s) \cdot p_i(\tau_s, \tau_S)(1 - p_i(\tau_s, \tau_S))\Bigg] + \sum_{i=s+2}^{S} \left[(\delta_i - \delta_i) p_i(\tau_i, \tau_S)\right] \\
+&x_{s+1}^{s+1}(\tau_{s+1}) \cdot p_{s+1}(\tau_{s+1}, \tau_S)(1 - p_{s+1}(\tau_{s+1}, \tau_S)) - \delta_{s+1} \cdot p_{s+1}(\tau_{s+1}, \tau_S)
\end{aligned} \tag{109}$$

We cannot conclude that the variance will decrease. This is to be expected, as it also was not the case for the $M/M/\infty$ queue.

4.3.3 A dynamic blueprint

Now we assume that scheduling patients spontaneously is also possible. This applies that $B_i^s = 0$ no longer necessarily holds.

We determine the expectation and variance of $X_i^s(\tau_S)$, $i \geq s+1$, where we use the realization of A_i^s, which is denoted by a_i^s:

$$\mathbb{E}[X_i^s(\tau_S)] = \mathbb{E}\left[\mathbb{E}\left[X_i^s(\tau_S) | N_i^s\right]\right] = \mathbb{E}\left[\sum_{j=1}^{N_i^s} p_j^i(\tau_i, \tau_S)\right] \tag{110}$$

Since N_i^s has a shifted Poisson distribution with parameters A_i^s and γ_i^s, its probability density function is as follows:

$$\mathbb{P}(N_i^s = n) = \frac{e^{-\gamma_i^s} \gamma_i^{s(n-A_i^s)}}{(n - A_i^s)!} \tag{111}$$

Hence, we get:

$$\mathbb{E}[X_i^s(\tau_S)] = \sum_{n=a_i^s}^{\infty} \left[\frac{e^{-\gamma_i^s} \gamma_i^{s(n-a_i^s)}}{(n - a_i^s)!} \sum_{j=1}^{n} \left[p_j^i(\tau_i, \tau_S) \right] \right] \tag{112}$$

Regarding the variance of $X_i^s(\tau_S)$, we use the law of total variance. We have for $i > s$:

$$\begin{aligned}
&\mathrm{Var}(X_i^s(\tau_S)) \\
=&\mathbb{E}\left[\mathrm{Var}(X_i^s | N_i^s)\right] + \mathrm{Var}(\mathbb{E}[X_i^s | N_i^s]) \\
=&\mathbb{E}\left[\sum_{j=1}^{N_i^s} p_j^i(\tau_i, \tau_S)(1 - p_j^i(\tau_i, \tau_S)) | A_i^s = a_i^s \right] + \mathrm{Var}\left(\sum_{j=1}^{N_i^s} p_j^i(\tau_i, \tau_S) | A_i^s = a_i^s \right) \\
=&\sum_{n=a_i^s}^{\infty} \left[\frac{e^{-\gamma_i^s} \gamma_i^{s(n-a_i^s)}}{(n - a_i^s)!} \sum_{j=1}^{n} \left[p_j^i(\tau_i, \tau_S)(1 - p_j^i(\tau_i, \tau_S)) \right] \right] \\
&+ \sum_{n=a_i^s}^{\infty} \left[\frac{e^{-\gamma_i^s} \gamma_i^{s(n-a_i^s)}}{(n - a_i^s)!} \left(\sum_{j=1}^{n} p_j^i(\tau_i, \tau_S) \right)^2 \right] - \left(\sum_{n=a_i^s}^{\infty} \left[\frac{e^{-\gamma_i^s} \gamma_i^{s(n-a_i^s)}}{(n - a_i^s)!} \sum_{j=1}^{n} p_j^i(\tau_i, \tau_S) \right] \right)^2
\end{aligned} \tag{113}$$

For type i patients, $i \leq s$, who were already in the hospital at the time of making the prediction, it is slightly different. Since no new patients will arrive, we use the realization of $X_i^s(\tau_s)$, denoted by $x_i^s(\tau_s)$, and apply the residual Length of Stay for each of these patients as well. Note that this is different from the $M/M/\infty$ queue. So, for $i \leq s$, we get a Poisson binomial distribution with parameters $x_i^s(\tau_s)$ and $p_j^i(\tau_i, \tau_S)$, $j = 1 \dots x_i^s(\tau_s)$.

$$\mathbb{E}[X_i^s(\tau_S)] = \sum_{j=1}^{x_i^s(\tau_s)} p_j^i(\tau_s, \tau_S) \tag{114}$$

$$\mathrm{Var}(X_i^s(\tau_S)) = x_i^s(\tau_s)) = \sum_{j=1}^{x_i^s(\tau_s)} \left[p_j^i(\tau_s, \tau_S)(1 - p_j^i(\tau_s, \tau_S))\right] \tag{115}$$

We use these expressions to determine the total expectation and variance. We are especially interested in the width of the confidence interval, which is determined by the variance. The expectation is given by:

$$\begin{aligned} &\mathbb{E}\left[Y^s(\tau_S)|X_i^s(\tau_s) = x_i^s(\tau_s), i = 0 \dots s \text{ and } A_i^s = a_i^s, i = s+1 \dots S\right] \\ =&\sum_{i=0}^{s}\left[\sum_{j=1}^{x_i^s(\tau_s)} p_j^i(\tau_s, \tau_S)\right] + \sum_{i=s+1}^{S}\left[\sum_{n=a_i^s}^{\infty}\left[\frac{e^{-\gamma_i^s}\gamma_i^{s(n-a_i^s)}}{(n-a_i^s)!}\sum_{j=1}^{n}\left[p_j^i(\tau_i, \tau_S)\right]\right]\right] \end{aligned} \tag{116}$$

The variance is given by:

$$\begin{aligned} &\mathrm{Var}(Y^s(\tau_S)|X_i^s(\tau_s) = x_i^s(\tau_s), i = 0 \dots s \text{ and } A_i^s = a_i^s, i = s+1 \dots S) \\ =&\sum_{i=0}^{s}\left[\sum_{j=1}^{x_i^s(\tau_s)}\left[p_j^i(\tau_s, \tau_S)(1 - p_j^i(\tau_s, \tau_S))\right]\right] \\ +&\sum_{i=s+1}^{S}\left[\sum_{n=a_i^s}^{\infty}\left[\frac{e^{-\gamma_i^s}\gamma_i^{s(n-a_i^s)}}{(n-a_i^s)!}\sum_{j=1}^{n}\left[p_j^i(\tau_i, \tau_S)(1 - p_j^i(\tau_i, \tau_S))\right]\right]\right. \\ +&\sum_{n=a_i^s}^{\infty}\left[\frac{e^{-\gamma_i^s}\gamma_i^{s(n-a_i^s)}}{(n-a_i^s)!}\left(\sum_{j=1}^{n} p_j^i(\tau_i, \tau_S)\right)^2\right] \\ -&\left.\left(\sum_{n=a_i^s}^{\infty}\left[\frac{e^{-\gamma_i^s}\gamma_i^{s(n-a_i^s)}}{(n-a_i^s)!}\sum_{j=1}^{n} p_j^i(\tau_i, \tau_S)\right]\right)^2\right] \end{aligned} \tag{117}$$

If we would compare the difference in variance of $Y^{s+1}(\tau_S)$ and $Y^s(\tau_S)$, we would likely not be able to say with absolute certainty that it decreases. Since 117 is already a difficult expression, this will not be explicitly done.

We will now analyze what happens when $\gamma_i^s = 0$ for all i, s. The physical meaning behind this is equal to the deterministic blueprint. We then see that

$$\frac{e^{-\gamma_i^s}\gamma_i^{s(n-a_i^s)}}{(n-a_i^s)!} = \begin{cases} 1 & n = a_i^s \\ 0 & \text{elsewhere} \end{cases} \tag{118}$$

and thus

$$\mathbb{E}[Y^s(\tau_S)] = \sum_{i=0}^{s}\left[\sum_{j=1}^{x_i^s(\tau_s)} p_j^i(\tau_s, \tau_S)\right] + \sum_{i=s+1}^{S}\left[\sum_{j=1}^{a_i^s} p_j^i(\tau_i, \tau_S)\right] \tag{119}$$

and

$$\begin{aligned} \mathrm{Var}(Y^s(\tau_S)) = &\sum_{i=0}^{s}\left[\sum_{j=1}^{x_i^s(\tau_s)} \left[p_j^i(\tau_s, \tau_S)(1 - p_j^i(\tau_s, \tau_S))\right]\right] \\ &+ \sum_{i=s+1}^{S}\left[\sum_{j=1}^{a_i}[p_j^i(\tau_i, \tau_S)(1 - p_j^i(\tau_i, \tau_S))]\right] \end{aligned} \tag{120}$$

which are the same expressions as in 90 and 91, as is expected. Thus, setting up $\gamma_i^s = 0$ for all i, s and making A_i^s a random variable would yield the same result as for the stochastic blueprint.

4.4 Various service distributions

Research suggests that a log-normal service distribution is realistic in hospitals [16]. Since the normal distribution is very similar, we analyze that first.

4.4.1 Normal service distribution

We start by analysing patient j, who has not undergone surgery yet and has a Length of Stay, which has a normal distribution with mean μ_j and variance $2\sigma_j^2$. We see that [22]:

$$p_j^i(\tau_i, \tau_S) = 1 - \frac{1}{2}\left[1 + \text{erf}\left(\frac{(\tau_S - \tau_i) - \mu_j}{2\sigma_j}\right)\right] \tag{121}$$

After surgery happening at time τ_s, more is known about this type i patient and their variance decreases to σ_j^2. This leads to

$$p_j^i(\tau_s, \tau_S) = \frac{1 - \frac{1}{2}\left[1 + \text{erf}\left(\frac{(\tau_S - \tau_i) - \mu_j}{\sqrt{2}\sigma_j}\right)\right]}{1 - \frac{1}{2}\left[1 + \text{erf}\left(\frac{(\tau_s - \tau_i) - \mu_j}{\sqrt{2}\sigma_j}\right)\right]} \tag{122}$$

4.4.2 Log-normal service distribution

We analyze the log-normal service distribution in a similar way to Section 4.4.1, where we decrease the variance of patient j as soon as they have undergone their surgery. We again assume that the Length of Stay of patient j is distributed with parameters μ_j and $2\sigma_j^2$ before surgery. We then have [23]:

$$p_j^i(\tau_i, \tau_S) = 1 - \frac{1}{2}\left[1 + \text{erf}\left(\frac{\ln(\tau_S - \tau_i) - \mu_j}{2\sigma_j}\right)\right] \tag{123}$$

After patient j has undergone surgery at time τ_s and their Length of Stay distribution now has variance σ_j^2, their new probability of still being inpatient at time τ_S is:

$$p_j^i(\tau_s, \tau_S) = \frac{1 - \frac{1}{2}\left[1 + \text{erf}\left(\frac{\ln(\tau_S - \tau_i) - \mu_j}{\sqrt{2}\sigma_j}\right)\right]}{1 - \frac{1}{2}\left[1 + \text{erf}\left(\frac{\ln(\tau_s - \tau_i) - \mu_j}{\sqrt{2}\sigma_j}\right)\right]} \tag{124}$$

We see that these probabilities are very similar to (121) and (122). The choice to go from an initial variance of $2\sigma^2$ to σ^2 is also not arbitrary, but allows for an easier mathematical expression. In practicality, this can be determined from data. Due to the error function in (123) and (124), we will show numerical results of applying this model on the stochastic blueprint.

4.5 Conclusion

We see that for all blueprint models, we can not conclude with certainty that the variance decreases as the prediction epoch moves forward. For the stochastic and dynamic blueprint, we looked at the case where all patients of type i have identical Length of Stay distributions. We see that applying the $M/G/\infty$ model for an exponential Length of Stay distribution yields identical results to Section 3, which should be the case.

Furthermore, we looked at the Length of Stay distribution being normal and being log-normal. We introduced a special update rule, which Section 5 analyzes.

5 Numerical results

This Section shows the numerical results of our $M/M/\infty$ model and our $M/G/\infty$ model, where a log-normal service distribution is applied. Patients are generated, where the Length of Stay of each patient is drawn from the corresponding distribution. The simulation runs for a certain amount of time epochs. The expected value (prediction) of the number of inpatient patients in the ward and its variance (used for the prediction interval) are determined every time epoch, based on the information that is currently available. The prediction and prediction interval are plotted over time.

For both models, we first analyze the deterministic blueprint case. To create this blueprint, we generate random values for a_i, where the realization of these values are known during the entire simulation. After that, we analyze the stochastic blueprint case for both models. Here, the blueprint is determined using a Poisson distribution with mean δ_i, as is done in Sections 3 and 4.

For the log-normal service distribution simulation, we analyze the impact of having an update rule (see Section 4.4.2), by comparing it to making the prediction without update rule.

As for the prediction interval of significance level α, we use the following relation:

$$P(Y^s(\tau_S) > Y^{su}(\tau_S)) = P(Y^s(\tau_S) < Y^{sl}(\tau_S)) = \frac{\alpha}{2} \tag{125}$$

where $Y^{su}(\tau_S)$ and $Y^{sl}(\tau_S)$ are the upper and lower bound of the prediction interval respectively. We approximate this using the expectation and variance of $Y^s(\tau_S)$, by using:

$$Y^{su}(\tau_S) = \mathbb{E}[Y^s(\tau_S)] + z \cdot \sqrt{\mathrm{Var}(Y^s(\tau_S))} \tag{126}$$

and

$$Y^{sl}(\tau_S) = \mathbb{E}[Y^s(\tau_S)] - z \cdot \sqrt{\text{Var}(Y^s(\tau_S))} \tag{127}$$

where z is the z-value corresponding with the prediction interval with significance level α.

5.1 M/M/∞ model simulation blueprint

5.1.1 A deterministic blueprint

We start with simulating the $M/M/\infty$ queue with a deterministic blueprint. The blueprint itself is randomly created, where a_i follows a Poisson distribution with mean α_i. The realization of this blueprint is entirely known for the complete duration of the simulation. All time epochs have equal length of 1 time unit.

Table 1: Simulation of the $M/M/\infty$ model with $\mu_i = \mu = 0.05$, a deterministic blueprint with $\alpha_i = \alpha = 5$ and horizon $S = 10$.

	0	1	2	3	4	5	6	7	8	9		N_s^s
$\mathbb{E}[Y]$	51.99	52.24	52.76	52.98	53.19	54.36	52.53	51.54	50.81	48.51		48
$\text{Var}(Y)$	11.38	11.32	11.16	10.62	9.99	9.38	7.92	6.48	4.70	2.37		

We can see in Table 1 that the decrease in variance is most significant in the later part of the prediction. The running time for this small instance is almost instantaneous (0.0 seconds). For very large instances ($S = 1000$), where $\alpha_i = \alpha = 30$, it results in a running time of 33.93 seconds. As such, the running time is not a limiting factor. For this large instance, the same principle applies that the variance decreases the most when the prediction is made relatively late.

We will now look at more realistic scenarios. We increase the horizon (S) and we increase the average number of customers that arrive in each time interval.

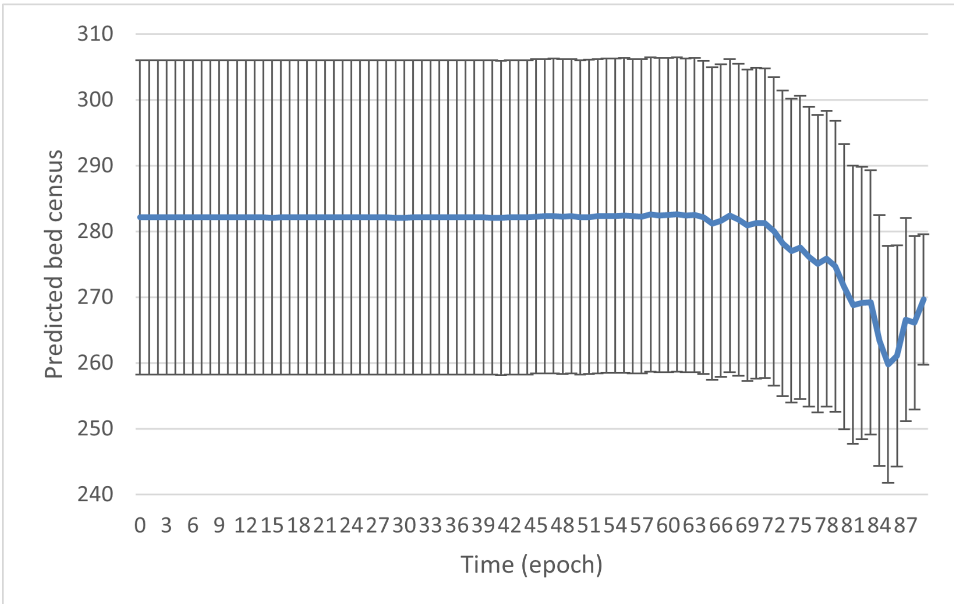

Figure 3: Prediction of the number of patients with a 95% prediction interval, using the $M/M/\infty$ model with $\mu_i = \mu = 0.1$, a deterministic blueprint with $\alpha_i = \alpha = 30$ and horizon $S = 90$. The realisation of $Y^S(\tau_S)$ equals 274.

The running time of the simulation in Figure 3 was 0.83 seconds. We chose for a horizon of $S = 90$, to simulate each time epoch being a day and predicting 3 months in advance. We again see that the variance (and thus the width of the prediction interval), as well as the predicted value stay roughly the same for the majority of the simulation. This is due to the fact that patients who were inpatient early are very likely to have already been discharged by time τ_S. To improve the insight gained by these simulations, we run another simulation with a smaller horizon S and higher relative service times (in comparison with the horizon).

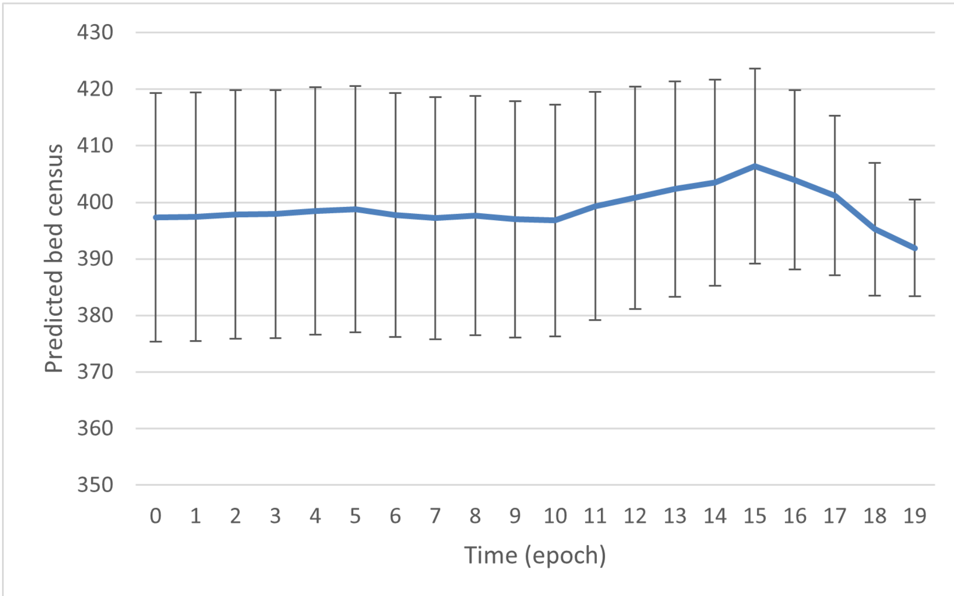

Figure 4: Prediction of the number of patients with a 95% prediction interval, using the $M/M/\infty$ model with $\mu_i = \mu = 0.05$, a deterministic blueprint with $\alpha_i = \alpha = 30$ and horizon $S = 20$. The realisation of $Y^S(\tau_S)$ equals 391. Titel aanpassen

In Figure 4, we see that the early behaviour of the prediction is still very stable. Especially the size of the prediction interval implies that there is little useful information gain. We can also see from the simulations that the variance does not necessarily has to decrease as the prediction epoch progresses. Furthermore, the prediction intervals of the early predictions are large compared with the prediction. In practice, this would imply that there will likely be too many nurses scheduled for work.

5.1.2 A stochastic blueprint

We will run simulations similar to the ones for the deterministic blueprint.

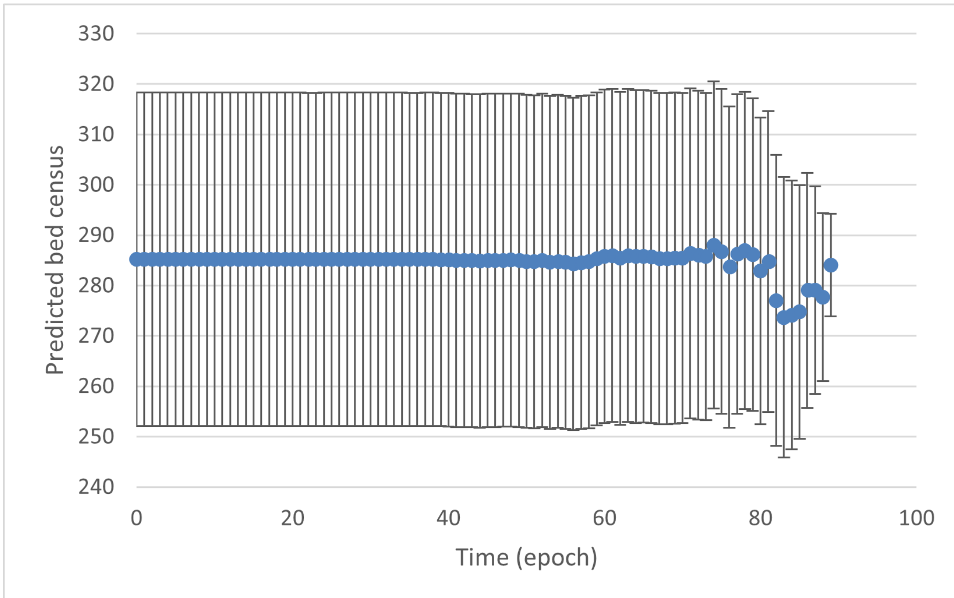

Figure 5: Prediction of the number of patients with a 95% prediction interval, using the $M/M/\infty$ model with $\mu_i = \mu = 0.1$, a Poisson distributed blueprint with $\delta_i = \delta = 30$ and horizon $S = 90$. The realisation of $Y^S(\tau_S)$ equals 281.

In Figure 5, we see similar behaviour comparing with the deterministic blueprint. The early predictions have little variation between them, both in the expected value and the variance of the prediction. Again, to gain better insights and to compare with the deterministic blueprint, we run the simulation with different parameters.

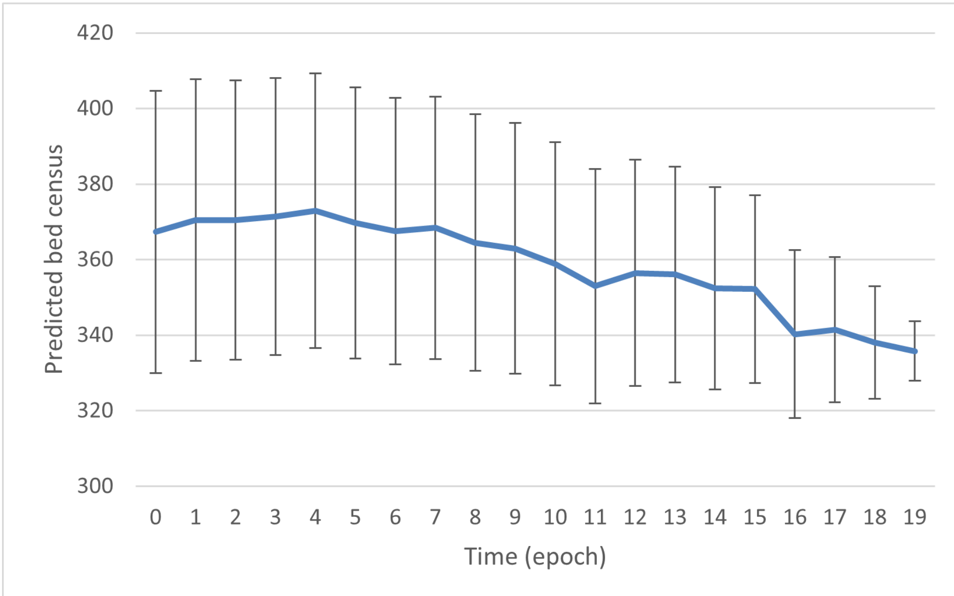

Figure 6: Prediction of the number of patients with a 95% prediction interval, using the $M/M/\infty$ model with $\mu_i = \mu = 0.05$, a Poisson distributed blueprint with $\delta_i = \delta = 30$ and horizon $S = 20$. The realisation of $Y^S(\tau_S)$ equals 333.

In Figure 6, we see that the prediction model behaves similarly to the deterministic blueprint. A noticeable difference is the difference in realizations of $Y^S(\tau_S)$. However, this is not impacted by the model, as the Length of Stay distribution of the patients remains the same. Therefore, we conclude that there is little difference in terms of using the deterministic or stochastic blueprint for the $M/M/\infty$ model.

5.2 M/G/∞ model simulation with log-normal service distribution

To compare with the $M/M/\infty$ model, we will do simulations with the same parameters as input, except that the variance of the service distribution is based upon real findings [16]. Here, the variance is roughly twice as large as the mean.

5.2.1 A deterministic blueprint

Similarly to the numerical approach of the $M/M/\infty$ model, we will first show the results of a large instance simulation of the $M/G/\infty$ model, using a log-normal distribution. We will show comparison between using the arbitrary update rule after patients underwent surgery and not using this update rule.

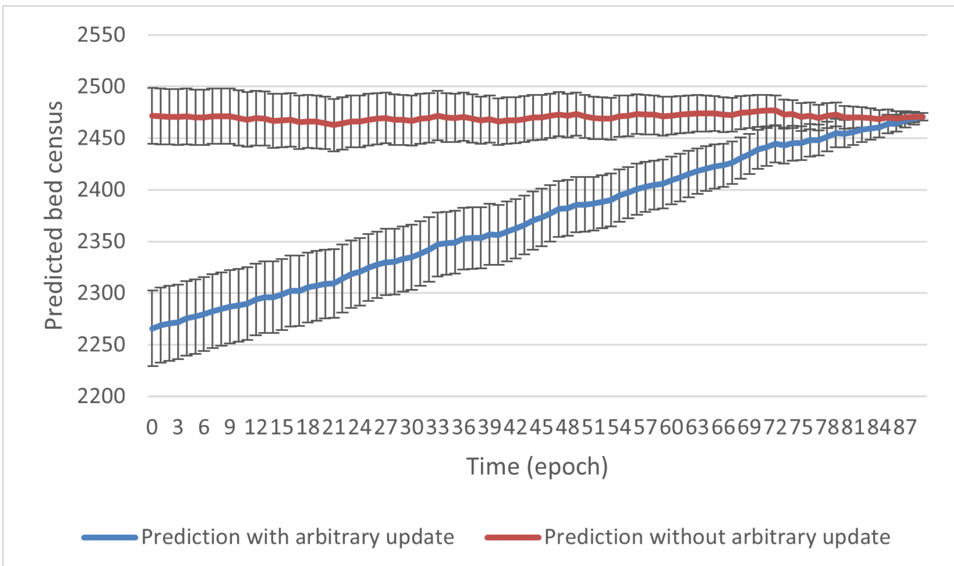

Figure 7: Prediction of the number of patients with a 95% prediction interval, using the $M/G/\infty$ model with $\mu_j = \mu = 10$, $\sigma_j^2 = \sigma^2 = 20$, a Poisson distributed blueprint with $\alpha_i = \alpha = 30$ and horizon $S = 90$. The realisation of $Y^S(\tau_S)$ equals 2472.

We can see that in Figure 7, there is a significant increase over time of the predicted value of the prediction made with the update rule. This implies that assuming a higher variance leads to a lower predicted value. Another point of interest is that the prediction interval of the prediction made with the update rule is larger compared with the prediction made without the update rule. This is to be expected, as a higher variance of the service distribution leads to more uncertainty about each patient.

Furthermore, there is a clear difference between the $M/M/\infty$ model simulation, as the prediction interval is relatively smaller for the $M/G/\infty$ model. However, increasing the variance to $\frac{1}{\mu^2} = 100$ leads to similar results (see Appendix C).

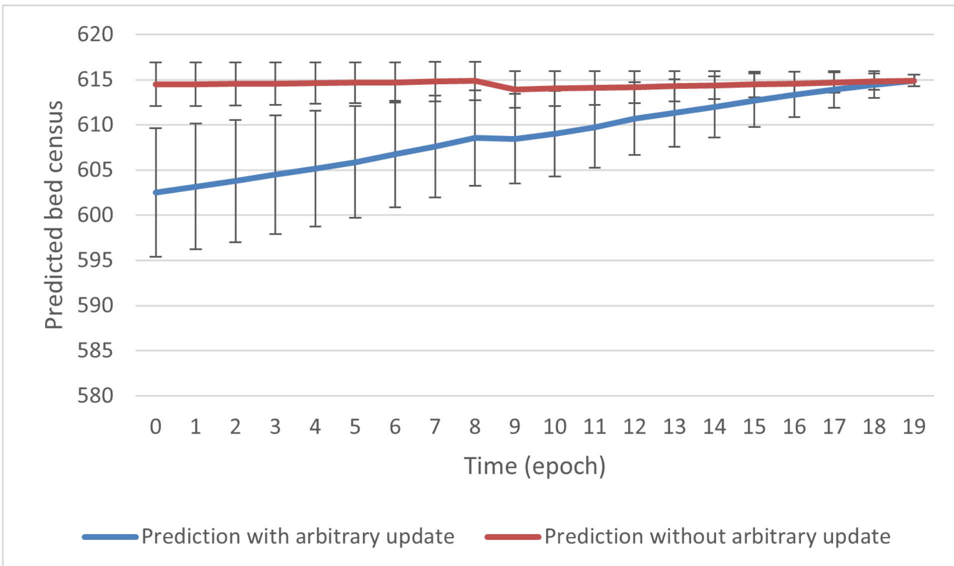

Figure 8: Prediction of the number of patients with a 95% prediction interval, using the $M/G/\infty$ model with $\mu_i = \mu = 20$, $\sigma_j^2 = \sigma^2 = 40$, a Poisson distributed blueprint with $\alpha_i = \alpha = 30$ and horizon $S = 20$. The realisation of $Y^S(\tau_S)$ equals 615.

In Figure 8, we can see that there is very little deviation throughout the prediction made without arbitrary update, which is dissimilar to the $M/M/\infty$ model. However, this might also be due to the prediction made at τ_0 being very close to the realization of $Y^S(\tau_S)$. Another simulation can be seen in Appendix C.

Similar to the large instance above, the prediction made with the update rule steadily increases as the prediction epoch progresses and has a bigger prediction interval everywhere compared with the prediction made without the update rule. Furthermore, we again see that the prediction interval of the $M/G/\infty$ model significantly smaller compared with the $M/M/\infty$ model. We did another simulation with $\sigma_j^2 = \sigma^2 = 400$ to more accurately compare with the $M/M/\infty$ queue. but similar results are found (see Appendix C).

5.2.2 A stochastic blueprint

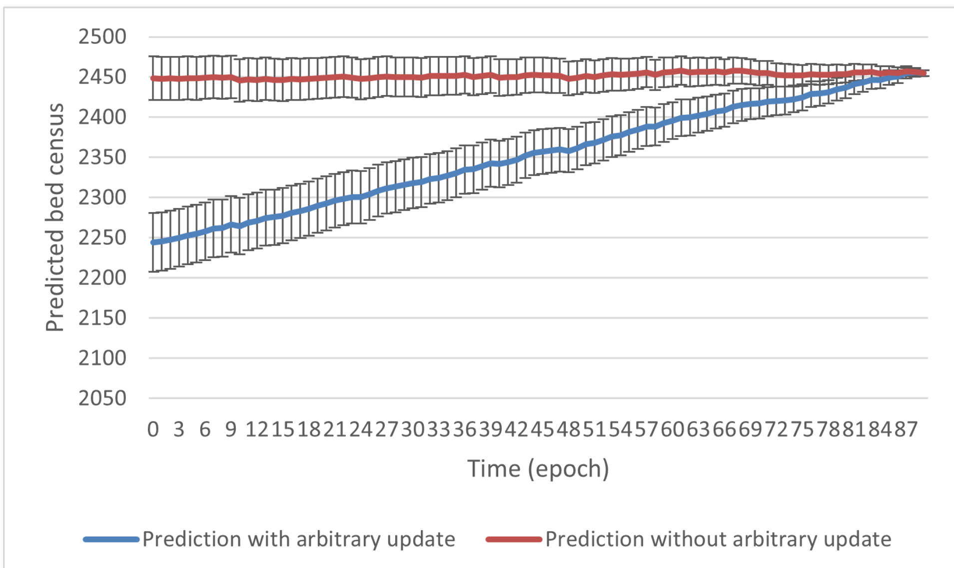

Figure 9: Prediction of the number of patients with a 95% prediction interval, using the $M/G/\infty$ model with $\mu_i = \mu = 10$, $\sigma_j^2 = \sigma^2 = 20$, a Poisson distributed blueprint with $\delta_i = \delta = 30$ and horizon $S = 90$. The realisation of $Y^S(\tau_S)$ equals 2455.

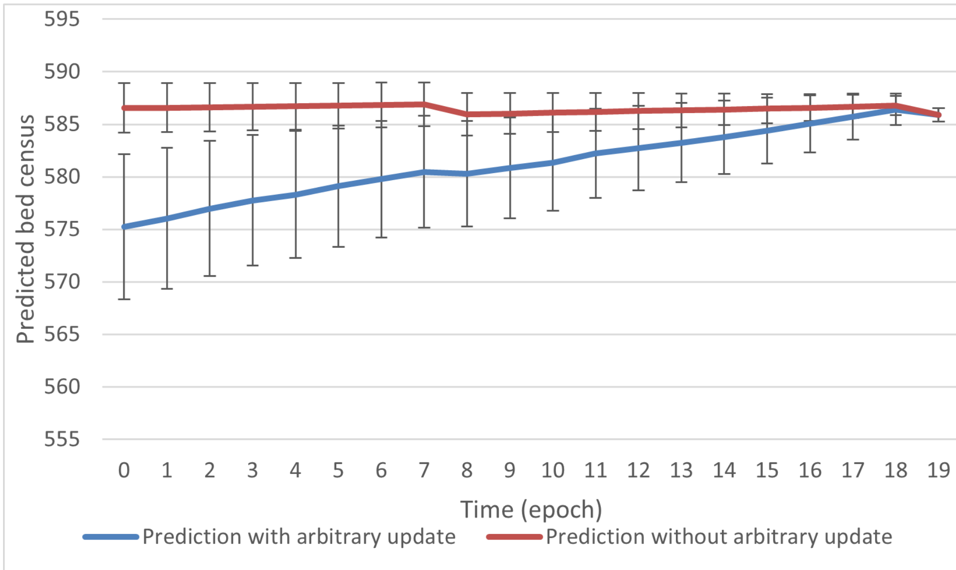

Figure 10: Prediction of the number of patients with a 95% prediction interval, using the $M/G/\infty$ model with $\mu_i = \mu = 20$, $\sigma_j^2 = \sigma^2 = 40$, a Poisson distributed blueprint with $\delta_i = \delta = 30$ and horizon $S = 20$. The realisation of $Y^S(\tau_S)$ equals 586.

Figures 9 and 10 bears a great resemblance to their deterministic counterparts (Figures 7 and 8). We see little difference in the size of the prediction intervals.

5.3 Conclusion

We have seen that the $M/M/\infty$ model results in a relatively large prediction interval for both the deterministic as the stochastic blueprint. The variance of the Length of Stay of each patient is also quite large due to the characteristics of the exponential distribution. However, we see that inputting similar values for the variance of the $M/G/\infty$ model with a log-normal distribution results in a relatively smaller prediction interval. In the $M/G/\infty$ case however, there is a significant increase in the bed census itself, especially for the large instance scenario.

For the $M/M/\infty$ model, there is a small difference in the size of the prediction interval between the deterministic and the stochastic blueprint. For the large instance, the stochastic model yields a larger prediction interval. This is expected, as there is more uncertainty regarding the number of patients that will undergo surgery. For the smaller instance, there is no significant difference.

We see that for the $M/G/\infty$ model that there is little difference in the deterministic and stochastic models. This is interesting, as the extra uncertainty that is provided by not knowing the number of patients that will undergo surgery does not have a significant impact.

Furthermore, we see that the using the arbitrary update rule for the log-normal distribution results in the prediction being too low. This can be explained by the fact that we did not update the expectation of these patients accordingly as well, by distinguishing between different types of surgeries.

6 Conclusion

6.1 Summary of results

The aim of this research was to create a model that is able to be updated to incorporate the new information gained to predict the number of patients being present in the ward in the future. We came up with two models, one for the $M/M/\infty$ queue and one for the $M/G/\infty$ queue. For each of these models, we successfully analyzed the impact of the additional information that is gained due to knowing the number of patients that are in the ward currently or are schedule for surgery. We further analyzed applying the normal distribution and the log-normal distribution for the Length of Stay on the $M/G/\infty$ model. We see that in no case of these models, a certain decrease in variance is obtained.

The usage of various ways of modelling the surgery schedule provided insights into the difference in the size of the prediction intervals. Distinguishing between patients that arrive in different time intervals helped gain those insights.

We have shown numerical results of our models. Here, we see that the variance is likely to decrease as the prediction epoch progresses. The results show that the information gain per time epoch depends on the parameters used for the simulation. Early on in the predictions, there is often little difference in the size of the prediction interval.

For the $M/M/\infty$ model, we see that using a deterministic blueprint yields a prediction where the size of the prediction interval is smaller. We cannot conclude that this is also the case for the $M/G/\infty$ model, where a log-normal distribution is applied.

6.2 Discussion

In this report we have done a theoretical analysis of the $M/M/\infty$ and the $M/G/\infty$ queues. The limitations of using these models in real life depend on the assumptions that are made in order to apply these models, e.g. assuming a Poisson arrival process or assuming that there are infinite beds. In order to make these models more applicable, we should analyze what happens when we lose these assumptions.

Furthermore, in the models we started the Length of Stay of a type i patient at time τ_i. This is not necessarily the case. This problem could be overcome by introducing a random variable that is uniformly distributed over their time interval. We expect that this solution has a relative small impact on the results of the model.

For the log-normal distribution, we introduced an update rule. However, this update rule is especially useful when looking at specific properties of a patient, e.g. type of surgery (arm, leg). Then, the uncertainty of every patient is bigger before surgery, as you do not know its type of surgery yet. This would imply that the probability density function of a patient has a certain probability to be equal to one distribution and a certain probability to be equal to another distribution.

6.3 Future research recommendations

As stated above, doing further analysis on more queues can provide more insights and general solutions, which can be used for real life models. The queues we recommend are $G/M/\infty$ and $G/G/\infty$. The analysis that is done in this report can be the foundation of analyzing those queues. Another recommendation is to look into the impact of having limiting beds (servers) available. The insights provided in this report are then limited.

Furthermore, one could research updating the parameters of the queues, by using Bayesian statistical models. This could provide more insight into the impact of the information gain by the type of surgery or the surgeon. Following up on that, we especially recommend researching the update rule, used in the log-normal Length of Stay distribution. The case where a patient has n possible log-normal distributions, each with probability p_n, where the distribution is known once a patient undergoes surgery could provide very useful in real life.

Lastly, we recommend showing numerical results of the dynamic blueprint, as the theoretical analysis has already been done for this. It might provide further insights for more acute surgeries.

7 Bibliography

[1] Paulien M. Bekker, René Koeleman. Scheduling admissions and reducing variability in bed demand. *Health Care Management Science*, 14(3): 237–249, 2011.

[2] Florin Gorunescu, Sally McClean, and Ph Millard. A queueing model for bed-occupancy management and planning of hospitals. *Journal of the Operational Research Society Journal of the Operational Research Society*, 53(53):19–24, 2002.

[3] Stef Baas, Sander Dijkstra, Aleida Braaksma, Plom van Rooij, Fieke J. Snijders, Lars Tiemessen, and Richard Boucherie. Real-time forecasting of covid-19 bed occupancy in wards and intensive care units. *Health care management science*, 24(2):402–419, 2021.

[4] P. R. Harper and A. K. Shahani. Modelling for the planning and management of bed capacities in hospitals. *The Journal of the Operational Research Society*, 53(1):11–18, 2002.

[5] Gary Harrison, Andrea Shafer, and Mark Mackay. Modelling variability in hospital bed occupancy. *Health care management science*, 8(4):325–334, 2005.

[6] Aleida Braaksma, Nikky Kortbeek, and Richard J. Boucherie. *Bed Census Predictions and Nurse Staffing*, pages 151–180. International Series in Operations Research and Management Science. Springer, 2021. ISBN 978-3-030-60211-6.

[7] Nikky Kortbeek, Aleida Braaksma, Ferry H.F. Smeenk, Piet J.M. Bakker, and Richardus J. Boucherie. Integral resource capacity planning for inpatient care services based on bed census predictions by hour. *Journal of the Operational Research Society*, 66(7):1061–1076, 2015.

[8] Van Houdenhoven M. Hurink J. L. Hans E. W. Wullink G. Kazemier G. van Oostrum, Jeroen M. A master surgical scheduling approach for cyclic scheduling in operating room departments.

[9] A. Goldenshluger. Nonparametric estimation of the service time distribution in the M/G/∞ queue. *Advances in Applied Probability*, 48(4): 1117–1138, 2016.

[10] H. Feng, P. Dube, and L. Zhang. On estimation problems for the G/G/∞ queue. *Performance Evaluation Review*, 39(3):40–42, 2011.

[11] J. Fang, J. Zhu, and X. Zhang. Prediction of length of stay on the intensive care unit based on bayesian neural network. *Journal of Physics: Conference Series*, 1631(1), 2020.

[12] Nikky Kortbeek, Aleida Braaksma, C.A.J. Burger, P.J.M Bakker, and Richardus J. Boucherie. Flexible nurse staffing based on hourly bed census predictions. *International journal of production economics*, 161: 167–180, 2015.

[13] A. Kolmogoroff. Über die analytischen methoden in der wahrscheinlichkeitsrechnung. *Mathematische Annalen*, 104:415–458, 1931.

[14] Vidyadhar G. Kulkarni. *Modeling and Analysis of Stochastic Systems*. Chapman & Hall, Ltd., 1st edition, 1995.

[15] William Feller & Vilim Feller. *An Introduction to Probability Theory and Its Applications, Volume 1*. Wiley, University of Michigan, 3rd edition, 1968.

[16] Gregory M Garrison and Jennifer L Pecina. Using the m/g/ queueing model to predict inpatient family medicine service census and resident workload. *Health Informatics Journal*, 22(3):429–439, 2016.

[17] Stephen G. Eick, William A. Massey, and Ward Whitt. The physics of the Mt/G/∞ queue. *Operations Research*, 41(4):731–742, 1993.

[18] G. F. Newell. The M/G/∞ queue. *SIAM Journal on Applied Mathematics*, 14, 1966.

[19] Khanik Kerobyan. Infinite-server M/G/∞ queueing models with catastrophes. July 2018.

[20] Achim Klenke. *Probability Theory*. Springer, London, 2nd edition, 2008.

[21] Harald Cramér. *Mathematical Methods of Statistics*. Princeton University Press, 1946.

[22] Eric W Weisstein. Normal distribution. mathworld.wolfram.com, 2020.

[23] Eric W Weisstein. Log normal distribution. mathworld.wolfram.com, 2020.

A State diagrams

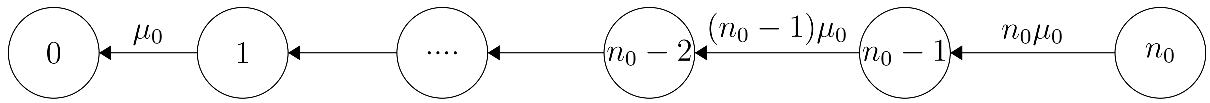

Figure 11: Pure-death process for the n_0 patients present at τ_0

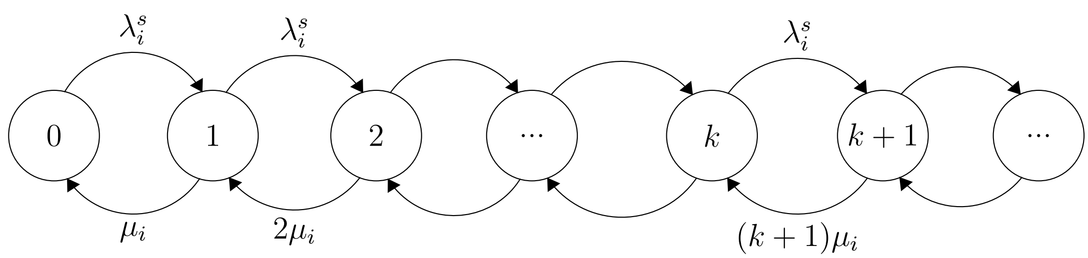

Figure 12: Birth-death process for type i patients in interval $[\tau_{i-1}, \tau_i)$

B Solutions of differential equations

Here, we show the derivations of several differential equations.

B.1 Solution for the pure-death process

(3) has an easy solution: $P_{n_0}^0 = e^{-n_0\mu_0 t}$. Plugging this into (2) for $k = n_0 - 1$ gives:

$$\frac{d}{dt}P_{n_0-1}^0 = n_0\mu_0 e^{-n_0\mu_0 t} - (n_0 - 1)\mu_0 P_{n_0-1}^0$$

which has solution $P_{n_0-1}^0 = n_0(e^{-\mu_0 t})^{n_0-1}(1 - e^{-\mu_0 t})$. Recursively, this gives the following potential solution:

$$P_k^0(t) = \binom{n_0}{k} e^{-\mu_0 kt}(1 - e^{-\mu_0 t})^{n_0-k}, k \geq 0 \tag{128}$$

which, if plugged into (2) and (3) solves the equation (note that $P_k^0(t)$ will be 0 for $k > n_0$).

B.2 Solution for the birth-death process

From (5) we get the following equation to solve for $\rho_1(t)$:

$$-\rho_1(t)'e^{-\rho_1(t)} = \mu_1\rho_1(t)e^{-\rho_1(t)} - \lambda e^{-\rho_1(t)} \tag{129}$$

$$-\rho_1(t)' = \mu_1\rho_1(t) - \lambda$$

This is split up into its homogeneous solution and its particular solution. The homogeneous solution is found as follows:

$$\rho_{1h}(t)' + \mu_1\rho_{1h}(t) = 0 \tag{130}$$

$$\rho_{1h}(t) = Ce^{-\mu_1 t}$$

The particular solution is very easily found by:

$$\rho_{1p}(t) = \frac{\lambda}{\mu_1} \tag{131}$$

which gives the solution for $\rho_1(t)$:

$$\rho_1(t) = \frac{\lambda}{\mu_1} + Ce^{-\mu_1 t} \tag{132}$$

Now we plug this $\rho_1(t)$ into (6) and solve for C.

$$-\rho_1'(t)e^{-\rho_1(t)}\frac{\rho_1^k(t)}{k!} + e^{-\rho_1(t)}\frac{k\rho_1^{k-1}(t)\rho_1'(t)}{k!} =$$

$$\lambda e^{-\rho_1(t)}\frac{\rho_1^{k-1}(t)}{(k-1)!} + (k+1)\mu_1 e^{-\rho_1(t)}\frac{\rho_1^{k+1}(t)}{(k+1)!} - (\lambda + k\mu_1)e^{-\rho_1(t)}\frac{\rho_1^k(t)}{k!} \tag{133}$$

This gives as result $C = \frac{\lambda}{\mu_1}$ and thus $\rho_1(t)$ equals (9).

C Simulation results

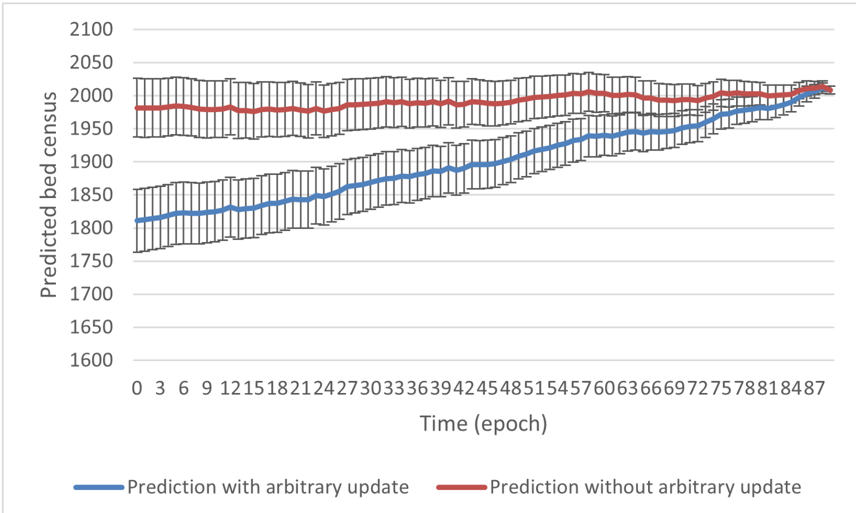

Figure 13: Prediction of the number of patients with a 95% prediction interval, using the $M/G/\infty$ model with $\mu_j = \mu = 10$, $\sigma_j^2 = \sigma^2 = 100$, a Poisson distributed blueprint with $\alpha_i = \alpha = 30$ and horizon $S = 90$. The realisation of $Y^S(\tau_S)$ equals 2011.

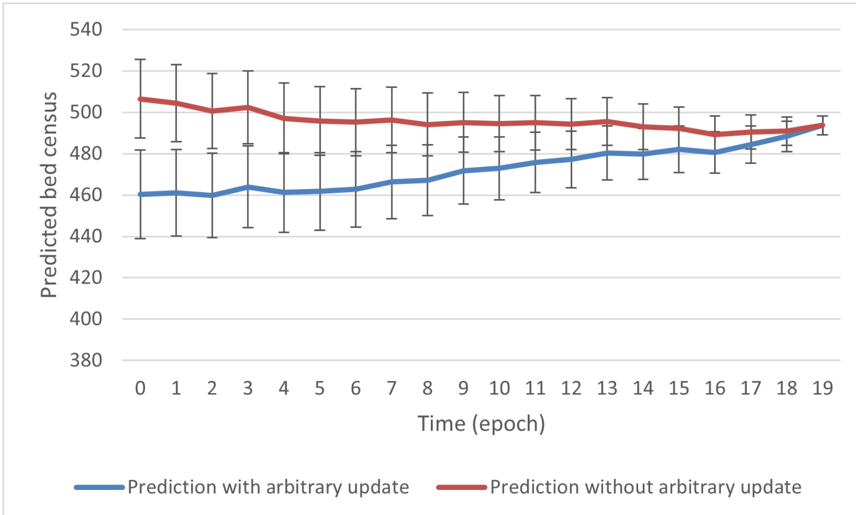

Figure 14: Prediction of the number of patients with a 95% prediction interval, using the $M/G/\infty$ model with $\mu_j = \mu = 20$, $\sigma_j^2 = \sigma^2 = 400$, a Poisson distributed blueprint with $\alpha_i = \alpha = 30$ and horizon $S = 90$. The realisation of $Y^S(\tau_S)$ equals 494.

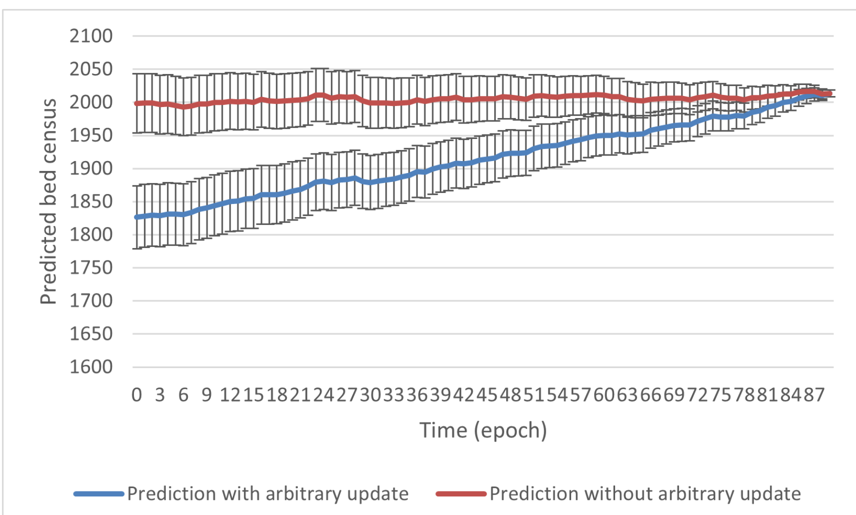

Figure 15: Prediction of the number of patients with a 95% prediction interval, using the $M/G/\infty$ model with $\mu_j = \mu = 10$, $\sigma_j^2 = \sigma^2 = 100$, a Poisson distributed blueprint with $\delta = \delta = 30$ and horizon $S = 90$. The realisation of $Y^S(\tau_S)$ equals 2014.

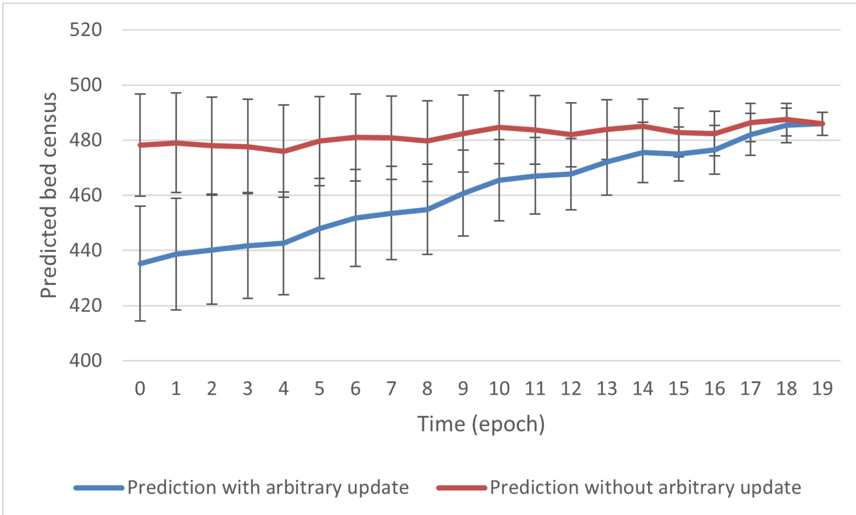

Figure 16: Prediction of the number of patients with a 95% prediction interval, using the $M/G/\infty$ model with $\mu_j = \mu = 20$, $\sigma_j^2 = \sigma^2 = 400$, a Poisson distributed blueprint with $\delta = \delta = 30$ and horizon $S = 90$. The realisation of $Y^S(\tau_S)$ equals 483.