

ESCUELA POLITÉCNICA NACIONAL

FACULTAD DE CIENCIAS

**MODELO DE TEORÍA DE COLAS Y SIMULACIÓN PARA LA
ASIGNACIÓN DEL RECURSO HUMANO EN UN SISTEMA
OPERATIVO DE SERVICIOS**

**TRABAJO DE TITULACIÓN PREVIO A LA OBTENCIÓN DEL TÍTULO DE
INGENIERO MATEMÁTICO**

PROYECTO DE INVESTIGACIÓN

WAGNER MARCELO SALAZAR GALLARDO

`wagner.salazar@epn.edu.ec`

Director: SANDRA ELIZABETH GUTIÉRREZ POMBOSA, PH.D.

`sandra.gutierrez@epn.edu.ec`

QUITO, JULIO 2022

DECLARACIÓN

Yo WAGNER MARCELO SALAZAR GALLARDO, declaro bajo juramento que el trabajo aquí escrito es de mi autoría; que no ha sido previamente presentado para ningún grado o calificación profesional; y que he consultado las referencias bibliográficas que se incluyen en este documento.

A través de la presente declaración cedo mis derechos de propiedad intelectual, correspondientes a este trabajo, a la Escuela Politécnica Nacional, según lo establecido por la Ley de Propiedad Intelectual, por su reglamento y por la normatividad institucional vigente.

Wagner Marcelo Salazar Gallardo

CERTIFICACIÓN

Certifico que el presente trabajo fue desarrollado por WAGNER MARCELO SALAZAR GALLARDO, bajo mi supervisión.

Sandra Elizabeth Gutiérrez Pombosa, PH.D.
Director del Proyecto

AGRADECIMIENTOS

A mi tutora Sandra, por el tiempo y esfuerzo dedicado a este proyecto. También, por ser la persona que me ayudó a crecer mi profesión en el campo laboral.

A mi profesor Ramiro Torres, por ser la persona que me infundió el agrado en la investigación operativa y la programación, siendo esta última una de las habilidades que más me costó desarrollar durante mi formación profesional.

A la Banda de los Gatos, por ser unos compañeros incondicionales en esta etapa universitaria y más que compañeros, amigos, en todo el sentido de la palabra.

A mis tíos y primos, que con su apoyo y palabras de motivación han sido un pilar importante durante mi formación profesional.

A todas las personas que estuvieron conmigo durante este trayecto, sepan que han sido de gran ayuda y me han permitido llegar hasta este momento.

DEDICATORIA

A mi madre, Norma, por tener siempre su confianza puesta en mí y hacerme saber que podía lograr este sueño que parecía distante, siempre de la mano de Dios. A mi padre, Marcelo, por brindarme el cariño, la motivación y el ejemplo para lograr lo que me propusiera y siempre dar un poco más de mí.

A mis hermanos Andrés y Matías, que nunca se rindan ni dejen de luchar por sus sueños.

Índice general

Resumen	VIII
Abstract	x
1. Estructura de un modelo de colas	1
1.1. Estructura básica de un modelo de colas	1
1.2. Proceso de colas elemental	3
1.3. Notaciones a usar en el proyecto	4
1.4. Relaciones entre L , W , L_q y W_q	5
1.5. Usos de modelos de teoría de colas	6
2. Modelos de teoría de colas con varios servidores	8
2.1. Proceso de nacimiento y muerte	8
2.2. Modelo M/M/s	11
2.3. Modelo M/D/s	14
2.4. Modelo M/G/s	17
3. Aplicación en un sistema operativo de servicios	21
3.1. Antecedentes	21
3.2. Depuración archivo de datos	28
3.2.1. Descripción del archivo	28
3.2.2. Tratamiento de datos atípicos	29
3.3. Ajuste de datos	32
3.4. Aplicación del modelo M/G/s	37

3.4.1. Cálculo de métricas principales	41
4. Simulación y Resultados	42
4.1. Simulación de la línea de espera	42
4.2. Evaluación de los resultados	44
4.3. Comparación costos actuales vs propuestos	45
5. Conclusiones y Recomendaciones	47
5.1. Conclusiones	47
5.2. Recomendaciones	48
Bibliografía	49
Anexos	51
5.3. Resultados y métricas	51
5.4. Simulaciones	55

Resumen

Las colas o líneas de espera forman parte de la vida cotidiana, pues a todos ha tocado esperar en algún momento en una fila de supermercado para pagar abastos o en una estación de buses para dirigirse a las labores; sin embargo, cuando estas filas son demasiado largas aparte de generar malestar en las personas tienen repercusiones económicas que afectan directamente a la empresa o entidad que esté brindando el servicio. Los dos problemas económicos principales que se pueden observar en una línea de espera son:

- i) Cuando hay demasiados servidores para brindar un servicio, aunque la cola será corta, se tendrían gastos innecesarios y excesivos de nómina.
- ii) Cuando no hay suficientes servidores la cola se vuelve muy larga, lo que conlleva a que clientes abandonen el servicio generando pérdidas para la empresa.

Los modelos de teoría de colas buscan capturar el comportamiento de una línea de espera y brindarnos información de dicho comportamiento como: número de clientes en el sistema, el tiempo de espera en el sistema, entre otros; con los cuales se puede trabajar para encontrar el mejor número de servidores y encontrar un balance entre el costo de servicio y la cantidad de espera.

Los modelos de teoría de colas son aplicables en distintos ámbitos y tienen diferentes utilidades y no solo se limitan al modelo *cliente esperando un servicio*; por ejemplo, en un aeropuerto la espera adecuada para los despegues ayuda a que no se retrasen vuelos, mientras que en trabajos de manufactura la asignación óptima de las máquinas ayuda a maximizar la producción.

Por otro lado, se considera un sistema operativo de servicios a cualquier entidad que brinda un servicio o asesoría a las personas; en el cual los clientes llegan, se les asigna un turno y esperan un tiempo hasta poder ser atendidos.

Para nuestro estudio se cuenta con una archivo de registro del Servicio de Rentas Internas (S.R.I), el cual contiene las horas llegadas y el tiempos de servicio desde

agosto de 2016 hasta mayo de 2018 con 53 oficinas y 33 tipos de servicios, donde se observan tiempos de espera en cola muy altos, los cuales han generado malestar en los clientes. Con el fin de solventar este problema se ha establecido una política calidad que consiste en reducir los tiempos de espera en fila para los diferentes servicios, de tal manera que estos sean a lo sumo de 20 minutos en las diferentes horas del día. Al analizar los tiempos de espera por oficinas se ha observado que la Plataforma Gubernamental Amazonas (PGA, agencia del Servicio de Rentas Internas S.R.I) no cumple esta política de calidad y además es la que más concurrencia de clientes tiene, por tanto, se escoge como caso de estudio para este trabajo.

Ahora bien, se ha evidenciado (gracias a un histórico del número de ventanillas abiertas por hora) que de las 21 ventanillas de servicio con las que se cuenta en la PGA se abre un número fijo entre 18 o 19 ventanillas en todas las horas del día; por lo cual, el aporte que brindará este proyecto de titulación es proponer un modelo de teoría de colas con el cual se puede explicar el comportamiento de este sistema de operaciones para poder decidir un número de ventanillas a abrir en cada hora del día de tal manera que se cumpla la política de calidad y analizar si se puede tener un ahorro en las horas de poca afluencia.

Para la elección de este modelo se usarán pruebas de bondad de ajuste para encontrar distribuciones de probabilidad a las que se ajusten los tiempos de llegada y de servicio y con ellos elegir un modelo de teoría de colas adecuado. Para validar resultados se realizarán simulaciones usando el número de servidores sugerido por el modelo, luego se extraerán los percentiles 1 y 99 de los resultados de la simulación para contrastar que al menos 98 % de las veces, el tiempo de espera en cola, sea menor a 20 minutos.

El resultado de este estudio muestra que el comportamiento de la línea de espera de la PGA se puede aproximar mediante un modelo de colas donde los tiempos entre llegadas se ajustan a una distribución exponencial y los tiempos de servicio se ajustan a una distribución Weibull, además se propone un número de servidores en cada hora del día que son: 11 servidores a las 8:00, 15 servidores a las 9:00, 18 servidores a las 10:00, 20 servidores a las 11:00, 16 servidores a las 12:00, 14 servidores a las 13:00, 14 servidores a las 14:00, 16 servidores a las 15:00 y 14 servidores a las 16:00. Con estos valores se logra que los tiempos de espera en fila sean menores a los 20 minutos y se tiene un ahorro en horas del 16,36 %.

Abstract

Queues or waiting lines are part of everyday life, since everyone has had to wait some day in a supermarket line to pay for groceries or at a bus station to go to work; however, when these lines are too long, apart from generating discomfort in people, they have economic repercussions that directly affect the company or entity that is providing the service. The two main economic problems that can be observed in a waiting line are:

- i) When there are too many servers to provide a service, although the queue will be short, there would be unnecessary and excessive payroll expenses.
- ii) When there are not enough servers, the queue becomes very long, which leads to clients abandoning the service, generating losses for the company.

Queuing theory models seek to capture the behavior of a waiting line and provide us with information on said behavior such as: number of customers in the system, waiting time in the system, among others; with which you can work to find the best number of servers and find a balance between the cost of service and the amount of waiting.

Queuing theory models are applicable in different fields and have different uses and are not only limited to the model *client waiting for a service*; For example, in an airport, adequate waiting for takeoffs helps to avoid delaying flights, while in manufacturing jobs, the optimal allocation of machines helps to maximize production.

On the other hand, any entity that provides a service or advice to people is considered an operating system of services; in which customers arrive, are assigned a shift and wait a while until they can be served.

For our study, there is a registry file of the Internal Revenue Service (S.R.I), which contains the service hours arrived and the service times from August 2016 to May 2018 with 53 offices and 33 types of types, where They observe very high waiting times in queue, which have generated discomfort in customers. In order to solve this

problem, a quality policy has been established that consists of reducing the waiting times in line for the different services, in such a way that these are at most 20 minutes at different times of the day. When analyzing the waiting times by offices, it has been observed that the Amazonas Governmental Platform (PGA, agency of the Internal Revenue Service S.R.I) does not comply with this quality policy and is also the one with the highest number of clients, therefore, it is chosen as case study for this work.

Now, it has been shown (thanks to a history of the number of windows open per hour) that of the 21 service windows that the PGA has, a fixed number between 18 or 19 windows are opened at all hours of the day ; Therefore, the contribution that this degree project will provide is to propose a queuing theory model with which the behavior of this operations system can be explained in order to decide a number of windows to open at each hour of the day in such a way that the quality policy is complied with and analyze whether savings can be achieved during low-peak hours.

For the choice of this model, goodness-of-fit tests will be used to find probability distributions to which the arrival and service times fit and with them choose an appropriate queuing theory model. To validate the results, simulations will be carried out using the number of servers suggested by the model, then the 1st and 99th percentiles of the simulation results will be extracted to verify that at least 98 % of the time, the waiting time in the queue is less than 20 minutes.

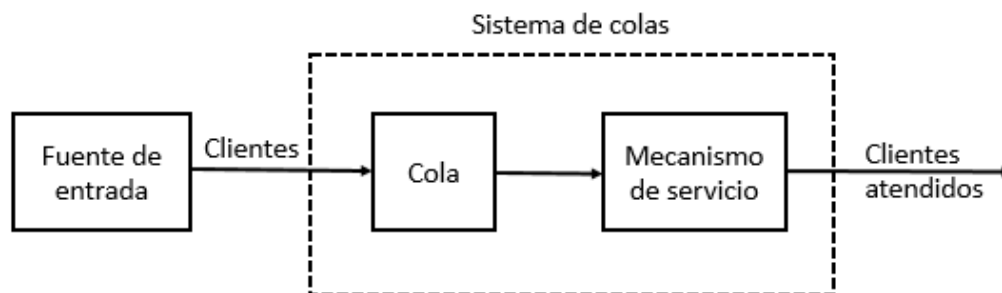
The result of this study shows that the behavior of the PGA waiting line can be approximated by means of a queuing model where the times between arrivals are adjusted to an exponential distribution and the service times are adjusted to a Weibull distribution, in addition, proposes a number of servers in each hour of the day which are: 11 servers at 8:00, 15 servers at 9:00, 18 servers at 10:00, 20 servers at 11:00, 16 servers at 12:00, 14 servers at 1:00 p.m., 14 servers at 2:00 p.m., 16 servers at 3:00 p.m., and 14 servers at 4:00 p.m. With these values it is achieved that waiting times in line are less than 20 minutes and there is a saving in hours of 16,36 %.

Capítulo 1

Estructura de un modelo de colas

1.1. Estructura básica de un modelo de colas

El proceso básico de colas consta de cuatro elementos principales que son: fuente de entrada, cola, disciplina de la cola y mecanismo de servicio.



Autor: Salazar W. Fuente: Hillier [7]

Figura 1.1 Proceso básico de colas

La Figura 1.1 muestra que para recibir un servicio los clientes se generan en una *f fuente de entrada*, posteriormente pasan a formar parte de una *cola*; ya en la cola el cliente a ser atendido se elige mediante la *disciplina de la cola* y su requerimiento es solventado mediante un *mecanismo de servicio*. Se detalla más sobre los elementos a continuación.

FUENTE DE ENTRADA

Es el lugar donde se generan los clientes potenciales a lo largo del tiempo, una característica importante de la fuente de entrada es el tamaño, es decir, el número de

clientes en la fuente de entrada. Puede suponerse que el tamaño es finito o infinito, en general el caso de tamaño infinito (cuando el tamaño es un número fijo relativamente grande) es más sencillo de tratar y es una suposición que se debe hacer en caso de que no se impongan restricciones sobre la fuente de entrada. El caso de tamaño finito es más complejo pues se debe tomar en cuenta la influencia que tiene el número de clientes en el sistema, en cualquier instante de tiempo, con el tamaño de la fuente de entrada. De aquí los clientes pasan a formar parte de la cola y un elemento importante a tener en cuenta es el *tiempo entre llegadas* (tiempo entre dos llegadas consecutivas).

COLA

Es el lugar donde esperan los clientes antes de recibir el servicio, una característica importante de la cola es el tamaño, es decir, el número máximo permisible de clientes que puede admitir. Las colas pueden ser finitas o infinitas, según si dicho número es finito o infinito. El caso de tamaño infinito es una suposición a hacerse incluso si la cola presenta una cota superior (relativamente grande) para el número permitido de clientes. El caso de tamaño finito es más complejo, y se entra en dicho caso cuando la cota superior es relativamente pequeña y se llega a dicho valor con frecuencia.

DISCIPLINA DE LA COLA

Es el orden en que se elige un cliente de la cola para ser atendido, existen algunos tipos entre los cuales destacan:

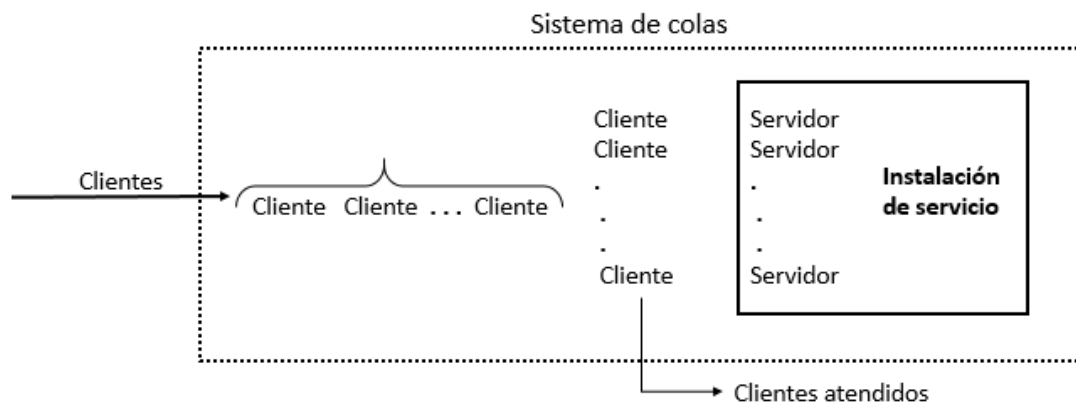
- **Disciplina FIFO** (First-In-First-Out): El cliente seleccionado para ser atendido es el primero en haber llegado a la cola.
- **Disciplina LIFO** (Last-In-First-Out): El cliente seleccionado para ser atendido es el último en haber llegado a la cola.
- **Disciplina RSS** (Random-Selection-Service): El cliente seleccionado para ser atendido es elegido de manera aleatoria.

MECANISMO DE SERVICIO

Es el procedimiento por el cual, una vez elegido el cliente a ser atendido, se le brinda el servicio solicitado. Existen dos elementos necesarios para determinar el mecanismo de servicio; la distribución de probabilidad a la que se ajusta el *tiempo de servicio* (tiempo transcurrido desde el inicio hasta la conclusión del servicio) y el número de servidores de dicho mecanismo. Existen los casos en que cada servidor ofrece un servicio particular, es decir, los tiempos de servicio serán diferentes para cada servidor; en este caso se necesita conocer la distribución a la que se ajusta el tiempo de servicio de cada servidor.

1.2. Proceso de colas elemental

El proceso de colas notado como elemental, es aquel que se ajusta al siguiente esquema:



Autor: Salazar W. **Fuente:** Hillier [7]

Figura 1.2 Sistema de colas elemental

La Figura 1.2 indica una única línea de espera formada frente a una estación de servicio con uno o varios servidores.

Ahora bien, extendiendo esta idea podemos considerar que:

- El número de estaciones de servicio no debe ser única, pues puede haber más estaciones como en el caso de ventanillas en un banco.
- Los servidores y clientes no necesariamente deben ser personas, por ejemplo, en el caso de un auto-lavado los clientes son autos y los servidores son las

diferentes estaciones de lavado.

- No es necesario una línea física delante de una estructura para recibir el servicio, por ejemplo; en el caso un hospital las habitaciones de los pacientes están distribuidas dentro de un área y los pacientes esperan que los médicos (servidores) se acerquen a ellos para brindarles atención.

El requisito elemental para poder hacer uso de teoría de colas es que los cambios en el número de clientes que esperan un servicio ocurran como si prevaleciera el esquema de la Figura 1.2.

1.3. Notaciones a usar en el proyecto

En Hillier [7] se encuentra la notación más relevante a usar en el proyecto, que se resume a continuación:

- $A/B/s$ = notación Kendall (Kendall [8]) para el tipo de modelo de colas.
 - A: distribución del tiempo entre llegadas.
 - B: distribución del tiempo de servicio.
 - s: número de servidores.
- n = Estado del sistema (número de clientes en el *sistema*).
- Sistema = conjunto de los clientes en cola y los clientes siendo atendidos.
- $N(t)$ = número de clientes en el sistema al tiempo t ($t \geq 0$).
- $P_n(t)$ = probabilidad de que exactamente n clientes estén en el sistema en el tiempo t .
- λ = tasa de llegada de clientes (número de clientes que llegan al sistema por unidad de tiempo).
- μ = tasa de servicio a clientes (número de clientes que completan su servicio por unidad de tiempo).
- $\rho = \lambda/(s\mu)$. *Factor de utilización*, es decir, fracción esperada de tiempo que los servidores individuales están ocupados.

Cuando un sistema de colas inicia su operación el estado del sistema está influenciado por el estado inicial y el tiempo transcurrido, se dice entonces que el sistema está en *condición transitoria*, ya pasado suficiente tiempo el estado del sistema es independiente del estado inicial (excepto cuando $\rho \geq 1$, pues la cola crecería indefinidamente), se dice entonces que el sistema está en *condición de estado estable*. La notación siguiente supone que el sistema está en estado estable:

- P_n = probabilidad de que exactamente n clientes estén en el sistema.
- L = número esperado de clientes en el sistema.
- L_q = número esperado de clientes en cola.
- W = tiempo esperado por los clientes en el sistema.
- W_q = tiempo esperado por los clientes en cola.
- $EW(A/B/s)$ = tiempo esperado por los clientes en cola en un modelo A/B/s.

En esencia W_q y $EW(A/B/s)$ tienen el mismo significado, sin embargo, se usa la segunda notación cuando se necesita aclarar el modelo de colas al que se hace referencia.

Las principales distribuciones que se usarán para ajustar los tiempos entre llegadas y de servicio serán notadas por:

- M = distribución exponencial.
- D = distribución degenerada (cuando los tiempos son constantes).
- G = distribución general (permite cualquier distribución arbitraria).

1.4. Relaciones entre L , W , L_q y W_q

En primer lugar, de Hillier [7] se presentan las definiciones de L y L_q que serán útiles más adelante.

$$L = \sum_{n=0}^{\infty} nP_n \quad (1.1)$$

$$L_q = \sum_{n=s}^{\infty} (n - s)P_n \quad (1.2)$$

Ahora bien, para presentar la relación entre las cuatro métricas fundamentales (L , L_q , W y W_q) se presenta la *fórmula de Little* que establece que, en un proceso de colas en estado estable, el número esperado de clientes en un sistema (L) es igual a la tasa de llegada de los clientes al sistema (λ) por el tiempo esperado por un cliente en el sistema (W).

$$L = \lambda W \quad (1.3)$$

Gracias a esta fórmula se tienen las relaciones presentadas en Hillier [7]:

$$L_q = \lambda W_q \quad (1.4)$$

y

$$W = W_q + \frac{1}{\mu} \quad (1.5)$$

Estas relaciones son sumamente importantes pues permiten encontrar el valor de las cuatro métricas fundamentales al encontrar una de ellas por algún método. Esto resulta de gran ayuda pues en la práctica algunas métricas resultan más sencillas de calcular que otras.

1.5. Usos de modelos de teoría de colas

Los modelos de teoría de colas se han usado anteriormente en diferentes sistemas operativos de servicios, pertenecientes a ramas como la medicina y actividades bancarias; siendo de gran utilidad para lograr reducir tiempos en líneas de espera, mejorar la satisfacción del cliente, optimizar costos, etc. A continuación, se muestran aplicaciones relevantes de los modelos de colas:

En el ámbito de la medicina se encuentran aplicaciones como en Ameh [1], donde se muestra que mediante un modelo de colas se logró que la mayoría de los pacientes (164, 78.1 %) pasarán 2 horas o menos en la cola antes de ser atendidos por un médico y menos de 1 hora para ver al médico, la mayoría de los pacientes (144, 68.5 %) estaban satisfechos con el tiempo que pasaban en la cola antes de ser atendidos por un médico. Otra aplicación se encuentra en Cazorla [3], donde se observa que al aumentar dos servidores por ventanilla en un hospital se logró subir de un 8 % de pacientes que califican la satisfacción como *muy buena*, a un 100 % de pacientes que la califican de *muy buena*. Una tercera aplicación se expone en Lee [11], donde se presenta un método de iteración basado en un modelo de teoría de colas

para modelar y analizar las transiciones entre el departamento de emergencias, la unidad de cuidados intensivos y la sala general de un hospital. Además, se muestra que el procedimiento de iteración es convergente y conduce a una precisión de estimación aceptable, tanto para hospitales pequeños como grandes.

Respecto a una entidad financiera se tienen aplicaciones como las mencionadas en Gómez [6], donde se logró determinar que con 3 servidores los clientes durarán en promedio 30 minutos en una entidad financiera, contando el tiempo de espera en cola y de atención. Otra aplicación se encuentra en Magnus [12], donde se aplica la teoría de colas en la gestión del tiempo en los bancos comerciales. Donde se diseña un procedimiento que optimice una medida establecida de rendimiento, como la suma de los costos de los clientes que esperan y el costo de las instalaciones inactivas. El resultado del análisis de colas se utiliza en el contexto de un modelo de optimización de costos, donde se busca la minimización de la suma de dos costos; los costos de ofrecer el servicio y los costos de espera.

Finalmente, en Portilla [13] se muestra el contraste entre los modelos de la teoría de colas y la simulación, debido a que con el modelo matemático se puede validar el modelo de simulación y este último permite al investigador profundizar mucho más en el análisis del sistema de colas objeto de estudio.

Capítulo 2

Modelos de teoría de colas con varios servidores

2.1. Proceso de nacimiento y muerte

Antes de empezar con los modelos de teoría de colas con más de un servidor, se presenta la idea de cómo se llega a los resultados de estado estable de algunos de ellos mediante el uso de cadenas de Markov a tiempo continuo.

En primer lugar, se define como un *proceso estocástico* a una colección de variables aleatorias $\{X_t : t \in T\}$ parametrizada por un conjunto T , llamado *espacio parametral*, en donde las variables toman valores en un conjunto S llamado *espacio de estados*. En el caso en que se toma como espacio parametral el conjunto discreto $T = \mathbb{N}$, se dice que el proceso es a tiempo discreto y se denota por $\{X_n : n = 0, 1, \dots\}$. Por otro lado, si se toma como espacio parametral el conjunto continuo $T = [0, \infty)$, se dice que el proceso es a tiempo continuo y se denota por $\{X_t : t \geq 0\}$.

Ahora bien, las *cadenas de Markov* son procesos estocásticos que cumplen con la propiedad de Markov. Informalmente, esta propiedad establece que los estados futuros dependen únicamente del estado actual, y no de la secuencia de estados para llegar ahí. Finalmente, una *cadena de Markov a tiempo continuo* es un tipo de cadena de Markov cuya principal diferencia es que los tiempos de estancia en un estado es una variable aleatoria con distribución exponencial.

Cabe señalar también que la *distribución estacionaria o límite* de una cadena de

Markov consiste en una distribución estable, es decir, donde pasado cierto tiempo el estado del sistema es independiente del estado inicial y se mantiene constante. Adicionalmente, el *generador infinitesimal* es una matriz cuyas entradas son las tasas de transición de un estado a otro en una cadena de Markov.

Un caso de cadena de Markov a tiempo continuo es el proceso de nacimiento y muerte con espacio de estados $S = \mathbb{N}$ y con generador infinitesimal dado por:

$$G = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & \cdots & 0 & 0 & \cdots \\ \mu_1 & -(\lambda_1 + \mu_1) & \lambda_1 & \cdots & 0 & 0 & \cdots \\ 0 & \mu_2 & -(\lambda_2 + \mu_2) & \cdots & 0 & 0 & \cdots \\ 0 & 0 & \mu_3 & \cdots & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \lambda_{n-1} & 0 & \cdots \\ 0 & 0 & 0 & 0 & -(\lambda_n + \mu_n) & \lambda_n & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

Donde los índices $\lambda_0, \lambda_1, \dots$ y μ_1, μ_2, \dots son las tasas de nacimiento y muerte, respectivamente. Notemos que podemos describir un sistema de líneas de espera como un proceso de nacimiento (con tasa λ_n) y muerte (con tasa μ_n), donde un nacimiento sería representado por la llegada de un cliente al sistema y una muerte sería la salida (o atención) de un cliente del sistema. Ahora bien, la idea de condición de estado estable es que pasado cierto tiempo el estado del sistema es independiente del estado inicial, lo cual sería equivalente a encontrar la distribución estacionaria (o distribución límite) del proceso de nacimiento y muerte.

Para encontrar la distribución estacionaria se emplea la siguiente proposición presentada en Rincón [14]:

PROPOSICIÓN 2.1. *La distribución $P = (P_1, P_2, \dots, P_n)$ es estacionaria para la cadena con generador infinitesimal G si y sólo si:*

$$PG = 0$$

De la proposición 2.1, el sistema de ecuaciones a resolver será:

$$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \end{pmatrix} = \begin{pmatrix} P_0 \\ P_1 \\ P_2 \\ P_3 \\ \vdots \\ P_{n-2} \\ P_{n-1} \\ P_n \\ \vdots \end{pmatrix}^T \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & \cdots & 0 & 0 \\ \mu_1 & -(\lambda_1 + \mu_1) & \lambda_1 & \cdots & 0 & 0 \\ 0 & \mu_2 & -(\lambda_2 + \mu_2) & \cdots & 0 & 0 \\ 0 & 0 & \mu_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \lambda_{n-2} & 0 \\ 0 & 0 & 0 & 0 & -(\lambda_{n-1} + \mu_{n-1}) & \lambda_{n-1} \\ 0 & 0 & 0 & 0 & \mu_n & -(\lambda_n + \mu_n) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

Es decir, las ecuaciones de balance del proceso de nacimiento y muerte son:

$$\begin{aligned}
\mu_1 P_1 &= \lambda_0 P_0 \\
\lambda_0 P_0 + \mu_2 P_2 &= (\lambda_1 + \mu_1) P_1 \\
\lambda_1 P_1 + \mu_3 P_3 &= (\lambda_2 + \mu_2) P_2 \\
&\vdots \\
\lambda_{n-2} P_{n-2} + \mu_n P_n &= (\lambda_{n-1} + \mu_{n-1}) P_{n-1} \\
\lambda_{n-1} P_{n-1} + \mu_{n+1} P_{n+1} &= (\lambda_n + \mu_n) P_n
\end{aligned}$$

Ahora bien, escribiendo cada elemento de P en función de P_0 tenemos:

$$\begin{aligned}
P_1 &= \left(\frac{\lambda_0}{\mu_1} \right) P_0 \\
P_2 &= \left(\frac{\lambda_1 \lambda_0}{\mu_2 \mu_1} \right) P_0 \\
P_3 &= \left(\frac{\lambda_2 \lambda_1 \lambda_0}{\mu_3 \mu_2 \mu_1} \right) P_0 \\
&\vdots \\
P_n &= \left(\frac{\lambda_{n-1} \lambda_{n-2} \cdots \lambda_0}{\mu_n \mu_{n-1} \cdots \mu_1} \right) P_0
\end{aligned}$$

Para simplificar la notación, definamos:

$$C_n = \begin{cases} 1 & n = 0 \\ \frac{\lambda_{n-1}\lambda_{n-2}\cdots\lambda_0}{\mu_n\mu_{n-1}\cdots\mu_1} & n = 1, 2, \dots \end{cases}$$

Así, las probabilidades estacionarias o de estado estable son:

$$P_n = C_n P_0, \quad n = 0, 1, 2, \dots$$

Ahora bien, sabemos que para la distribución estacionaria se cumple que:

$$\sum_{n=0}^{\infty} P_n = 1 \Rightarrow \left(\sum_{n=0}^{\infty} C_n \right) P_0 = 1 \Rightarrow P_0 = \left(\sum_{n=0}^{\infty} C_n \right)^{-1}$$

Cuando un modelo de líneas de espera se basa en un proceso de nacimiento y muerte, de manera que el estado del sistema n representa el número de clientes en el sistema de colas; las métricas principales (L, L_q, W, W_q) se pueden obtener después de calcular las P_n mediante las fórmulas de la Sección 1.4.

Es importante notar que cuando el producto de la tasa de servicio y el número de servidores $s\mu$ excede la tasa de llegadas λ , es decir, cuando

$$\rho = \frac{\lambda}{s\mu} < 1$$

el modelo que representa el sistema de colas tarde o temprano alcanzará la condición de estado estable, pues informalmente se estará atendiendo más rápido de lo que llegan los clientes y llegará un punto en que la cola no crecerá más.

En el caso contrario, si $\rho \geq 1$ la cola crecerá indefinidamente y no se alcanzará la condición de estado estable.

2.2. Modelo M/M/s

En este modelo se parte de la suposición de que los tiempos entre llegadas son independientes e idénticamente distribuidos de acuerdo con una distribución exponencial (es decir, el proceso de entrada de clientes es de Poisson) y los tiempos de servicio son independientes e idénticamente distribuidos de acuerdo con otra distribución exponencial, donde los parámetros de estas distribuciones son la tasa de llegada de clientes λ y la tasa de servicio a clientes μ respectivamente.

Cuando el sistema tiene un solo servidor ($s = 1$) se sigue que los parámetros del proceso de nacimiento y muerte son $\lambda_n = \lambda$, $n = 0, 1, 2, \dots$ y $\mu_n = \mu$, $n = 1, 2, \dots$. Sin embargo, para el caso de múltiples servidores ($s > 1$) y ($n > 1$), μ_n no es lo mismo que μ . En este caso se tiene:

$$\mu_n = \begin{cases} n\mu & n < s \\ s\mu & n \geq s \end{cases}$$

De donde, calculando los factores C_n se tiene:

- Para $n < s$:

$$\begin{aligned} C_n &= \left(\frac{\lambda_{n-1} \lambda_{n-2} \cdots \lambda_0}{\mu_n \mu_{n-1} \cdots \mu_1} \right) \\ &= \frac{\lambda^n}{(\mu_n)^n} \\ &= \frac{\lambda^n}{(n\mu)^n} \\ &= \frac{\lambda^n}{1\mu 2\mu \cdots n\mu} \\ &= \frac{\lambda^n}{n! \mu^n} \\ &= \left(\frac{\lambda}{\mu} \right)^n \left(\frac{1}{n!} \right) \end{aligned}$$

- Para $n \geq s$:

$$\begin{aligned} C_n &= \left(\frac{\lambda_{n-1} \lambda_{n-2} \cdots \lambda_0}{\mu_n \mu_{n-1} \cdots \mu_1} \right) \\ &= \frac{\lambda^n}{(\mu_1 \mu_2 \cdots \mu_s) (\mu_{s+1} \mu_{s+2} \cdots \mu_n)} \\ &= \frac{\lambda^s \lambda^{n-s}}{1\mu 2\mu \cdots s\mu \underbrace{s\mu s\mu s\mu}_{(n-s) \text{ veces}}} \\ &= \frac{\lambda^s \lambda^{n-s}}{(s! \mu^s) (s\mu)^{n-s}} \end{aligned}$$

En resumen, los factores C_n son:

$$C_n = \begin{cases} 1 & n = 0 \\ \frac{(\lambda/\mu)^n}{n!} & n = 1, 2, \dots, s-1 \\ \frac{\lambda^s}{(s!\mu^s)} \left(\frac{\lambda}{s\mu}\right)^{n-s} & n = s, s+1, \dots \end{cases}$$

Así, si $\lambda < s\mu$ entonces,

$$\begin{aligned} P_0 &= \left(\sum_{n=0}^{\infty} C_n \right)^{-1} \\ &= \left[1 + \sum_{n=1}^{s-1} \frac{(\lambda/\mu)^n}{n!} + \frac{(\lambda/\mu)^s}{s!} \sum_{n=s}^{\infty} \left(\frac{\lambda}{s\mu}\right)^{n-s} \right]^{-1} \\ &= \left[\sum_{n=0}^{s-1} \frac{(\lambda/\mu)^n}{n!} + \frac{(\lambda/\mu)^s}{s!} \sum_{n=s}^{\infty} \left(\frac{\lambda}{s\mu}\right)^{n-s} \right]^{-1} \rightarrow 1 = (\lambda/\mu)^0/0! \\ &= \left[\sum_{n=0}^{s-1} \frac{(\lambda/\mu)^n}{n!} + \frac{(\lambda/\mu)^s}{s!} \frac{1}{1 - \lambda/(s\mu)} \right]^{-1} \rightarrow \sum_{n=a}^{\infty} x^{n-a} = \frac{1}{1-x}, \quad |x| < 1 \end{aligned}$$

De donde,

$$P_n = \begin{cases} \frac{(\lambda/\mu)^n}{n!} P_0 & 0 \leq n < s \\ \frac{\lambda^s}{(s!\mu^s)} \left(\frac{\lambda}{s\mu}\right)^{n-s} P_0 & n \geq s \end{cases}$$

Así, se han encontrado las probabilidades estacionarias o de estado estable para el modelo M/M/s.

Se calcula ahora la métrica L_q mediante la ecuación (1.2).

$$\begin{aligned} L_q &= \sum_{n=s}^{\infty} (n-s) P_n \\ &= \sum_{j=0}^{\infty} j P_{j+s} \quad \rightarrow \text{cambio variable } j = n-s \\ &= \sum_{j=0}^{\infty} j \left(\frac{1}{s!}\right) \left(\frac{\lambda}{\mu}\right)^s \left(\frac{\lambda}{s\mu}\right)^j P_0 \\ &= P_0 \frac{(\lambda/\mu)^s}{s!} \sum_{j=0}^{\infty} j \rho^j \quad \rightarrow \text{def. factor de utilización} \\ &= P_0 \frac{(\lambda/\mu)^s}{s!} \sum_{j=0}^{\infty} \rho \frac{d}{d\rho} \rho^j \quad \rightarrow \text{derivada de una potencia} \end{aligned}$$

$$\begin{aligned}
&= P_0 \frac{(\lambda/\mu)^s}{s!} \rho \frac{d}{d\rho} \sum_{j=0}^{\infty} \rho^j \\
&= P_0 \frac{(\lambda/\mu)^s}{s!} \rho \frac{d}{d\rho} \left[\frac{1}{1-\rho} \right] \quad \rightarrow \quad \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}, \quad |x| < 1 \\
&= P_0 \frac{(\lambda/\mu)^s}{s!} \left[\frac{\rho}{(1-\rho)^2} \right]
\end{aligned}$$

Finalmente, gracias a la fórmula de Little (1.3) y a las relaciones (1.4), (1.5) se tienen las métricas para el modelo M/M/s:

$$L_q = P_0 \frac{(\lambda/\mu)^s}{s!} \left[\frac{\rho}{(1-\rho)^2} \right] \quad (2.1)$$

$$W_q = EW(M/M/s) = \frac{L_q}{\lambda} = \frac{P_0 (\lambda/\mu)^s}{\lambda s!} \left[\frac{\rho}{(1-\rho)^2} \right] \quad (2.2)$$

$$W = W_q + \frac{1}{\mu} = \frac{P_0 (\lambda/\mu)^s}{\lambda s!} \left[\frac{\rho}{(1-\rho)^2} \right] + \frac{1}{\mu} \quad (2.3)$$

$$L = \lambda \left(W_q + \frac{1}{\mu} \right) = \lambda \left(\frac{P_0 (\lambda/\mu)^s}{\lambda s!} \left[\frac{\rho}{(1-\rho)^2} \right] + \frac{1}{\mu} \right) \quad (2.4)$$

2.3. Modelo M/D/s

En este modelo se parte de la suposición de que los tiempos entre llegadas son independientes e idénticamente distribuidos de acuerdo con una distribución exponencial (es decir, el proceso de entrada de clientes es de Poisson) y los tiempos de servicio son constantes; las tasas de llegada de clientes y las tasas de servicio a clientes son λ y μ respectivamente. Para este modelo no se tienen resultados exactos de las métricas principales, por ello, se trabajará con la aproximación para W_q de Kimura [10] y el resto de las métricas se las puede obtener de las relaciones de la Sección 1.4.

Antes de presentar la aproximación de Kimura [10], se presentan fórmulas que se han usado en los últimos años y las desventajas que presentan. Se asume que $\rho < 1$.

Usualmente para calcular $EW(M/D/s)$ se usa la fórmula:

$$EW(M/D/s) = \frac{1}{2s^2\mu\rho} \left((s\rho)^2 - s(s-1) + \sum_{j=0}^{s-1} [s(s-1) - j(j-1)] P_j^{det} \right) \quad (2.5)$$

donde, P_j^{det} , ($j = 0, \dots, s-1$) es la probabilidad de estado estable de tener j clientes en el sistema $M/D/s$. Kimura [10] indica que la principal desventaja de esta aproximación es la necesidad de recurrir a técnicas numéricas de ecuaciones lineales a gran escala, pues la fórmula (2.5) contiene s probabilidades P_j^{det} desconocidas, las cuales se pueden obtener, pero tras resolver un sistema infinito de ecuaciones lineales.

Otra fórmula para calcular $EW(M/D/s)$ es la propuesta por Crommelin [5]:

$$EW(M/D/s) = \frac{1}{\mu} \sum_{j=1}^{\infty} \sum_{k=j_s+1}^{\infty} \left(\frac{(js\rho)^{k-1}}{(k-1)!} - \frac{(js\rho)^k}{\rho k!} \right) e^{-js\rho} \quad (2.6)$$

A diferencia de la fórmula (2.5) no se tienen parámetros indeterminados, sin embargo, la principal desventaja de esta aproximación es que tiene un costo computacional muy alto pues la serie en (2.6) converge muy lentamente especialmente en tráfico pesado ($\rho > 0,6$); y también cuando el número de servidores es alto, pues se deben calcular factoriales sumamente altos y programas como R tienen un límite de poder realizar cálculos hasta con el número (170)!

Como se puede ver, las fórmulas (2.5) y (2.6) tienen una dificultad numérica en sus cálculos, y es esa misma dificultad la que ha impulsado la búsqueda de aproximaciones sencillas y precisas para $EW(M/D/s)$. En Kimura [9] se evalúa distintas aproximaciones y se muestra que la aproximación de Cosmetatos [4]:

$$EW(M/D/s) \simeq \frac{1}{2} \{1 + f(s)g(\rho)\} EW(M/M/s) \quad (2.7)$$

tiene la mejor calidad para propósitos prácticos, donde

$$f(s) = \frac{(s-1)(\sqrt{4+5s}-2)}{16s}$$

$$g(\rho) = \frac{1-\rho}{\rho}$$

En Kimura [9] se demuestra que la aproximación (2.7) funciona bien si el número de servidores es pequeño ($s \leq 10$) y el tráfico es pesado $\rho > 0,6$. Sin embargo,

presenta desventajas cuando $s \rightarrow \infty$ o $\rho \rightarrow 0$, pues la aproximación tiende a sobreestimar el valor real.

Así, con el fin de solventar estos problemas se presenta en Kimura [10] la aproximación

$$EW(M/D/s) \simeq \frac{1}{2} \{1 + f(s)g(\rho)h(s,\rho)\}EW(M/M/s) \quad (2.8)$$

donde $h(s,\rho)$ es la función de corrección dada por

$$h(s,\rho) = \zeta(s, a(\rho))\eta(b(s), \rho)$$

con

$$\zeta(s, x) = \sqrt{1 - \exp\left(-\frac{2x}{s-1}\right)} \quad x \geq 0$$

$$\eta(y, \rho) = 1 - \exp\left(-\frac{\rho y}{1-\rho}\right) \quad y \geq 0$$

Las funciones $a(\rho)$ y $b(s)$ son definidas como

$$a(\rho) = \frac{25,6}{\{g(\rho)\eta(\beta, \rho)\}^2}$$

y

$$b(s) = \frac{s-1}{(s+1)f(s)\zeta(s, \alpha)}$$

donde α y β satisfacen la relación

$$\alpha\beta^2 = 25,6$$

en general se usa $\alpha = 2,2$ (y por tanto $\beta = \sqrt{25,6/2,2}$)

La aproximación de Kimura [10] tiene las siguientes propiedades:

- i Da el resultado exacto para $s = 1$.
- ii Se vuelve asintóticamente exacto cuando $s \rightarrow \infty$ o $\rho \rightarrow 0$ o $\rho \rightarrow 1$.
- iii Sus errores porcentuales relativos son menores al 1 % para casi todos los casos cuando $s \leq 20$ y como máximo 5 % para otros casos.

2.4. Modelo M/G/s

En este modelo se parte de la suposición de que los tiempos entre llegadas son independientes e idénticamente distribuidos de acuerdo con una distribución exponencial (es decir, el proceso de entrada de clientes es de Poisson) y los tiempos de servicio son independientes e idénticamente distribuidos de acuerdo con alguna distribución arbitraria; las tasas de llegada de clientes y las tasas de servicio a clientes son λ y μ respectivamente. Para este modelo las métricas de rendimiento no se conocen y siguen siendo un problema abierto, por ello, se trabajará con la aproximación para W_q de Boxma [2] y el resto de las métricas se las puede obtener de las relaciones de la Sección 1.4

Se denota la distribución de los tiempos de servicio por $B(\cdot)$ y sus tres primeros momentos por $\beta_1, \beta_2, \beta_3$. Así, se puede reescribir el factor de utilización $\rho = \beta_1\lambda/s < 1$.

Una de las primeras modificaciones que se encuentra en Boxma [2] es asumir que se formarán colas separadas delante de cada servidor, así, el cliente que llega tiene probabilidad $1/s$ de unirse a la j -ésima cola ($j = 1, \dots, s$) independientemente de los tamaños de cola. Es decir, se ha reemplazado el sistema M/G/s por s sistemas M/G/1 independientes cada uno con tasa de arribo λ/s y distribución de tiempo de servicio $B(\cdot)$, de aquí se tiene la desigualdad:

$$EW(M/G/s) \leq EW(M/G/1)$$

Esta desigualdad se puede interpretar de la siguiente forma: en un sistema M/G/s a diferencia de un sistema de s operadores M/G/1 independientes, nunca pasa que un cliente debe esperar mientras un servidor está libre, en este sentido se dice que los servidores cooperan en el sistema M/G/s. De donde, se introduce una medida para la cooperación de los servidores en un sistema M/G/s; el *coeficiente de cooperación* C_{Gs} definido por:

$$C_{Gs} = \frac{EW(M/G/s)}{EW(M/G/1)} \leq 1$$

Ahora bien, dado que en un sistema M/M/s hay expresiones explícitas para $EW(M/M/s)$ y $EW(M/M/1)$ que se las puede encontrar en Hillier [7], se conoce el coeficiente de cooperación en un sistema M/M/s (C_{Ms}).

Luego, se define el *coeficiente de cooperación normado* N_{Gs} en un sistema M/G/s

definido como:

$$N_{Gs} = \frac{C_{Ms}}{C_{Gs}}$$

De donde,

$$\begin{aligned} EW(M/G/s) &= \left(\frac{EW(M/G/1)}{EW(M/M/1)} \right) \left(\frac{1}{N_{Gs}} \right) EW(M/M/s) \\ &= \left(\frac{\beta_2}{2\beta_1^2} \right) \left(\frac{1}{N_{Gs}} \right) EW(M/M/s) \end{aligned} \quad (2.9)$$

Así, lo que se pretende en Boxma [2] es encontrar aproximaciones para el coeficiente de cooperación normado en función del factor de utilización $N_{Gs}(\rho)$, para esto se presentan los siguientes resultados:

TEOREMA 2.2. *Sea $\beta_3 < \infty$, entonces*

$$\lim_{\rho \rightarrow 1} N_{Gs}(\rho) = 1$$

DEFINICIÓN 2.1. *Se tiene un intercambio de procesadores modificado M/G/s (IPM) cuando los clientes llegan a un sistema de s servidores de acuerdo a un proceso de Poisson homogéneo con tasa λ y los servicios requeridos son independientes y todos tienen distribución $B(\cdot)$.*

Siempre y cuando como mucho s clientes estén presentes, cada cliente es servido con tasa de servicio de una unidad de trabajo por unidad de tiempo; cuando el número de clientes x presente excede s, todos los clientes son servidos simultáneamente con tasa de servicio $f(x) = s/x$.

DEFINICIÓN 2.2. *Para el IPM M/G/s se define \mathbf{t} como el tiempo transcurrido entre la llegada de un cliente y la primera vez en que un servidor se desocupa; bajo el supuesto que C y los clientes posteriores no son atendidos.*

En un sistema ordinario M/G/s, \mathbf{t} es igual al tiempo de espera de un cliente. En Boxma [2] se prueba que para un IPM M/G/s; se tiene que $t_{Gs} = E(\mathbf{t})$ (esperanza de \mathbf{t}), se puede calcular mediante:

$$t_{Gs} = \frac{EW(M/M/s)}{\beta_1} \int_{v=0}^{\infty} [1 - H(v)] \left[\frac{1 - H(v)}{1 - \rho H(v)} \right]^{s-1} \left[s - \frac{(s-1)(\rho - \rho H(v))}{1 - \rho H(v)} \right] dv$$

donde

$$H(v) = \int_{u=0}^v \frac{1 - B(u)}{\beta_1} du, \quad v > 0$$

PROPOSICIÓN 2.3.

$$\lim_{\rho \rightarrow 0} \frac{t_{Gs}}{EW(M/G/s)} = 1$$

TEOREMA 2.4. *Bajo la proposición 2.3 tenemos:*

$$\lim_{\rho \rightarrow 0} N_{Gs}(\rho) = \frac{\beta_2}{2\beta_1 s \int_{v=0}^{\infty} (1 - H(v))^s dv}$$

PROPOSICIÓN 2.5. *Para un sistema M/D/s, se tienen la siguiente aproximación para el coeficiente de cooperación normado N_{Ds} :*

$$N_{Ds}(\rho) = \left[1 + (1 - \rho)(s - 1) \frac{\sqrt{(4 + 5s) - 2}}{16\rho s} \right]^{-1}$$

De donde:

$$\lim_{\rho \rightarrow 0} N_{Ds}(\rho) = \frac{s + 1}{2s}$$

Las demostraciones de los teoremas (2.2) y (2.4) se las encuentra en Boxma [2].

Ahora bien, una vez presentados los resultados que ayudarán a encontrar la aproximación para el coeficiente N_{Gs} , se presenta la aproximación propuesta por Boxma [2]:

$$N_{Gs} \simeq A(s) + (1 - A(s))N_{Ds}(\rho) \quad (2.10)$$

Para encontrar la forma de $A(s)$, notemos que:

$$\begin{aligned} N_{Gs} &= A(s) + (1 - A(s))N_{Ds}(\rho) \\ \lim_{\rho \rightarrow 0} N_{Gs}(\rho) &= A(s) + (1 - A(s)) \lim_{\rho \rightarrow 0} N_{Ds}(\rho) \\ \frac{\beta_2}{2\beta_1 s \int_{v=0}^{\infty} (1 - H(v))^s dv} &= A(s) + (1 - A(s)) \frac{s + 1}{2s} \quad \rightarrow \quad \text{Teo 2,4 y Prop 2,5} \\ \frac{\beta_2}{2\beta_1 s \int_{v=0}^{\infty} (1 - H(v))^s dv} &= A(s) \left[1 - \frac{s + 1}{2s} \right] + \frac{s + 1}{2s} \\ A(s) &= \frac{\beta_2}{(s - 1)\beta_1 \int_{v=0}^{\infty} (1 - H(v))^s dv} - \frac{s + 1}{s - 1} \end{aligned}$$

Finalmente se presenta la aproximación para un modelo M/G/s propuesta en Boxma [2] definida como:

$$EW(M/G/s) \simeq \left(\frac{\beta_2}{2\beta_1^2} \right) \frac{2EW(M/M/s)EW(M/D/s)}{2A(s)EW(M/D/s) + (1 - A(s))EW(M/M/s)} \quad (2.11)$$

Cabe señalar que en Boxma [2] para la aproximación $EW(M/D/s)$ se usa la propuesta por Crommelin [5], sin embargo, por todas las propiedades presentadas en la Sección 2.3 se propone cambiar la fórmula (2.11) usando la aproximación $EW(M/D/s)$ propuesta por Kimura [10].

Además, se puede notar que si se emplea la aproximación Crommelin [5], para un caso aleatorio en que $j = 9$ y $s = 20$ al momento de aplicar la fórmula (2.6) aparecería un valor $(k - 1)! = (js + 1 - 1) = 180!$ lo cual excede el límite de capacidad en el programa R. Por otro lado, al usar la aproximación de Kimura [10], al momento de aplicar la fórmula (2.8) con el mismo valor de $s = 20$ el máximo valor que intervendría en los cálculos es $s! = 20!$; por lo cual se puede observar que la aproximación de Kimura [10] se la puede llevar de manera eficiente a la práctica.

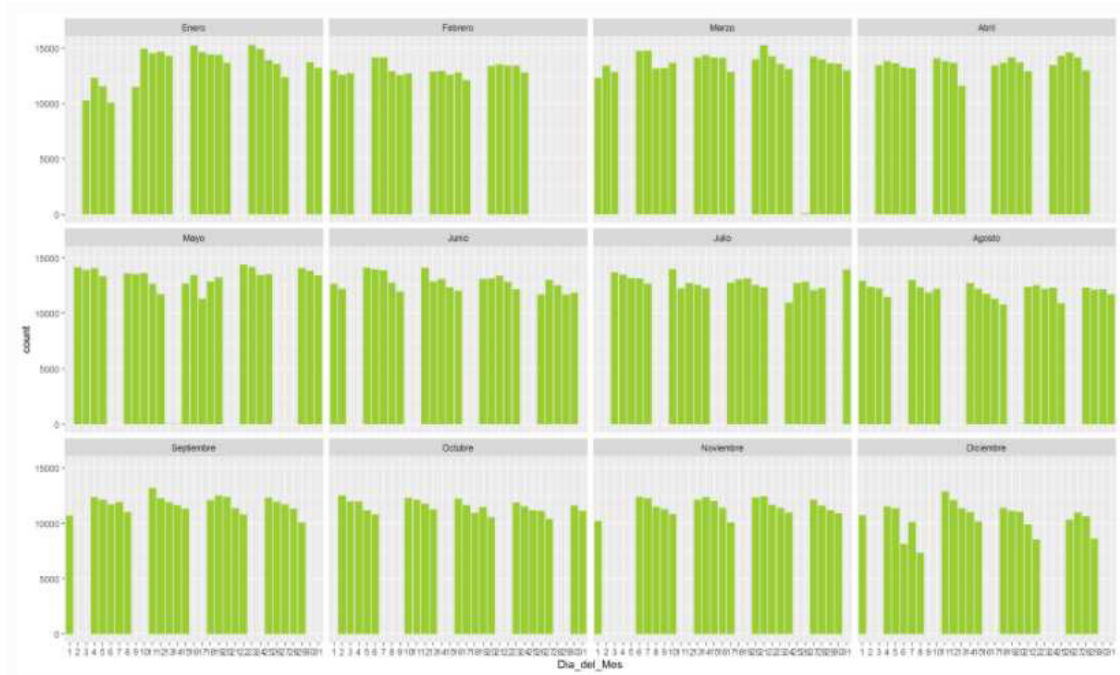
Capítulo 3

Aplicación en un sistema operativo de servicios

3.1. Antecedentes

Se contó con un archivo con datos proporcionado por el S.R.I con los campos: año, mes, día, horas, minutos, oficina, nombre del servicio, ticket ID, tiempo de espera y tiempo de transacción; desde agosto de 2016 hasta mayo de 2018, donde se tomaba en cuenta 53 oficinas y 33 tipos de servicios.

Usando dicho archivo se han realizado análisis exploratorios en un proyecto previo de investigación de la Escuela Politécnica Nacional (PIGR_19_11) entre 2020 - 2022. Los resultados de dicho análisis exploratorio se resumen a continuación:



Autor: Estudio proyecto PIGR-19-11. **Fuente:** S.R.I.

Figura 3.1 Día del mes

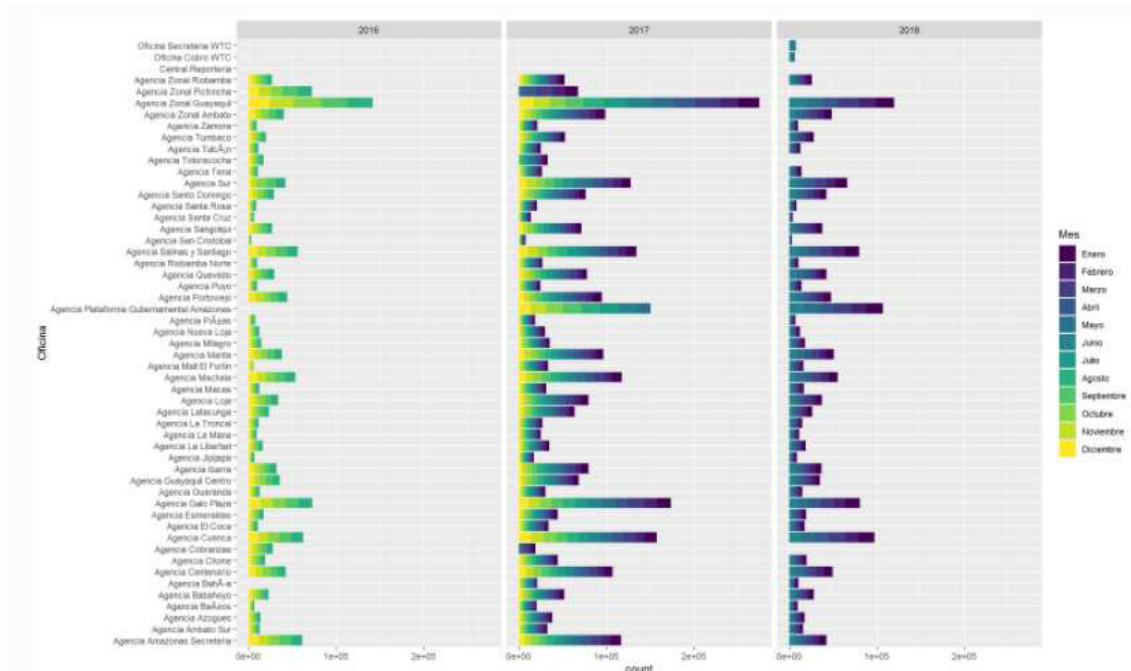
La Figura 3.1 indica la cantidad de trámites que se realizan según el día del mes, donde se puede apreciar que las columnas faltantes se dan por los fines de semana, feriados, etc. No se distingue una afluencia diferente en la cantidad de trámites con respecto a inicios, mediados y finales de mes, pero si dentro de una semana; por ejemplo, se ve que en el inicio de semana hay más trámites que al final.



Autor: Estudio proyecto PIGR-19-11. Fuente: S.R.I.

Figura 3.2 Oficinas

Se observa también en la Figura 3.2 que hay 53 oficinas pertenecientes al S.R.I.



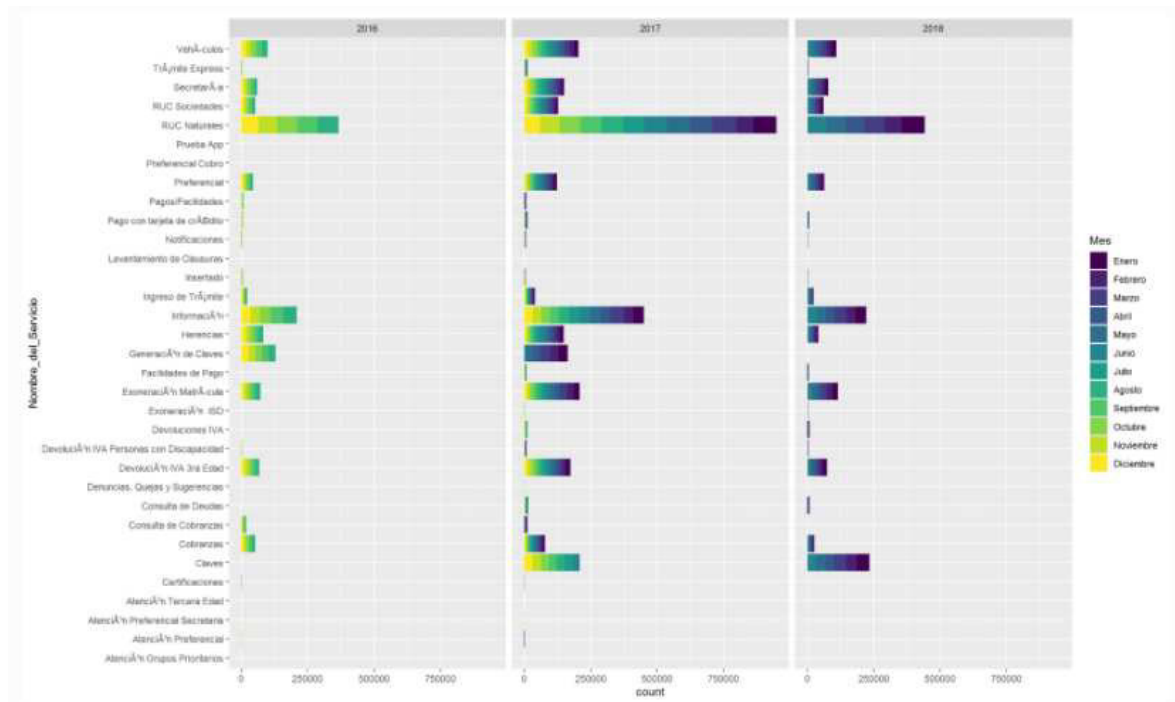
Autor: Estudio proyecto PIGR-19-11. Fuente: S.R.I.

Figura 3.3 Tipos de trámite

La Figura 3.3 muestra una comparación entre el número de trámites y la agencia en los últimos 3 años, en los años 2016 y 2018 hay menos trámites pues no están contabilizados los trámites de todo el año. Aquí se observa que:

- Agencia Zonal Guayaquil.
- Galo Plaza.
- Agencia Cuenca.
- Agencia Plataforma Gubernamental Amazonas.
- Agencia Salinas y Santiago.

Son las agencias que más trámites han realizado. Cabe recalcar que la agencia Plataforma Gubernamental Amazonas, de aquí en adelante PGA; fue abierta en 2017 por lo que no tiene datos en 2016, aun así, se mantiene como una de las agencias que más trámites genera, por lo cual el proyecto de titulación se centra en la PGA.



Autor: Estudio proyecto PIGR-19-11. Fuente: S.R.I.

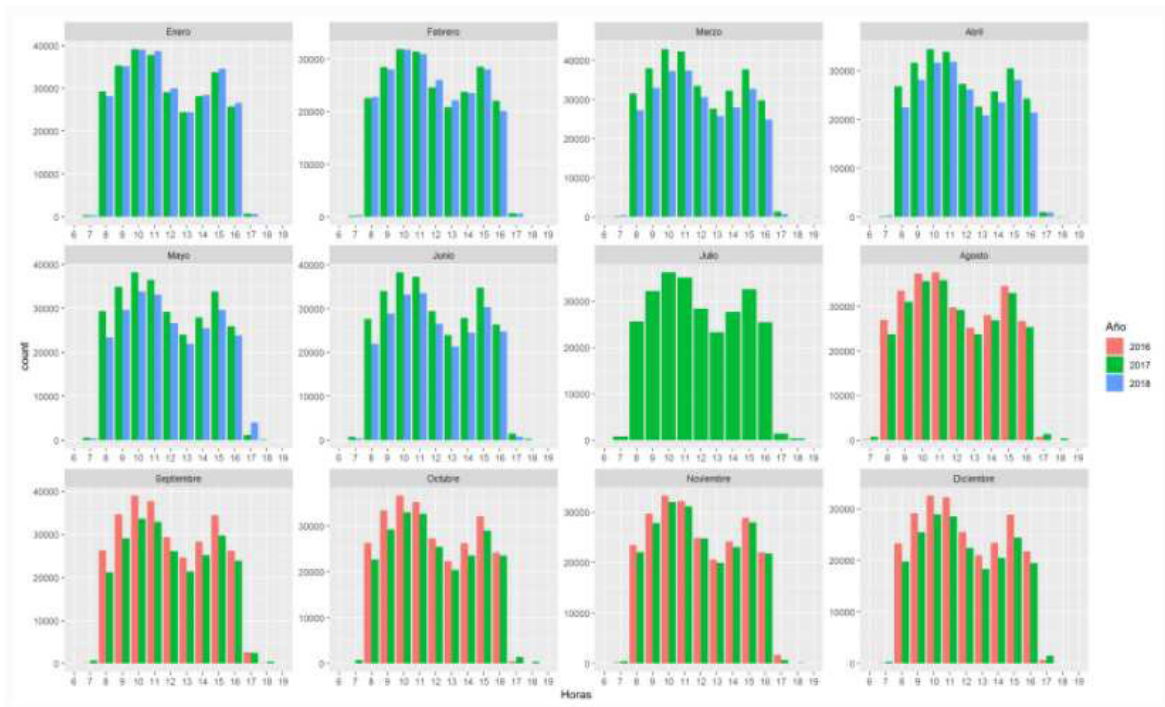
Figura 3.4 Trámites por agencia

La Figura 3.4 muestra cuales son los servicios más solicitados según el mes. Aquí se observa que:

- RUC Naturales.
- Información.
- Claves.

- Vehículos.
- Exoneración matricula.

Son los servicios más solicitados y estos se producen en todos los meses pues su barra tiene todos los colores, además se puede observar que el servicio *claves* se empezó a solicitar desde mediados de 2017.

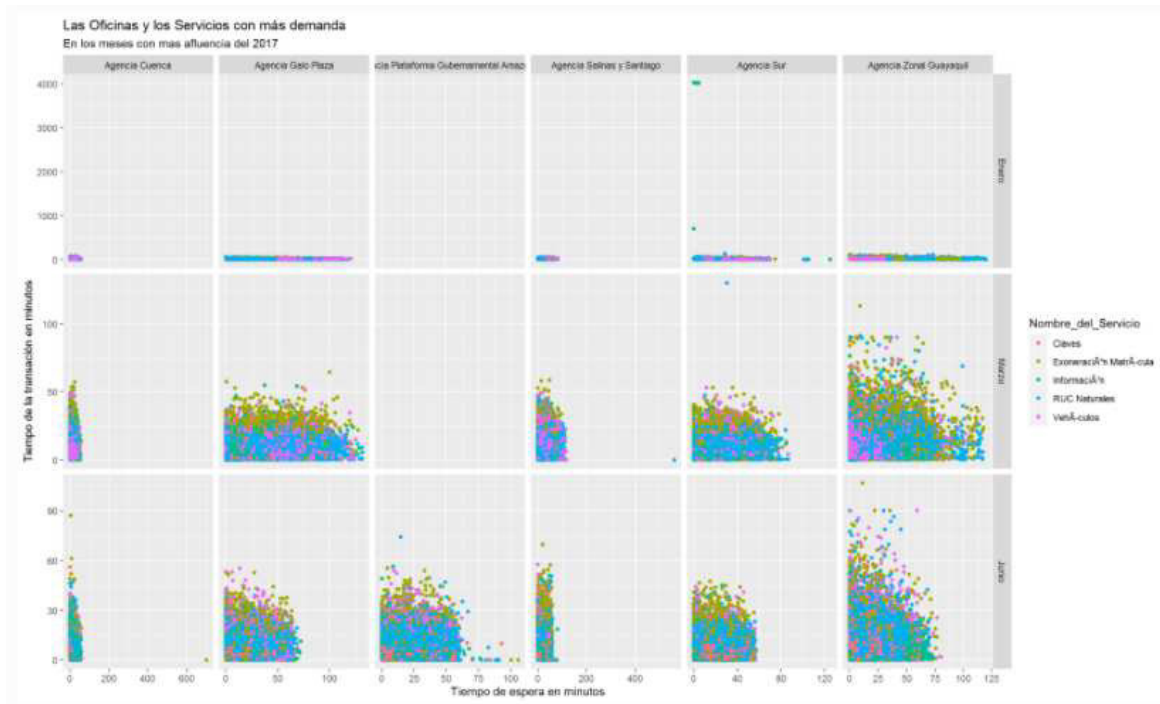


Autor: Estudio proyecto PIGR-19-11. Fuente: S.R.I.

Figura 3.5 Trámites por horas

La Figura 3.5 muestra el número de trámites según la hora del día. Aquí se observa un comportamiento con respecto a las horas; teniendo la máxima afluencia a las 10:00 y 15:00 horas y mínima afluencia a las 8:00, 13:00 y 16:00 horas. Esto es independiente del mes y año pues se ve este comportamiento en todos los sub-gráficos y en todos los colores de las barras.

Posteriormente se analizan las columnas de tiempo de espera y tiempo de transacción conjuntamente para ver su tendencia. Cabe resaltar que para los siguientes gráficos se considera únicamente las oficinas con más trámites y los servicios más solicitados en los meses con más afluencia, pues se busca observar el comportamiento en los meses más críticos.

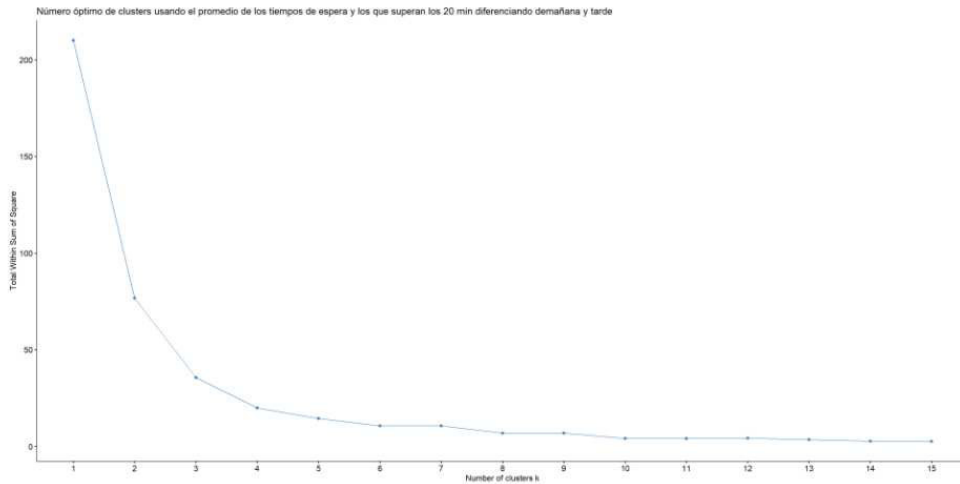


Autor: Estudio proyecto PIGR-19-11. **Fuente:** S.R.I.

Figura 3.6 Tiempos espera/transacción

Se observa en la Figura 3.6 una tendencia horizontal, la cual indica que sin importar el tiempo de espera en cola, la transacción puede durar mucho o poco tiempo (tiempo de espera no influye en el tiempo de transacción). Además, gracias a las nubes de puntos se observa que el tiempo de transacción no depende del servicio pues en el mismo servicio (color) hay eventos que ocupan mucho y poco tiempo. Finalmente se puede observar en el eje x que los tiempos de espera son sumamente altos (superando incluso los 50 minutos).

A continuación, se generan grupos de agencias similares mediante clusters, que poseen como variables los tiempos de espera y el número de clientes que superan un tiempo de espera de 20 minutos.

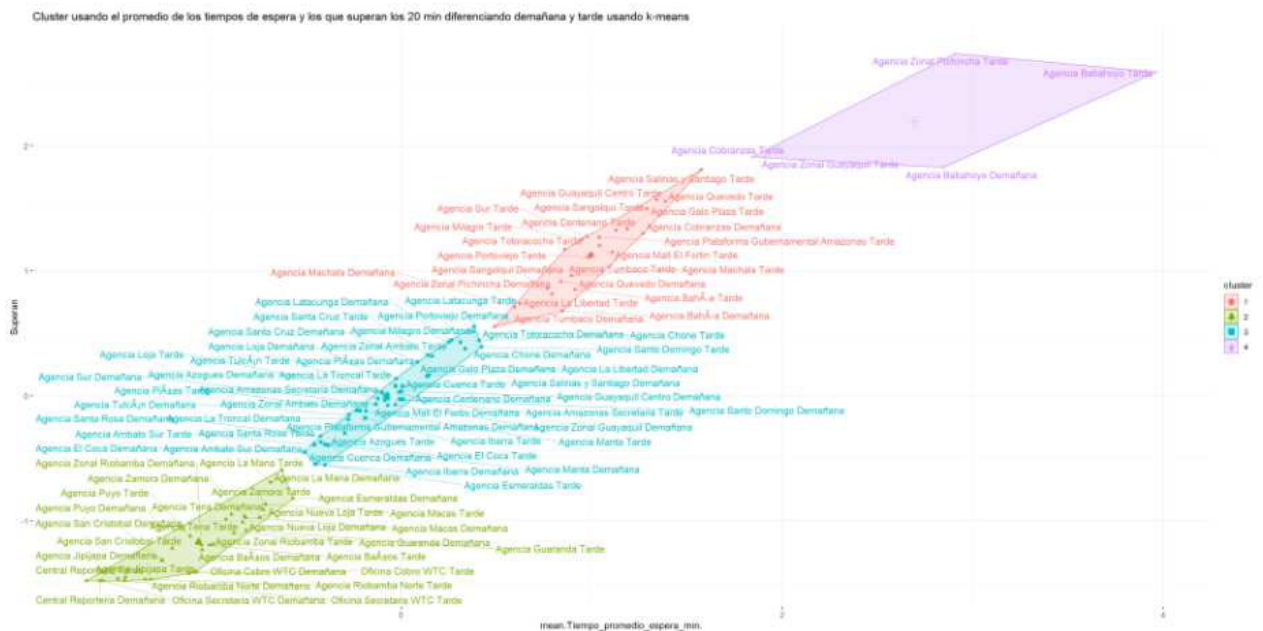


Autor: Estudio proyecto PIGR-19-11. **Fuente:** S.R.I.

Figura 3.7 Óptimo de clusters

La Figura 3.7 indica que el mejor número de clusters son 4.

Seguidamente se realizan los clusters en la Figura 3.8.



Autor: Estudio proyecto PIGR-19-11. **Fuente:** S.R.I.

Figura 3.8 Clusters

Así, se puede buscar describir el comportamiento de una oficina y las oficinas pertenecientes a ese cluster tendrán un comportamiento similar.

Tras una revisión del análisis exploratorio realizado a los datos de las agencias del S.R.I, se ha escogido trabajar con la agencia PGA por ser la agencia con más datos y usar los meses con más afluencia (segundo semestre de 2017). Además, el problema principal que se puede observar es el alto tiempo de espera para que un cliente sea atendido. Dicho problema se aborda en este proyecto de titulación, y se busca determinar un modelo matemático que permita escoger un número de servidores adecuado para operar en las diferentes horas del día en que se brinda atención en la PGA, reduciendo así los tiempos de espera en a lo sumo 20 minutos; disminuyendo tiempos de espera mayores a 50 minutos como los de la Figura 3.6.

3.2. Depuración archivo de datos

3.2.1. Descripción del archivo

El archivo de datos proporcionada por el S.R.I cuenta con las variables: año, mes, día del mes, hora, minutos, oficina, nombre del servicio, ticket ID, tiempo de espera (minutos) y tiempo de transacción (minutos); desde el mes de mayo hasta el mes de diciembre de 2017, de estas variables se pretende extraer los tiempos entre llegadas y los tiempos de servicio para ver su ajuste a alguna distribución y con ello elegir un modelo de teoría de colas.

En primer lugar, para la obtención de los tiempos entre llegadas se ha incluido la variable *segundos*, la cual se ha generado mediante una distribución uniforme $U \sim [0, 59]$ con el fin de poder separar 2 llegadas simultaneas en un mismo minuto y así evitar tiempos entre llegadas con valor 0. Posteriormente se han dividido los datos por días y en cada día se han sumado las variables hora, minutos y segundos obteniendo un solo valor que nos indica los tiempos de llegada (segundos). Finalmente se restan los valores de los tiempos de llegada entre las posiciones $i + 1$ e i obteniendo así los tiempos entre llegadas. Se lo puede visualizar de mejor manera en la Figura 3.9:

Dia_del_Mes	Horas	Minutos	Segundos	Tiempo_llegada_segundos
12	17	25	3.0542618	62703.054262
12	17	25	8.5979175	62708.597918
12	17	25	10.7277706	62710.727771
12	17	25	11.6477590	62711.647759
12	17	25	14.4231543	62714.423154
12	17	25	23.5020445	62723.502045
12	17	25	25.2864971	62725.286497

↓

$$\text{Horas} * 3600 + \text{Minutos} * 60 + \text{Segundos}$$

Autor: Salazar W. Fuente: Salazar W.

Figura 3.9 Esquema tiempos entre llegadas

Los tiempos de servicio se los tiene directamente de la variable *tiempo de transacción* del archivo de datos y solo se la transforma a segundos.

3.2.2. Tratamiento de datos atípicos

TIEMPOS ENTRE LLEGADAS

Una primera exploración a los tiempos entre llegadas se la puede visualizar en los siguientes gráficos:

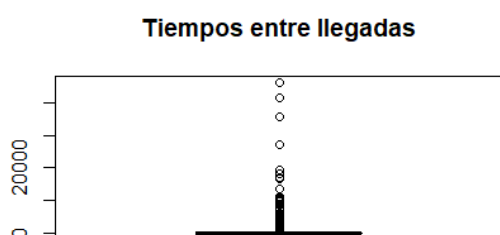


Figura 3.10 D.C tiempos entre llegadas

Autor: Salazar W.

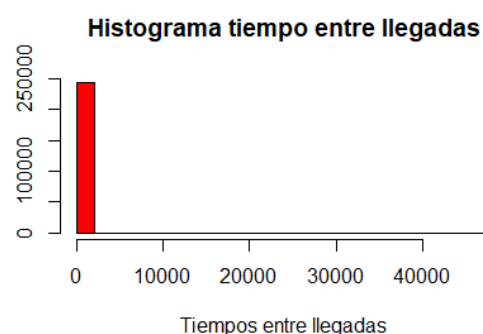


Figura 3.11 Histograma tiempos entre llegadas

Fuente: Salazar W.

Donde el diagrama de caja 3.10 indica la presencia de atípicos y el histograma 3.11 no brinda una idea intuitiva de la distribución a la que se pueden ajustar los

datos. Para tratar los atípicos se recurre a la tabla de frecuencias de la Figura 3.13 donde se puede observar que hasta los 308 segundos se acumulan el 99.35 % de los datos, así todos los valores mayores a 308 segundos se han reemplazado por una distribución uniforme $U \sim [0, 307]$.

Análisis de Una Variable - Tiempos entre llegada							
Tabla de Frecuencias para Tiempos entre llegada							
	Limite	Limite			Frecuencia	Frecuencia	Frecuencia
Clase	Inferior	Superior	Punto Medio	Frecuencia	Relativa	Acumulada	Rel. Acum.
	menor o igual	0		1	0,0000	1	0,0000
1	0	153,95	76,975	232427	0,9626	232428	0,9626
2	153,95	307,9	230,925	7457	0,0309	239885	0,9935
3	307,9	461,85	384,875	985	0,0041	240870	0,9975

Autor: Salazar W. Fuente: Salazar W.

Figura 3.13 Tabla de frecuencias tiempos entre llegadas

Tras esta corrección se obtienen los siguientes gráficos, donde el histograma 3.15 da indicios de que los datos pueden ajustarse a una distribución exponencial y en el diagrama de caja 3.14 se puede observar un sesgo propio de la distribución exponencial.

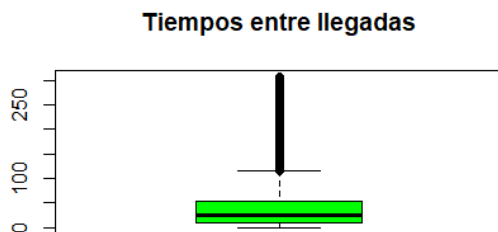


Figura 3.14 D.C tiempos entre llegadas

Autor: Salazar W.

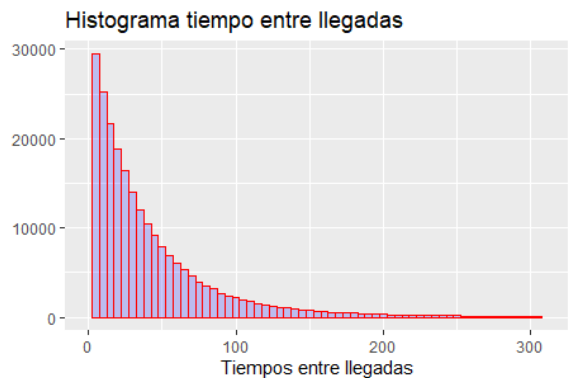


Figura 3.15 Histograma tiempos entre llegadas

Fuente: Salazar W.

Con esta corrección de atípicos los datos de tiempos entre llegadas se pueden someter a un test de bondad de ajuste para encontrar una distribución a la cual se ajusten los datos.

TIEMPOS DE SERVICIO

De manera similar para los tiempos de servicio se presentan gráficos previos pa-

ra intentar describir su comportamiento, el diagrama de caja 3.17 indica la posible presencia de atípicos y el histograma 3.18 no brinda información sobre la distribución a la que se pueden ajustar los datos.

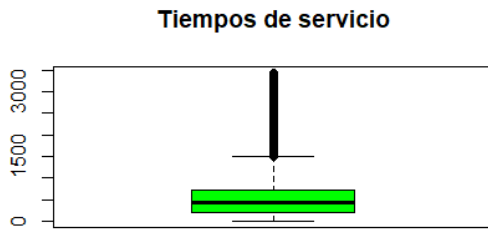


Figura 3.17 D.C tiempos de servicio

Autor: Salazar W.

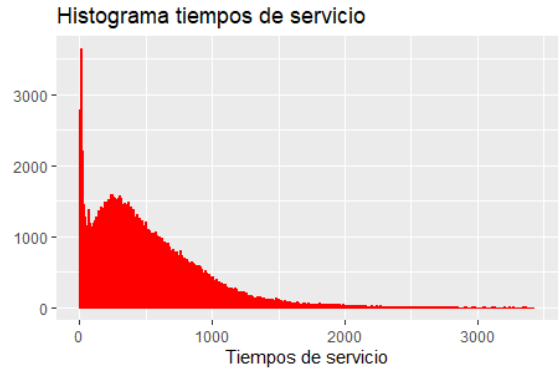


Figura 3.18 Histograma tiempos de servicio

Fuente: Salazar W.

De la misma manera que para los tiempos entre llegadas se realiza una tabla de frecuencia, de donde se encuentra que hasta los 1918 segundos se acumulan el 99.0% de los datos, así todos los valores mayores a 1918 se reemplazan por una distribución uniforme $U \sim [0, 1918]$. Adicionalmente se puede notar que existe una acumulación de datos entre 0 y 60 segundos lo cual se considera también como atípicos, pues a pesar de que hay servicios que duran menos de un minuto como las actualizaciones de clave, al no influir en la generación de tiempos de espera altos los vamos a desestimar del modelo para poder encontrar una distribución a la que se ajusten de mejor manera los datos restantes. Entonces se procede a retirar dichos valores obteniendo así los siguientes gráficos:

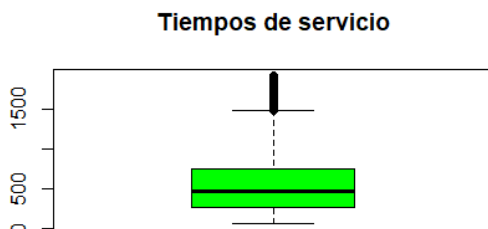


Figura 3.20 DC tiempos de servicio

Autor: Salazar W.

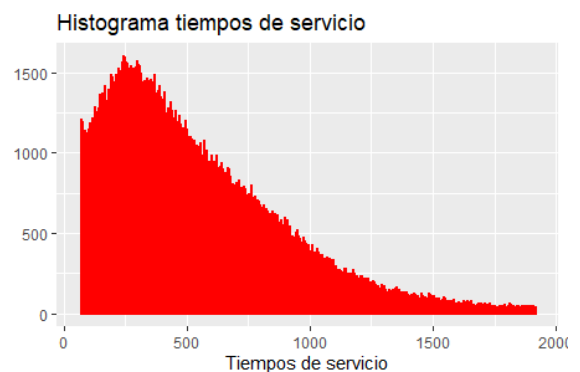


Figura 3.21 Histograma tiempos de servicio

Fuente: Salazar W.

En el diagrama de caja 3.20 se observa un sesgo que se puede visualizar más cla-

ramente en el histograma 3.21. Con esta corrección de atípicos los datos de tiempos de servicio se pueden someter a un test de bondad de ajuste para encontrar una distribución a la cual se ajusten los datos.

3.3. Ajuste de datos

TIEMPOS ENTRE LLEGADAS

En este punto se separan los datos para describir el comportamiento en las diferentes horas y meses, posteriormente se usa la función «useFitdist» del programa R obteniendo que la distribución a la que más se ajustan los datos es una distribución exponencial. Para validar esto se realiza el test de bondad de ajuste «ks.test» (Kolmogorov-Smirnov) a cada segmento de datos, contrastando las hipótesis:

H_0 : Los datos analizados siguen una distribución exponencial

H_1 : Los datos analizados no siguen una distribución exponencial

De donde se obtienen los *p-valores* de la Figura 3.23:

		H8	H9	H10	H11	H12	H13	H14	H15	H16
P valores	Junio	0,7070	0,2440	0,0060	0,5070	0,0300	0,1840	0,0380	0,6990	0,0540
	Julio	0,3970	0,1370	0,1460	0,2930	0,1860	0,0070	0,6400	0,1780	0,3550
	Agosto	0,2300	0,1190	0,0200	0,3490	0,4700	0,1740	0,2850	0,3870	0,0780
	Septiembre	0,1450	0,0100	0,7060	0,0220	0,7550	0,8330	0,5680	0,0940	0,2410
	Octubre	0,2470	0,0550	0,2780	0,2760	0,0090	0,3800	0,4450	0,6440	0,7590
	Noviembre	0,5660	0,6710	0,6050	0,3740	0,0870	0,8450	0,4420	0,4280	0,5210
	Diciembre	0,5800	0,0080	0,5730	0,1710	0,6100	0,8090	0,6590	0,1240	0,7580

Autor: Salazar W. **Fuente:** Salazar W.

Figura 3.23 P-valores para tiempos entre llegadas

Notemos que si usamos los p-valores, la regla de decisión está asociada al nivel de significación α de la siguiente manera:

Si p-valor $\geq \alpha \Rightarrow$ No rechazar H_0

Si p-valor $< \alpha \Rightarrow$ Rechazar H_0

Por tanto, si tomamos un nivel $\alpha = 0,05$ entonces el 87.30 % de los valores de la Figura 3.23 cumplen la regla para no rechazar H_0 , más aún si tomamos un nivel $\alpha = 0,01$ entonces el 95.24 % de los valores de la Figura 3.23 nos indican no rechazar H_0 . Por tanto, podemos concluir que los tiempos entre llegadas se ajustan a una distribución exponencial.

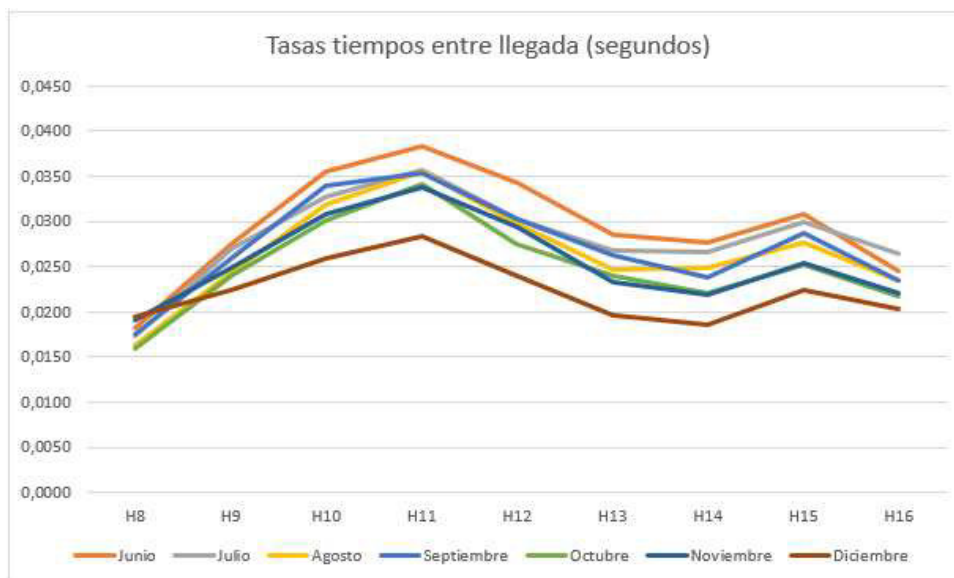
Ahora bien, dado que los tiempos entre llegadas siguen una distribución exponencial (Tiempos entre llegada $\sim Exp(\lambda)$) se buscan los parámetros λ o tasas de llegada (clientes / segundo) obteniendo los valores de la Figura 3.24:

	H8	H9	H10	H11	H12	H13	H14	H15	H16
Tasas (clientes/segundo)	0,0181	0,0274	0,0355	0,0383	0,0343	0,0286	0,0276	0,0308	0,0245
Junio	0,0174	0,0269	0,0327	0,0358	0,0303	0,0268	0,0266	0,0299	0,0265
Julio	0,0163	0,0243	0,0318	0,0355	0,0297	0,0247	0,0249	0,0276	0,0234
Agosto	0,0174	0,0259	0,0339	0,0353	0,0302	0,0262	0,0238	0,0288	0,0234
Septiembre	0,0160	0,0239	0,0301	0,0341	0,0275	0,0240	0,0221	0,0251	0,0217
Octubre	0,0190	0,0248	0,0307	0,0338	0,0294	0,0233	0,0218	0,0254	0,0221
Noviembre	0,0195	0,0223	0,0259	0,0283	0,0240	0,0196	0,0186	0,0225	0,0203
Diciembre	0,0177	0,0251	0,0315	0,0344	0,0293	0,0247	0,0236	0,0272	0,0231

Autor: Salazar W. **Fuente:** Salazar W.

Figura 3.24 Tasas de llegadas (cliente/seg)

Para obtener el valor de las tasas de llegada en cada hora el día, se toma el promedio de las tasas a esa hora en cada mes, como se puede ver en la Figura 3.24. Gráficamente las tasas de los tiempos entre llegadas tienen el comportamiento de la Figura 3.25:



Autor: Salazar W. Fuente: Salazar W.

Figura 3.25 Gráfico tasas de llegadas (cliente/seg)

Finalmente, las tasas se expresan en minutos, y se puede observar las tasas finales con las que se trabajara en cada hora en la Figura 3.26, además se indica que al día llegan en promedio 852 personas para recibir atención en la PGA.

	H8	H9	H10	H11	H12	H13	H14	H15	H16	
Junio	1,09	1,64	2,13	2,30	2,06	1,71	1,66	1,85	1,47	
Julio	1,04	1,62	1,96	2,15	1,82	1,61	1,59	1,79	1,59	
Agosto	0,98	1,46	1,91	2,13	1,78	1,48	1,49	1,65	1,41	
Septiembre	1,05	1,55	2,04	2,12	1,81	1,57	1,43	1,73	1,40	
Octubre	0,96	1,44	1,81	2,04	1,65	1,44	1,33	1,51	1,30	
Noviembre	1,14	1,49	1,84	2,03	1,76	1,40	1,31	1,53	1,32	
Diciembre	1,17	1,34	1,55	1,70	1,44	1,17	1,12	1,35	1,22	
	1,06	1,50	1,89	2,07	1,76	1,48	1,42	1,63	1,39	851,96

Autor: Salazar W. Fuente: Salazar W.

Figura 3.26 Tasas de llegadas (cliente/min)

Con este primer análisis se obtiene que los tiempos entre llegadas se ajustan a una distribución exponencial, por tanto, el modelo de teoría de colas a usar sería un (M/B/s), pues aún no se tiene información de los tiempos de servicio.

TIEMPOS DE SERVICIO

Similar a lo realizado en la sección anterior se separan los datos para describir

el comportamiento en las diferentes horas y meses, sin embargo, la función «useFit-dist» nos sugiere que la distribución a la que más se ajustan los datos es una distribución Weibull. Para validar esto se realiza el test de bondad de ajuste «ks.test», de donde se obtienen los *p-valores* de la Figura 3.27:

	H8	H9	H10	H11	H12	H13	H14	H15	H16	
P valores	Junio	0,5250	0,2970	0,4330	0,6280	0,7230	0,5250	0,4820	0,4370	0,4980
	Julio	0,4280	0,2140	0,3160	0,0470	0,7130	0,4210	0,5610	0,6330	0,3260
	Agosto	0,9260	0,5390	0,6360	0,2180	0,4870	0,5820	0,7820	0,7820	0,2250
	Septiembre	0,5560	0,4530	0,6450	0,2170	0,1980	0,7250	0,5170	0,7520	0,8060
	Octubre	0,7970	0,4000	0,1890	0,8720	0,5900	0,5420	0,8630	0,5960	0,4600
	Noviembre	0,7320	0,3360	0,6930	0,3580	0,7570	0,6730	0,9630	0,7530	0,7290
	Diciembre	0,2860	0,3170	0,0690	0,2120	0,2690	0,0350	0,7430	0,6000	0,6010

Autor: Salazar W. **Fuente:** Salazar W.

Figura 3.27 P-valores tiempos de servicio (min)

Notemos que ahora las hipótesis a contrastar con el «ks.test» son:

H_0 : Los datos analizados siguen una distribución Weibull.

H_1 : Los datos analizados no siguen una distribución Weibull.

Donde, al tomar un nivel $\alpha = 0,05$ entonces todos los valores de la Figura 3.27 cumplen la regla para no rechazar la H_0 . Por tanto, podemos concluir que los tiempos de servicio se ajustan a una distribución Weibull.

Ahora bien, dado que los tiempos de servicio siguen una distribución Weibull (Tiempos de servicio $\sim Weibull(\beta, \theta)$), donde β es el parámetro de forma y θ es el parámetro escala. Se presentan los valores para ambos parámetros en cada hora del día:

	H8	H9	H10	H11	H12	H13	H14	H15	H16	
Shape	Junio	1,62963	1,57137	1,50423	1,52853	1,47542	1,39627	1,47598	1,47097	1,43091
	Julio	1,59939	1,51631	1,54818	1,50065	1,43562	1,36814	1,45682	1,46347	1,40794
	Agosto	1,57343	1,48222	1,46595	1,47522	1,40019	1,43446	1,43984	1,47291	1,42711
	Septiembre	1,54358	1,50709	1,50955	1,46878	1,41249	1,40154	1,44835	1,43184	1,48554
	Octubre	1,62005	1,57034	1,45327	1,4984	1,48113	1,45252	1,50464	1,51563	1,48605
	Noviembre	1,6542	1,55413	1,45077	1,50512	1,44589	1,44427	1,48693	1,51517	1,47831
	Diciembre	1,56545	1,52557	1,49863	1,43243	1,39059	1,4016	1,43021	1,43018	1,44286
		1,5980	1,5324	1,4901	1,4870	1,4345	1,4141	1,4633	1,4715	1,4512

Autor: Salazar W. **Fuente:** Salazar W.

Figura 3.28 Parámetro de forma tiempos de servicio

	H8	H9	H10	H11	H12	H13	H14	H15	H16	
Scale	Junio	11,5183	10,9289	10,302	10,124	9,61322	9,76695	10,6458	10,0611	9,37004
	Julio	10,9659	10,1127	10,297	9,70737	9,42371	8,92973	9,62179	10,1249	9,60136
	Agosto	10,6496	10,2139	9,88435	9,83221	9,15121	9,60844	9,68675	9,86679	9,32507
	Septiembre	10,2225	10,1437	9,88789	9,80059	9,22882	9,52674	9,58739	9,88687	10,0104
	Octubre	10,4714	10,293	9,83822	9,65313	9,94782	9,44592	10,2613	10,4955	10,6266
	Noviembre	10,9017	10,2826	9,56267	9,83406	9,47132	9,40787	10,3645	10,6054	10,4314
	Diciembre	11,3051	10,4265	10,0756	9,98087	9,63134	9,84387	10,2396	10,2121	10,0374
		10,8621	10,3431	9,9782	9,8475	9,4953	9,5042	10,0582	10,1789	9,9146

Autor: Salazar W. Fuente: Salazar W.

Figura 3.29 Parámetro de escala tiempos de servicio

Para obtener el valor de los parámetros en cada hora el día, se toma el promedio de los parámetros a esa hora en cada mes, como se puede ver en las Figuras 3.28 y 3.29. Finalmente se presenta la Figura 3.30 con las tasas de servicio μ para cada hora del día.

	H8	H9	H10	H11	H12	H13	H14	H15	H16	
Tasas (clientes/minuto)	Junio	0,09706	0,10202	0,10766	0,10979	0,11515	0,11235	0,10379	0,10985	0,11755
	Julio	0,10183	0,10993	0,10813	0,11436	0,11703	0,12259	0,11489	0,10913	0,11452
	Agosto	0,1047	0,10844	0,11193	0,11259	0,12014	0,11474	0,11384	0,11216	0,11824
	Septiembre	0,10895	0,10949	0,11225	0,11296	0,11929	0,11532	0,11519	0,11152	0,11063
	Octubre	0,10674	0,10839	0,11238	0,11484	0,11134	0,11698	0,10816	0,10577	0,10423
	Noviembre	0,10275	0,10841	0,11155	0,11289	0,11662	0,11741	0,10686	0,10469	0,1061
	Diciembre	0,09855	0,10665	0,11025	0,1105	0,11408	0,11172	0,10756	0,10789	0,11001
		0,10294	0,10762	0,11116	0,11256	0,11623	0,11587	0,11004	0,10871	0,11161

Autor: Salazar W. Fuente: Salazar W.

Figura 3.30 Tasas de servicio (cliente/min)

Con este segundo análisis se obtiene que los tiempos de servicio se ajustan a una distribución Weibull.

Finalmente se concluye que el modelo de teoría de colas a usar sería un modelo (M/G/s).

3.4. Aplicación del modelo M/G/s

La aproximación de la esperanza del tiempo de espera de los clientes en cola viene dada por:

$$EW(M/G/s) \simeq \left(\frac{\beta_2}{2\beta_1^2} \right) \frac{2EW(M/M/s)EW(M/D/s)}{2A(s)EW(M/D/s) + (1 - A(s))EW(M/M/s)}$$

En primer lugar, se calcula el término $A(s)$, que viene dado por:

$$A(s) = \frac{\beta_2}{(s-1)\beta_1 \int_{v=0}^{\infty} (1-H(v))^s dv} - \frac{s+1}{s-1} \quad (3.1)$$

donde:

$$H(v) = \int_{u=0}^v \frac{1-B(u)}{\beta_1} du, \quad v > 0$$

Ahora bien, como los tiempos de servicio siguen una distribución $Weibull(\beta, \theta)$, entonces su función de distribución viene dada por:

$$B(u) = 1 - \exp \left\{ - \left(\frac{u}{\theta} \right)^\beta \right\}$$

Donde β_1 y β_2 son los dos primeros momentos de la variable aleatoria *Tiempos de servicio* que siguen una distribución $B(u)$. Ahora bien, notemos que el parámetro de forma β indica el comportamiento de las tasas de servicio (si crecen, decrecen o se mantienen constantes), para nuestro estudio se tienen tasas de servicio constantes, es decir, no cambian dependiendo el número de clientes en el sistema; por tal motivo si un servidor atiende a razón de 6 clientes por hora, entonces 11 servidores atienden a 66 clientes en una hora. Así, al tener una tasa de servicio constante se toma $\beta = 1$, con ello:

$$B(u) = 1 - \exp \left\{ - \left(\frac{u}{\theta} \right) \right\}$$

Así:

$$\begin{aligned} H(v) &= \int_{u=0}^v \frac{e^{-\left(\frac{u}{\theta}\right)}}{\beta_1} du \\ &= - \left(\frac{\theta}{\beta_1} \right) \left(e^{-\left(\frac{v}{\theta}\right)} - 1 \right) \end{aligned}$$

Regresando a la ecuación (3.1), calculamos:

$$\begin{aligned}\int_{v=0}^{\infty} (1 - H(v))^s dv &= \int_{v=0}^{\infty} \left[1 + \left(\frac{\theta}{\beta_1} \right) \left(e^{-\left(\frac{v}{\theta}\right)} - 1 \right) \right]^s dv \\ &= \int_{v=0}^{\infty} \left[\left(\frac{\theta}{\beta_1} \right) e^{-\left(\frac{v}{\theta}\right)} + \left(1 - \frac{\theta}{\beta_1} \right) \right]^s dv\end{aligned}$$

Para abordar esta integral se usa la fórmula del *binomio de Newton*, la cual establece que el desarrollo de la n -ésima potencia de un binomio $x + y$, con $x, y \in \mathbb{R}$ y $n \in \mathbb{Z}^+$, puede ser expandida en una suma de la forma:

$$\begin{aligned}(x + y)^n &= \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k \\ &= \binom{n}{0} x^n + \binom{n}{1} x^{n-1} y + \binom{n}{2} x^{n-2} y^2 + \dots + \binom{n}{n} y^n\end{aligned}$$

donde $\binom{n}{k}$ es el *coeficiente binomial*, el cual representa el número de formas de escoger k elementos de un conjunto con n elementos. Así, regresando a nuestra integral, dado que $s \in \mathbb{N}$ tenemos:

$$\begin{aligned}\left[\left(\frac{\theta}{\beta_1} \right) e^{-\left(\frac{v}{\theta}\right)} + \left(1 - \frac{\theta}{\beta_1} \right) \right]^s &= \binom{s}{0} \left[\left(\frac{\theta}{\beta_1} \right) e^{-\left(\frac{v}{\theta}\right)} \right]^s + \binom{s}{1} \left[\left(\frac{\theta}{\beta_1} \right) e^{-\left(\frac{v}{\theta}\right)} \right]^{s-1} \left(1 - \frac{\theta}{\beta_1} \right) \\ &\quad + \binom{s}{2} \left[\left(\frac{\theta}{\beta_1} \right) e^{-\left(\frac{v}{\theta}\right)} \right]^{s-2} \left(1 - \frac{\theta}{\beta_1} \right)^2 + \dots + \binom{s}{s} \left(1 - \frac{\theta}{\beta_1} \right)^s\end{aligned}$$

De donde,

$$\begin{aligned}\int \left[\left(\frac{\theta}{\beta_1} \right) e^{-\left(\frac{v}{\theta}\right)} + \left(1 - \frac{\theta}{\beta_1} \right) \right]^s dv &= \int \left[\left(\frac{\theta}{\beta_1} \right) e^{-\left(\frac{v}{\theta}\right)} \right]^s dv + \int \binom{s}{1} \left[\left(\frac{\theta}{\beta_1} \right) e^{-\left(\frac{v}{\theta}\right)} \right]^{s-1} \left(1 - \frac{\theta}{\beta_1} \right) dv + \\ &\quad \int \binom{s}{2} \left[\left(\frac{\theta}{\beta_1} \right) e^{-\left(\frac{v}{\theta}\right)} \right]^{s-2} \left(1 - \frac{\theta}{\beta_1} \right)^2 dv + \dots + \int \left(1 - \frac{\theta}{\beta_1} \right)^s dv\end{aligned}$$

Ahora bien, fijándonos en la última integral de la suma y definiendo Ψ como la suma de todas las integrales anteriores notemos que:

$$\begin{aligned}\int_0^{\infty} \left[\left(\frac{\theta}{\beta_1} \right) e^{-\left(\frac{v}{\theta}\right)} + \left(1 - \frac{\theta}{\beta_1} \right) \right]^s dv &= \Psi + \lim_{k \rightarrow \infty} \int_0^k \left(1 - \frac{\theta}{\beta_1} \right)^s dv \\ &= \Psi + \lim_{k \rightarrow \infty} \left[k \left(1 - \frac{\theta}{\beta_1} \right)^s \right]\end{aligned}$$

$$\begin{aligned}
&= \Psi + \lim_{k \rightarrow \infty} \left[k \left(1 - \frac{\theta}{\beta_1} \right)^s \right] \rightarrow \infty \\
&= \begin{cases} \infty & \text{si } s \text{ es par} \\ -\infty & \text{si } s \text{ es impar} \end{cases}
\end{aligned}$$

Dado que el término final diverge entonces la suma de todas las integrales también diverge, por tanto

$$\int_{v=0}^{\infty} \left[1 + \left(\frac{\theta}{\beta_1} \right) (e^{-\frac{v}{\theta}} - 1) \right]^s dv$$

diverge. Así:

$$A(s) = \frac{\beta_2}{(s-1)\beta_1 \int_{v=0}^{\infty} (1-H(v))^s dv} \frac{s+1}{s-1} \quad (3.2)$$

$$= -\frac{s+1}{s-1} \quad (3.3)$$

Posteriormente, para los cálculos de $EW(M/M/s)$ y $EW(M/D/s)$, se usan la fórmulas (2.2) y (2.7):

$$EW(M/M/s) = \frac{P_0 (\lambda/\mu)^s}{\lambda s!} \left[\frac{\rho}{(1-\rho)^2} \right]$$

$$EW(M/D/s) \simeq \frac{1}{2} \{1 + f(s)g(\rho)\} EW(M/M/s)$$

Para usar estas fórmulas se usan los valores de la tabla de la Figura 3.31 segmentada por horas, donde los parámetros β_1 ($E[X]$) y β_2 ($E[X^2]$) son calculados mediante la fórmula:

$$E[X^n] = \theta^n \Gamma \left(1 + \frac{n}{\beta} \right)$$

donde X es la variable aleatoria *Tiempos de servicio* la cual sigue una distribución Weibull.

Parámetro	H8	H9	H10	H11	H12	H13	H14	H15	H16
β_1	9,7393	9,3020	9,0122	8,8840	8,5952	8,6254	9,0864	9,1850	8,9368
β_2	134,4704	125,7690	119,7951	116,4156	111,5568	113,3344	122,6458	124,7962	119,5930
λ	1,0604	1,5047	1,8912	2,0658	1,7604	1,4827	1,4176	1,6294	1,3868
μ	0,1029	0,1076	0,1112	0,1126	0,1162	0,1159	0,1100	0,1087	0,1116
β	1,5980	1,5324	1,4901	1,4870	1,4345	1,4141	1,4633	1,4715	1,4512
θ	10,8621	10,3431	9,9782	9,8475	9,4953	9,5042	10,0582	10,1789	9,9146

Autor: Salazar W. Fuente: Salazar W.

Figura 3.31 Parámetros por hora

Finalmente, del uso de estas fórmulas se obtienen los resultados de la tabla de la Figura 3.32 para las 8:00 horas (el resultado de las demás horas se la tiene en Anexos):

s	ρ	A(s)	EW(M/M/s)	EW(M/D/s)	EW(M/G/s)
11	0,9389	-1,200	10,722044	5,473650	7,9605
12	0,8607	-1,182	2,954317	1,559372	2,3663
13	0,7944	-1,167	1,202371	0,657781	1,0476
14	0,7377	-1,154	0,546310	0,310370	0,5221
15	0,6885	-1,143	0,257470	0,152134	0,2722
16	0,6455	-1,133	0,122039	0,075082	0,1441
17	0,6075	-1,125	0,057259	0,036703	0,0762
18	0,5738	-1,118	0,026350	0,017595	0,0399
19	0,5436	-1,111	0,011829	0,008215	0,0205
20	0,5164	-1,105	0,005162	0,003718	0,0103
21	0,4918	-1,100	0,002186	0,001625	0,0050
22	0,4694	-1,095	0,000897	0,000686	0,0023
23	0,4490	-1,091	0,000356	0,000279	0,0010
24	0,4303	-1,087	0,000137	0,000109	0,0004
25	0,4131	-1,083	0,000051	0,000041	0,0002

Autor: Salazar W. Fuente: Salazar W.

Figura 3.32 Resultados 8:00 horas

Estos cálculos se los realiza desde 11 servidores que es cuando el factor de utilización ρ es menor a uno, hasta 25 servidores.

3.4.1. Cálculo de métricas principales

Es claro que el objetivo del estudio se enfoca en el tiempo de espera en cola en un sistema, sin embargo, para tener un contexto más amplio del problema se calculan las métricas L , L_q y W , mediante las ecuaciones (1.3), (1.4), (1.5) y los valores de los parámetros de la tabla de la Figura 3.31.

Con esto se tienen los resultados de la Figura 3.33. Con estos primeros resultados de las cuatro métricas principales se puede estimar un número de servidores tal que los tiempos de espera sean bajos, sin embargo, no se puede asegurar el cumplimiento de la política de calidad, para ello se realizarán simulaciones con las cuales se valida que los resultados obtenidos teóricamente converjan a los de la simulación y posteriormente extraer los percentiles 1 y 99 para verificar que esos valores efectivamente no sobrepasan los 20 minutos y asegurar el cumplimiento de la política de calidad.

s	W	Lq	L
11	17,675	8	19
12	12,081	3	13
13	10,762	1	11
14	10,236	1	11
15	9,987	0	11
16	9,858	0	10
17	9,791	0	10
18	9,754	0	10
19	9,735	0	10
20	9,725	0	10
21	9,719	0	10
22	9,717	0	10
23	9,715	0	10
24	9,715	0	10
25	9,715	0	10

Autor: Salazar W. Fuente: Salazar W.

Figura 3.33 Métricas 8:00 horas

Capítulo 4

Simulación y Resultados

4.1. Simulación de la línea de espera

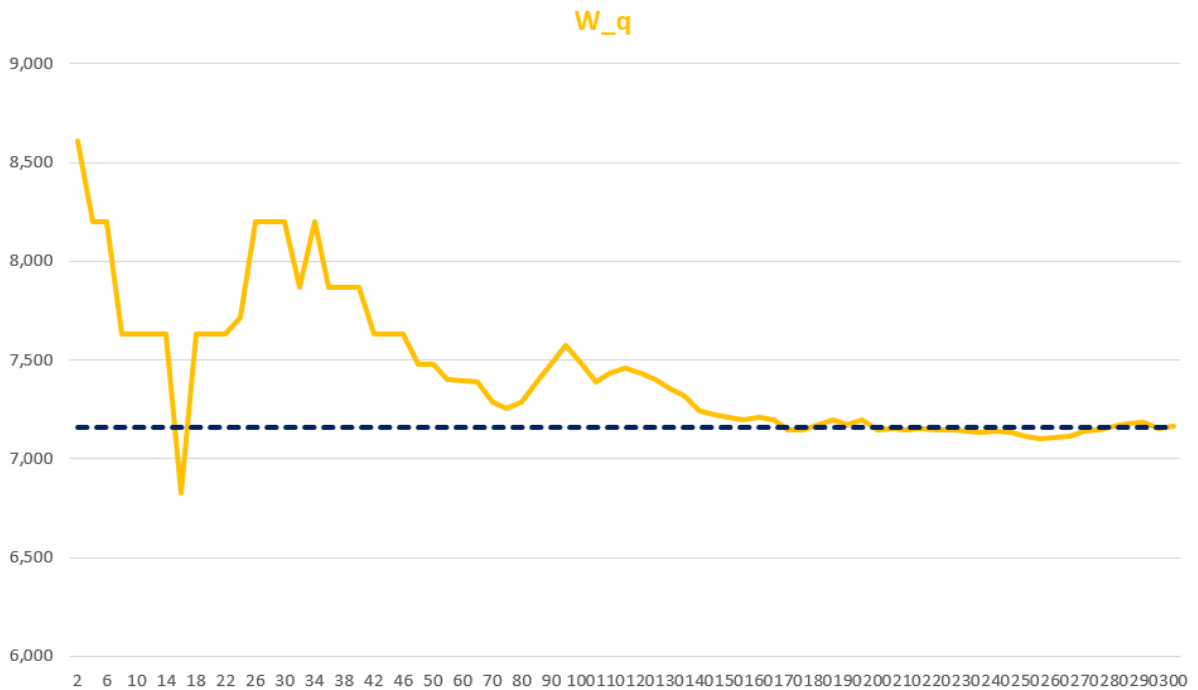
La simulación para la línea de espera se la realiza usando la librería *simmer* del programa R, bajo los supuestos de que los tiempos entre llegadas se ajustan a una distribución exponencial $Exp(\lambda)$ y los tiempos de servicio se ajustan a una distribución Weibull $Weibull(\beta, \theta)$ con los parámetros de la tabla de la Figura 3.31 para cada hora.

Para cada número de servidores se realizaron entre 200 y 300 simulaciones del comportamiento de una línea de espera, la línea de espera se puede visualizar en la Figura 4.1:

name	start_time	end_time	activity_time	finished
cliente0	2.358636	5.669003	3.310367	TRUE
cliente2	2.855846	8.645409	5.789562	TRUE
cliente3	4.498892	17.071480	12.572588	TRUE
cliente4	4.583696	17.079740	12.496045	TRUE
cliente8	9.874275	18.113838	8.239563	TRUE
cliente5	4.774722	19.792586	15.017864	TRUE
cliente9	10.459934	19.795573	9.335638	TRUE
cliente7	6.996444	21.110065	14.113621	TRUE
cliente6	5.491649	21.392584	15.900935	TRUE
cliente10	10.489845	22.591687	12.101842	TRUE
cliente13	11.704018	25.636526	8.565047	TRUE
cliente12	11.214387	25.809133	14.594746	TRUE
cliente11	11.200610	26.389247	15.188637	TRUE
cliente16	13.638977	28.295602	8.503016	TRUE

Autor: Salazar W. Fuente: Salazar W.

Figura 4.1 Simulación



Autor: Salazar W. **Fuente:** Salazar W.

Figura 4.2 Convergencia de la simulación

Notemos que lo que se pretende con la simulación es estimar la media de W_q , se puede observar en la Figura 4.2 que con un tamaño de muestra menor a 30 simulaciones los valores de W_q son muy variables; sin embargo, mientras el número de simulaciones es mayor se observa la convergencia a un valor. Se puede notar gráficamente que la elección de 200 o 300 simulaciones es apropiada pues se obtiene una buena aproximación al ser una muestra grande y representativa.

Posteriormente, de cada simulación se extrae el tiempo promedio de espera en fila W_q , obteniendo una muestra entre 200 y 300 datos de la cual se puede extraer la esperanza de los tiempos promedio en fila (cuartil 50) y los percentiles 1 y 99, para saber entre que valores se encuentran el 98 % de los datos. Estos valores se resumen en la tabla de la Figura 4.3.

Servidores	50%	1%	99%	Nsim
11	7,166	4,323	14,892	300
12	2,118	1,519	3,380	300
13	0,890	0,635	1,278	300
14	0,416	0,275	0,648	300
15	0,201	0,143	0,289	300
16	0,099	0,060	0,166	300
17	0,047	0,022	0,082	300
18	0,021	0,010	0,044	300
19	0,009	0,003	0,020	300
20	0,004	0,001	0,016	300
21	0,001	0,000	0,007	300
22	0,001	0,000	0,005	300

Autor: Salazar W. Fuente: Salazar W.

Figura 4.3 Simulación 8:00 horas

Tras la simulación se pueden evidenciar dos resultados principales:

- i) Los resultados de la simulación difieren de los resultados obtenidos teóricamente (tabla de la Figura 4.3) en a lo sumo un minuto, por lo cual se puede validar que los resultados obtenidos con el modelo M/G/s son correctos.
- ii) Para el caso de las 8:00 horas si se elige el número de 11 servidores para brindar el servicio, se puede evidenciar que el 98 % de los datos de la simulación tienen valores menores a 20 minutos, por ello este es el mejor número para cumplir la política de calidad.

Se puede notar que, si se elige un número de servidores mayor a 11, aún se garantiza el cumplimiento de la política de calidad, sin embargo, hacerlo incurre en un gasto de nómina adicional. Así, el criterio para la elección del mejor número de servidores en cada hora del día es *el primer menor número de servidores con el que se satisfaga la política de calidad*.

4.2. Evaluación de los resultados

Los resultados presentados en la sección anterior corresponden a las 8:00 horas, tras realizar los mismos cálculos para cada hora de atención se presenta un resumen en la Figura 4.4, del número de servidores elegido a cada hora del día junto al intervalo que contiene el 98 % de los datos para asegurar el cumplimiento de la política de calidad.

Hora	Servidores	50%	1%	99%	Nsim
8	11	7,166	4,323	14,892	300
9	15	4,749	2,961	9,673	300
10	18	5,063	2,926	11,216	300
11	20	2,456	1,552	4,075	200
12	16	5,743	3,451	12,220	200
13	14	3,703	2,341	6,404	250
14	14	4,186	2,560	7,780	300
15	16	4,822	2,890	9,292	300
16	14	2,560	1,695	4,540	250

Autor: Salazar W. Fuente: Salazar W.

Figura 4.4 Elección de servidores en cada hora de atención

Finalmente, se puede verificar que los resultados teóricos para todas las horas del día difieren de las simulaciones en a lo sumo 48 segundos, por lo cual se valida que los resultados son correctos; además se puede notar que el 98 % de los datos de la simulación siempre son menores a 20 minutos, por lo cual se asegura también el cumplimiento de la política de calidad con esta elección de servidores.

4.3. Comparación costos actuales vs propuestos

No se cuenta con el detalle de la remuneración a los servidores dentro del sistema operativo de servicios, por ello la unidad de medida para la comparación de costos será en *horas*, es decir, se medirá cuantas horas al mes de trabajo se requieren con el número de servidores actual y se contrastará con el número de horas al mes que se necesitará con el número de servidores propuesto. El costo se lo puede representar de esta manera pues el número de horas multiplicado por el salario de una hora será igual a la remuneración que se debe satisfacer.

Así, el número de servidores activos en cada hora del día actualmente, y el costo en horas generado al sistema se resume la tabla de la Figura 4.5:

Hora	Servidores	Costo al día (horas)	Horas al mes (horas)
8:00	18	18	360
9:00	18	18	360
10:00	19	19	380
11:00	19	19	380
12:00	19	19	380
13:00	19	19	380
14:00	18	18	360
15:00	18	18	360
16:00	17	17	340

Autor: Salazar W. Fuente: Salazar W.

Figura 4.5 Costos actuales

Finalmente, de la Figura 4.6 se puede evidenciar que con el número de servidores propuestos en la Figura 4.4, se tiene un ahorro en horas en la mayoría de las horas de atención, con excepción de las 11:00 donde se ve la necesidad de aumentar un servidor a los que actualmente operan. En total se ha generado una disminución de horas del 16,36 % gracias a la utilización del modelo.

Hora	Servidores	Costo al día (horas)	Costo al mes (horas)	Diferencia (horas)
8:00	11	11	220	140
9:00	15	15	300	60
10:00	18	18	360	20
11:00	20	20	400	-20
12:00	16	16	320	60
13:00	14	14	280	100
14:00	14	14	280	80
15:00	16	16	320	40
16:00	14	14	280	60

Autor: Salazar W. Fuente: Salazar W.

Figura 4.6 Costos propuestos

Capítulo 5

Conclusiones y Recomendaciones

5.1. Conclusiones

- El número de servidores que se propone en cada hora del día son: 11 servidores a las 8:00, 15 servidores a las 9:00, 18 servidores a las 10:00, 20 servidores a las 11:00, 16 servidores a las 12:00, 14 servidores a las 13:00, 14 servidores a las 14:00, 16 servidores a las 15:00 y 14 servidores a las 16:00. Con estos valores se verifica el cumplimiento de la política de calidad, además de un ahorro en horas del 16,36 %.
- A partir del análisis precedente, es posible observar que con el cambio sugerido (usar la aproximación EW(M/D/s) de Kimura [10] en lugar de la de Crommelin [5] en la aproximación EW(M/G/s) de Boxma [2]) para el modelo M/G/s se ha podido usar estas aproximaciones en la práctica, pues ya no se tiene la limitación del cálculo de factoriales sumamente grandes en la implementación de las fórmulas.
- En virtud de lo estudiado, es posible sostener que un estudio técnico que permite describir el comportamiento de la línea de espera en las instalaciones de la PGA, ayuda a la elección de un correcto número de servidores; que por ende genera un ahorro en horas en esta agencia como se puede ver en la tabla de la Figura 4.6.
- Se ha podido evidenciar el cumplimiento de la política de calidad en este sistema operativo de servicios, se lo puede visualizar en la tabla de la Figura 4.4 gracias a la extracción de los percentiles 1 y 99 que nos indican entre que valores estarán el 98 % de los datos de la simulación. Además, se puede notar que

no se cae en costos excesivos, por el contrario, se puede observar un ahorro.

- Si bien es cierto que las fórmulas para el modelo M/G/s funcionan para tiempos de servicio con cualquier distribución arbitraria, es necesario encontrar a que distribución de probabilidad se ajustan en una aplicación concreta; en este caso los cálculos de la Sección 3.4 se realizan con tiempos de servicio ajustados a una distribución *Weibull*(β, θ), los cuales cambiarían si los datos de ajustaran a otra distribución.

5.2. Recomendaciones

- Sería recomendable que pese a los resultados finales obtenidos, se considere aumentar un servidor más a las 8:00, 10:00 y 12:00 horas, pues si bien es cierto que el 98% de las simulaciones tiene valores menores a 20 minutos; siguen siendo superiores a 11 minutos, y por eventos no capturados por el modelo los tiempos de espera podrían subir.
- Para la realización del proyecto se contó con una archivo de observación de siete meses, por lo cual, sería recomendable proporcionar una archivo con mayor número de datos para capturar de mejor manera el comportamiento de la línea de espera.
- El presente proyecto se lo realizó para la agencia PGA del S.R.I por cual sería recomendable replicarlo en otras agencias de este sistema operativo de servicios, con el fin de realizar un intercambio de servidores entre agencias donde haya mucha afluencia y las que tengan poca afluencia de clientes, y así mejorar la política de calidad en las diferentes agencias sin la necesidad de contratación de nuevos servidores.

Bibliografía

- [1] Ameh, N., Oyefabi, M., y Sabo, B. Application of queuing theory to patient satisfaction at a tertiary hospital in Nigeria. *Nigerian Medical Journal*, 54(1):64, 2013.
- [2] Boxma, O. J., Cohen, J. W., y Huffels, N. Approximations of the mean waiting time in an M/G/s Queueing system. *Operations Research*, 27(6):1115–1127, December 1979.
- [3] Cazorla, F. Análisis estadístico mediante teoría de colas para determinar el nivel de satisfacción del paciente atendido en el departamento de admisiones del hospital provincial general docente de Riobamba. February 2014. Recuperado a partir de: <http://dspace.esoch.edu.ec/handle/123456789/3207> (visitado 02-06-2022).
- [4] Cosmetatos, G. Notes approximate explicit formulae for the average queueing time in the processes (M/D/r) and (D/M/r). *INFOR: Information Systems and Operational Research*, 13(3):328–331, October 1975.
- [5] Crommelin, C. Delay probability formulae. *Post Office Electrical Engineers Journal*, 26:266–274, 1934.
- [6] Gómez, A. Aplicación de teoría de colas en una entidad financiera: herramienta para el mejoramiento de los procesos de atención al cliente. *Revista Universidad EAFIT*, 44(150):51–63, March 2011.
- [7] Hillier, F. y Lieberman, J. *Introducción a la investigación de operaciones*, chapter Teoría de colas, págs. 709–771. McGrawHill, novena edición, 2010.
- [8] Kendall, D. G. Some problems in the theory of queues. *Journal of the Royal Statistical Society. Series B (Methodological)*, 13(2):151–185, 1951.
- [9] Kimura, T. Approximations for the mean delay in the m/d/s queue. *Proc. of the Seminar on Queueing Theory and Its Applications*, págs. 173–184, 1987.

- [10] Kimura, T. Refining Cosmetatos' approximation for the mean waiting time in the m/d/s queue. *Journal of the Operational Research Society*, 42(7):595–603, July 1991.
- [11] Lee, H. K., Musa, A. J., Bain, P. A., Nelson, K., Baker, C., y Li, J. A queueing network model for analysis of patient transitions within hospitals. *IEEE Transactions on Automation Science and Engineering*, 16(1):6–20, January 2019.
- [12] Magnus, U., Joseph, O., y Anthony, O. The application of queuing theory in the effective management of time in money deposit banks-a study of Zenith bank PLC in Enugu Metropolis. *The International Journal Research Publication's, RJSSM*, 5(8):19–32, December 2015.
- [13] Portilla, L., Arias, L., y Fernandez, S. Análisis de líneas de espera a través de teoría de colas y simulación. *Scientia Et Technica*, 17(46):56–61, 2010.
- [14] Rincón, L. *Introducción a los procesos estocásticos*, chapter Cadenas de Markov a tiempo continuo, págs. 151–167. Facultad de Ciencias UNAM, primera edición, 2012.
- [15] Vallejos, Y., Alfonzo, P. L., y Mariño, S. Teoría de colas. propuesta de un simulador didáctico. *Revista Publicando*, 4(13):5–20, December 2017.

Anexos

5.3. Resultados y métricas

s	ρ	A(s)	EW(M/M/s)	EW(M/D/s)	EW(M/G/s)	s	W	Lq	L
15	0,933	-1,143	6,548	3,368	5,062	15	14,355	8	22
16	0,875	-1,133	2,324	1,232	1,921	16	11,214	3	17
17	0,823	-1,125	1,064	0,582	0,945	17	10,237	1	15
18	0,778	-1,118	0,531	0,300	0,510	18	9,802	1	15
19	0,737	-1,111	0,274	0,160	0,286	19	9,579	0	14
20	0,700	-1,105	0,143	0,086	0,163	20	9,455	0	14
21	0,666	-1,100	0,074	0,046	0,093	21	9,385	0	14
22	0,636	-1,095	0,038	0,025	0,053	22	9,345	0	14
23	0,609	-1,091	0,019	0,013	0,030	23	9,322	0	14
24	0,583	-1,087	0,009	0,007	0,016	24	9,309	0	14
25	0,560	-1,083	0,005	0,003	0,009	25	9,301	0	14

(a) Resultados 9:00 horas

(b) Métricas 9:00 horas

Figura 5.1 Hora 9:00

Autor: Salazar W. Fuente: Salazar W.

s	ρ	A(s)	EW(M/M/s)	EW(M/D/s)	EW(M/G/s)	s	W	Lq	L
18	0,947	-1,118	6,817	3,495	5,306	18	14,303	10	27
19	0,897	-1,111	2,464	1,299	2,038	19	11,034	4	21
20	0,852	-1,105	1,166	0,632	1,029	20	10,025	2	19
21	0,812	-1,100	0,607	0,339	0,573	21	9,570	1	18
22	0,775	-1,095	0,329	0,189	0,334	22	9,330	1	18
23	0,741	-1,091	0,181	0,107	0,198	23	9,195	0	17
24	0,710	-1,087	0,099	0,061	0,118	24	9,115	0	17
25	0,682	-1,083	0,054	0,034	0,071	25	9,067	0	17

(a) Resultados 10:00 horas

(b) Métricas 10:00 horas

Figura 5.3 Hora 10:00

Autor: Salazar W. Fuente: Salazar W.

s	ρ	A(s)	EW(M/M/s)	EW(M/D/s)	EW(M/G/s)
19	0,966	-1,111	11,445	5,817	8,742
20	0,918	-1,105	3,337	1,741	2,698
21	0,874	-1,100	1,509	0,809	1,297
22	0,834	-1,095	0,779	0,429	0,714
23	0,798	-1,091	0,424	0,241	0,416
24	0,765	-1,087	0,237	0,138	0,250
25	0,734	-1,083	0,133	0,080	0,151

(a) Resultados 11:00 horas

s	W	Lq	L
19	17,626	18	36
20	11,583	6	24
21	10,181	3	21
22	9,598	1	20
23	9,301	1	19
24	9,134	1	19
25	9,035	0	19

(b) Métricas 11:00 horas

Figura 5.5 Hora 11:00

Autor: Salazar W. **Fuente:** Salazar W.

s	ρ	A(s)	EW(M/M/s)	EW(M/D/s)	EW(M/G/s)
16	0,946	-1,133	7,715	3,951	6,133
17	0,890	-1,125	2,547	1,342	2,158
18	0,841	-1,118	1,156	0,628	1,049
19	0,796	-1,111	0,582	0,326	0,568
20	0,757	-1,105	0,306	0,177	0,323
21	0,721	-1,100	0,163	0,097	0,187
22	0,688	-1,095	0,087	0,053	0,109
23	0,658	-1,091	0,046	0,029	0,063
24	0,630	-1,087	0,024	0,016	0,036
25	0,605	-1,083	0,012	0,008	0,021

(a) Resultados 12:00 horas

s	W	Lq	L
16	14,736	11	26
17	10,761	4	19
18	9,652	2	17
19	9,172	1	16
20	8,926	1	16
21	8,790	0	15
22	8,712	0	15
23	8,666	0	15
24	8,640	0	15
25	8,624	0	15

(b) Métricas 12:00 horas

Figura 5.7 Hora 12:00

Autor: Salazar W. **Fuente:** Salazar W.

s	ρ	A(s)	EW(M/M/s)	EW(M/D/s)	EW(M/G/s)	s	W	Lq	L
14	0,914	-1,154	4,748	2,460	3,911	14	12,541	6	19
15	0,853	-1,143	1,785	0,954	1,579	15	10,209	2	15
16	0,799	-1,133	0,820	0,453	0,784	16	9,414	1	14
17	0,752	-1,125	0,404	0,231	0,420	17	9,051	1	13
18	0,711	-1,118	0,204	0,121	0,233	18	8,863	0	13
19	0,673	-1,111	0,104	0,064	0,130	19	8,760	0	13
20	0,639	-1,105	0,052	0,033	0,073	20	8,703	0	13
21	0,609	-1,100	0,026	0,017	0,040	21	8,670	0	13
22	0,581	-1,095	0,013	0,009	0,022	22	8,652	0	13
23	0,556	-1,091	0,006	0,004	0,012	23	8,642	0	13
24	0,533	-1,087	0,003	0,002	0,006	24	8,636	0	13
25	0,512	-1,083	0,001	0,001	0,003	25	8,633	0	13

(a) Resultados 13:00 horas

(b) Métricas 13:00 horas

Figura 5.9 Hora 13:00

Autor: Salazar W. Fuente: Salazar W.

s	ρ	A(s)	EW(M/M/s)	EW(M/D/s)	EW(M/G/s)	s	W	Lq	L
14	0,920	-1,154	5,562	2,874	4,439	14	13,527	6	19
15	0,859	-1,143	2,028	1,081	1,737	15	10,825	2	15
16	0,805	-1,133	0,924	0,509	0,854	16	9,942	1	14
17	0,758	-1,125	0,454	0,259	0,457	17	9,544	1	14
18	0,716	-1,118	0,230	0,136	0,253	18	9,340	0	13
19	0,678	-1,111	0,117	0,072	0,141	19	9,229	0	13
20	0,644	-1,105	0,059	0,037	0,079	20	9,167	0	13
21	0,613	-1,100	0,029	0,019	0,044	21	9,131	0	13
22	0,586	-1,095	0,014	0,010	0,024	22	9,112	0	13
23	0,560	-1,091	0,007	0,005	0,013	23	9,100	0	13
24	0,537	-1,087	0,003	0,002	0,007	24	9,094	0	13
25	0,515	-1,083	0,001	0,001	0,003	25	9,091	0	13

(a) Resultados 14:00 horas

(b) Métricas 14:00 horas

Figura 5.11 Hora 14:00

Autor: Salazar W. Fuente: Salazar W.

s	ρ	A(s)	EW(M/M/s)	EW(M/D/s)	EW(M/G/s)	s	W	Lq	L
16	0,935	-1,133	6,609	3,400	5,201	16	14,399	8	23
17	0,880	-1,125	2,368	1,254	1,988	17	11,187	3	18
18	0,831	-1,118	1,098	0,600	0,989	18	10,187	2	17
19	0,788	-1,111	0,557	0,314	0,540	19	9,739	1	16
20	0,748	-1,105	0,293	0,170	0,308	20	9,506	1	15
21	0,713	-1,100	0,156	0,094	0,178	21	9,376	0	15
22	0,680	-1,095	0,082	0,051	0,103	22	9,302	0	15
23	0,651	-1,091	0,043	0,028	0,060	23	9,258	0	15
24	0,624	-1,087	0,022	0,015	0,034	24	9,233	0	15
25	0,599	-1,083	0,011	0,008	0,019	25	9,218	0	15

(a) Resultados 15:00 horas

(b) Métricas 15:00 horas

Figura 5.13 Hora 15:00

Autor: Salazar W. Fuente: Salazar W.

s	ρ	A(s)	EW(M/M/s)	EW(M/D/s)	EW(M/G/s)	s	W	Lq	L
13	0,953	-1,167	12,825	6,527	9,981	13	18,940	14	26
14	0,885	-1,154	3,264	1,713	2,721	14	11,681	4	16
15	0,826	-1,143	1,351	0,732	1,214	15	10,174	2	14
16	0,775	-1,133	0,638	0,358	0,622	16	9,582	1	13
17	0,729	-1,125	0,316	0,184	0,337	17	9,297	0	13
18	0,689	-1,118	0,159	0,096	0,186	18	9,146	0	13
19	0,652	-1,111	0,080	0,050	0,103	19	9,063	0	13
20	0,620	-1,105	0,039	0,026	0,057	20	9,017	0	13
21	0,590	-1,100	0,019	0,013	0,031	21	8,991	0	12
22	0,563	-1,095	0,009	0,006	0,017	22	8,977	0	12
23	0,539	-1,091	0,004	0,003	0,009	23	8,969	0	12
24	0,516	-1,087	0,002	0,001	0,004	24	8,964	0	12
25	0,496	-1,083	0,001	0,001	0,002	25	8,962	0	12

(a) Resultados 16:00 horas

(b) Métricas 16:00 horas

Figura 5.15 Hora 16:00

Autor: Salazar W. Fuente: Salazar W.

5.4. Simulaciones

Servidores	50%	1%	99%	Nsim
15	4,749	2,961	9,673	300
16	1,714	1,216	2,591	300
17	0,821	0,556	1,131	300
18	0,409	0,273	0,618	300
19	0,217	0,140	0,317	300
20	0,115	0,067	0,184	300
21	0,060	0,035	0,104	300
22	0,032	0,015	0,057	300

Autor: Salazar W. Fuente: Salazar W.

Figura 5.17 Simulación 9:00 horas

Servidores	50%	1%	99%	Nsim
18	5,063	2,926	11,216	300
19	1,857	1,257	3,041	300
20	0,897	0,598	1,298	300
21	0,477	0,321	0,636	300
22	0,260	0,171	0,392	300

Autor: Salazar W. Fuente: Salazar W.

Figura 5.18 Simulación 10:00 horas

Servidores	50%	1%	99%	Nsim
19	7,779	4,214	22,965	200
20	2,456	1,552	4,075	200
21	1,149	0,810	1,607	200
22	0,608	0,425	0,875	200

Autor: Salazar W. Fuente: Salazar W.

Figura 5.19 Simulación 11:00 horas

Servidores	50%	1%	99%	Nsim
16	5,743	3,451	12,220	200
17	2,028	1,385	3,079	200
18	0,923	0,627	1,417	200
19	0,461	0,326	0,628	200
20	0,248	0,172	0,338	200
21	0,134	0,087	0,203	200
22	0,072	0,045	0,114	200

Autor: Salazar W. Fuente: Salazar W.

Figura 5.20 Simulación 12:00 horas

Servidores	50%	1%	99%	Nsim
14	3,703	2,341	6,404	250
15	1,394	0,976	1,983	250
16	0,652	0,425	0,925	250
17	0,331	0,209	0,475	250
18	0,168	0,109	0,256	250
19	0,088	0,056	0,129	250
20	0,045	0,025	0,077	250
21	0,022	0,010	0,045	250
22	0,011	0,004	0,022	250

Autor: Salazar W. Fuente: Salazar W.

Figura 5.21 Simulación 13:00 horas

Servidores	50%	1%	99%	Nsim
14	4,186	2,560	7,780	300
15	1,581	1,142	2,381	300
16	0,715	0,517	1,075	300
17	0,360	0,249	0,520	300
18	0,187	0,120	0,284	300
19	0,096	0,060	0,147	300
20	0,048	0,029	0,085	300
21	0,024	0,011	0,044	300
22	0,012	0,004	0,025	300

Autor: Salazar W. Fuente: Salazar W.

Figura 5.22 Simulación 14:00 horas

Servidores	50%	1%	99%	Nsim
16	4,822	2,890	9,292	300
17	1,778	1,241	2,819	300
18	0,826	0,598	1,292	300
19	0,439	0,286	0,635	300
20	0,232	0,156	0,351	300
21	0,123	0,076	0,188	300
22	0,068	0,037	0,116	300

Autor: Salazar W. Fuente: Salazar W.

Figura 5.23 Simulación 15:00 horas

Servidores	50%	1%	99%	Nsim
13	9,149	5,417	27,013	250
14	2,560	1,695	4,540	250
15	1,060	0,714	1,613	250
16	0,497	0,354	0,773	250
17	0,258	0,177	0,389	250
18	0,127	0,079	0,201	250
19	0,064	0,037	0,107	250
20	0,033	0,018	0,058	250
21	0,016	0,006	0,034	250
22	0,007	0,002	0,019	250

Autor: Salazar W. Fuente: Salazar W.

Figura 5.24 Simulación 16:00 horas