

---

## TP 3 - SIMULACIÓN DE UN SISTEMA MM1.

---

**Alexis Inga**

Cátedra Simulación

*Comisión 401*

Universidad Tecnológica Nacional - FRRO

Zeballos 1341, S2000

alexisinga97@gmail.com

**Eliseo Gamba**

Cátedra Simulación

*Comisión 401*

Universidad Tecnológica Nacional - FRRO

Zeballos 1341, S2000

eliseogamba00@gmail.com

**Navarini Alejandro**

Cátedra Simulación

*Comisión 401*

Universidad Tecnológica Nacional - FRRO

Zeballos 1341, S2000

ale.navarini99@gmail.com

11 de julio de 2020

### ABSTRACT

En este documento procederemos a simular un sistema de colas MM1 y al posterior análisis de sus medidas de rendimiento en diversos casos.

**Keywords** simulación, teoría de colas, mm1, medidas de rendimiento

### 1. Introducción

Todos hemos sido parte de una cola de espera en nuestras vidas, para ser atendidos dentro de cualquier tipo de sistema en la demanda de un servicio. La Teoría de Colas surgió a principios del siglo XX, cuando el problema de la congestión de tráfico que se presentaban en las redes telefónicas fue estudiado por primera vez desde un enfoque científico por el danés Agner Kraup Erlang. A partir de ese entonces esta teoría se ha aplicado a multitud de problemas de la vida real. Proporciona tanto una base teórica del tipo de servicio que se puede esperar de un determinado recurso, como la forma en la cual puede ser diseñado. La teoría de colas es una colección de modelos matemáticos que describen sistemas de línea de espera. Dichos modelos sirven para encontrar un balance entre el costo del servicio y el costo asociado a la espera por ese servicio. El objetivo de nuestro trabajo es realizar la simulación de un sistema de colas MM1(según la notación de Kendal), en el mismo vamos a contar con una tasa entre arribos y una tasa de servicios Markovianas (M), que siguen una distribución de Poisson y con un solo servidor.

## 2. Descripción

Nuestro trabajo consiste en mostrar el funcionamiento de una cola MM1 la cual consiste en:

- 1- Una población infinita de clientes
- 2- Una tasa entre arribos que sigue un proceso de Poisson con una tasa promedio de  $\lambda$  clientes por unidad de tiempo
- 3- Una tasa de servicio que sigue un proceso de Poisson con una tasa promedio de  $\mu$  clientes por unidad de tiempo
- 4- Un proceso de colas que consiste en una sola cola de espera con capacidad infinita con una disciplina FIFO y más adelante en el trabajo se procederá a analizar la simulación con una capacidad finita de cola.

De nuestro modelo de colas MM1 realizaremos una simulación con los siguientes parámetros: tasa de servicio 1, con la tasa entre arribos variando entre el 25 %, 50 %, 75 %, 100 % y 125 % de la tasa de servicio, la simulación se realizará hasta que hayan pasado 10.000 clientes por el sistema y realizaremos 10 corridas con cada parámetro de los nombrados. La misma esta implementada tanto en código Python como en Anylogic, y de esta procederemos a analizar sus medidas de rendimiento.

## 3. Marco teórico

Una cola simple con un solo servidor a consiste en un servidor que realiza los requerimientos de los clientes, una línea de espera o cola donde los clientes esperan recibir servicios y una fuente que genera las llegadas de los clientes al sistema. Por lo general cuentan con una disciplina FIFO(First In First Out). D.G. Kendall sugirió de una notación de utilidad para clasificar la amplia diversidad de los diferentes modelos de línea de espera que se han desarrollado. La notación de Kendall, de tres símbolos es como sigue: A/B/K Donde: A: indica la distribución de probabilidades de las llegadas B: Indica la distribución de probabilidades de tiempos de servicio • K: Indica el número de canales. Para nuestro caso, para un sistema de cola simple, la notacion de Kendall sera M/M/1, las M denotan una distribucion de probabilidad de Poisson tanto para las llegadas como para el servicio, y el 1 indica que contamos con un solo servidor.

El tiempo en el sistema puede variar con el número de clientes en la cola

Sistema de la cola = es el conjunto formado por la cola y el mecanismo de servicio, junto con la disciplina de la cola, que es lo que nos indica el criterio de qué cliente de la cola elegir para pasar al mecanismo de servicio

- $\lambda$ = Número de entidades promedio que ingresan al sistema en un lapso de tiempo.
- $\mu$ = Número de entidades promedio que pueden ser atendidas por el servidor en un lapso de tiempo.

### 3.1. Medidas de Rendimiento

- Promedio de clientes en el sistema: promedio de los n clientes en el sistema que están en cola y que son atendidos.

$$L = \frac{\rho}{1 - \rho} \quad (1)$$

- Promedio de clientes en cola: promedio de la cantidad de clientes en cola a la espera de ser atendidos.

$$L_q = \frac{\rho^2}{1 - \rho} \quad (2)$$

- Tiempo promedio en sistema: tiempo promedio Desde que n clientes llegan y salen del sistema.

$$W = \frac{L}{\lambda} \quad (3)$$

- Tiempo promedio en cola: tiempo promedio que tuvieron los  $n$  clientes que pasaron por el sistema esperando en cola.

$$W_q = \frac{L_q}{\lambda} \quad (4)$$

- Utilización del servidor: es el servicio para todo el sistema (número esperado clientes que completan su servicio por unidad de tiempo) cuando hay  $n$  clientes en el sistema.

$$\rho = \frac{\lambda}{\mu} \quad (5)$$

- Probabilidad de  $n$  clientes en cola

$$P(N = n) = (1 - \rho)\rho^n \quad n \geq 0 \quad (6)$$

#### 4. Desarrollo

- Para los parámetros  $\lambda = 0.25$  y  $\mu = 1$ , con 10.000 clientes pasando por el sistema y 10 corridas obtenemos

- Promedio de clientes en el sistema
  - Teórico: 0.333
  - Anylogic: 0.351
  - Python: 0.33
- Promedio de clientes en cola
  - Teórico: 0.083
  - Anylogic : 0.093
  - Python: 0.08
- Tiempo promedio en sistema.
  - Teórico: 1.333s
  - Anylogic: 1.42s
  - Python: 1.33s
- Tiempo promedio en cola
  - Teórico: 0.333s
  - Anylogic: 0.38s
  - Python: 0.33s
- Utilización del servidor
  - Teórico: 0.25
  - Anylogic: 0.26
  - Python: 0.25
- Probabilidad de  $n$  clientes en cola (tomando  $n$  como 1)
  - Teórico: 0.0468
  - Python: 0.05

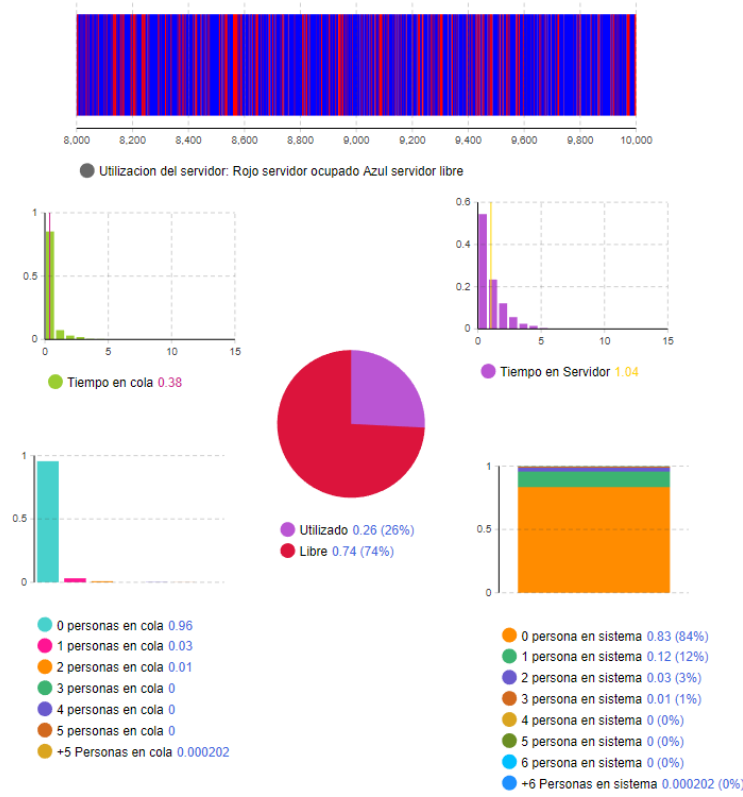


Figura 1: Gráficas de Anylogic obtenidas para  $\lambda = 0.25$  y  $\mu = 1$

Utilización del servidor: Rojo servidor ocupado, Azul servidor libre: En esta gráfica se refleja la utilización del servidor a lo largo de la simulación, de esta gráfica es interesante analizar los periodos en los cuales el servidor tienen tiempos ociosos y horas picos, para estos parámetros lo que podemos observar es que el servidor se encuentra libre la mayor parte del tiempo y además no vemos horas pico ni tiempos ociosos muy extensos, se observa que la utilización del servidor es esporádica.

Tiempo en cola, y tiempo en servidor. Estas gráficas reflejan el tiempo que pasan los clientes en cola y en el servidor, siendo la acumulación de ellas el tiempo que pasarían en el sistema, podemos observar que ambas muestran un comportamiento de una distribución exponencial, donde la media de el tiempo en cola es aproximadamente un tercio de la media del tiempo del servidor, lo que podemos deducir de estos valores es que las personas no tienen la necesidad de hacer periodos extensos de cola y que la mayor parte del tiempo que pasan dentro del sistema están siendo atendidos en el servidor.

En el histograma de personas en cola podemos observar el porcentaje de el tiempo que el sistema se encuentra con esa cantidad de personas esperando ser atendidos. Con respecto a la gráfica de personas en cola podemos observar que el 98 % de el tiempo de la simulación no hay nadie en la cola esto quiere decir que nuestra tasa de servicio es suficiente para satisfacer la demanda con una respuesta rápida.

En nuestro gráfico de personas en el sistema podemos apreciar el porcentaje del tiempo que el sistema cuenta con esa cantidad de personas ya sea esperando en la cola como siendo atendido en el servidor. La conclusión a la que llegamos de ver este gráfico es que por lo general el sistema se encuentra desocupado y ante un arribo la respuesta es eficiente logrando que no se acumule gente en cola.

En el gráfico de torta se representa la utilización del servidor pero en este caso nos permite ver cuanto fue utilizado el servidor y cuanto estuvo libre durante la totalidad del servidor. Y se ve tal como venimos comentando en los puntos anteriores que en su generalidad el servidor se encuentra libre.

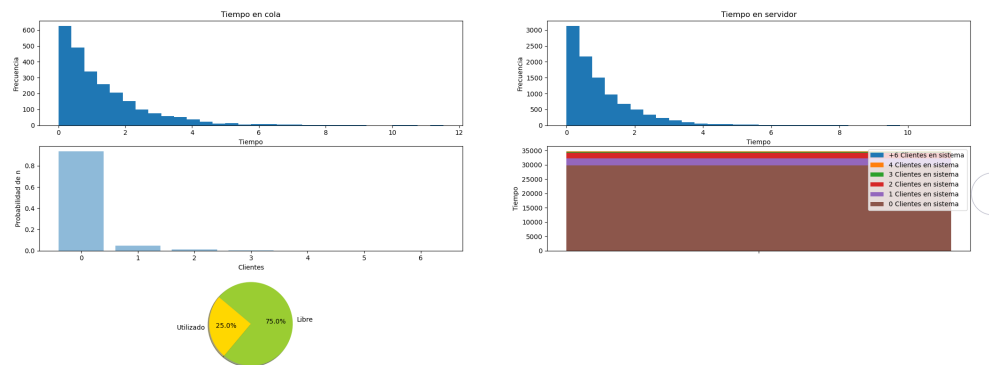


Figura 2: Gráficas de Python obtenidas para  $\lambda = 0.25$  y  $\mu = 1$

Las gráficas realizadas con Python tienen un comportamiento similar a las realizadas con Anylogic, por lo cual no es necesario hacer comentarios adicionales con respecto a lo dicho sobre las gráficas anteriores.

- Para los parámetros  $\lambda = 0.50$  y  $\mu = 1$ , con 10.000 clientes pasando por el sistema y 10 corridas obtenemos
  - Promedio de clientes en el sistema
    - Teórico: 1
    - Anylogic: 1
    - Python: 1.01
  - Promedio de clientes en cola
    - Teórico: 0.5
    - Anylogic: 0.5
    - Python: 0.51
  - Tiempo promedio en sistema.
    - Teórico: 2
    - Anylogic: 1.99s
    - Python: 2.03s
  - Tiempo promedio en cola
    - Teórico: 1
    - Anylogic: 0.99s
    - Python: 1.02s
  - Utilización del servidor
    - Teórico: 0.5
    - Anylogic: 0.5
    - Python: 0.5
  - Probabilidad de n clientes en cola (tomando n como 1)
    - Teórico: 0.125
    - Python: 0.12

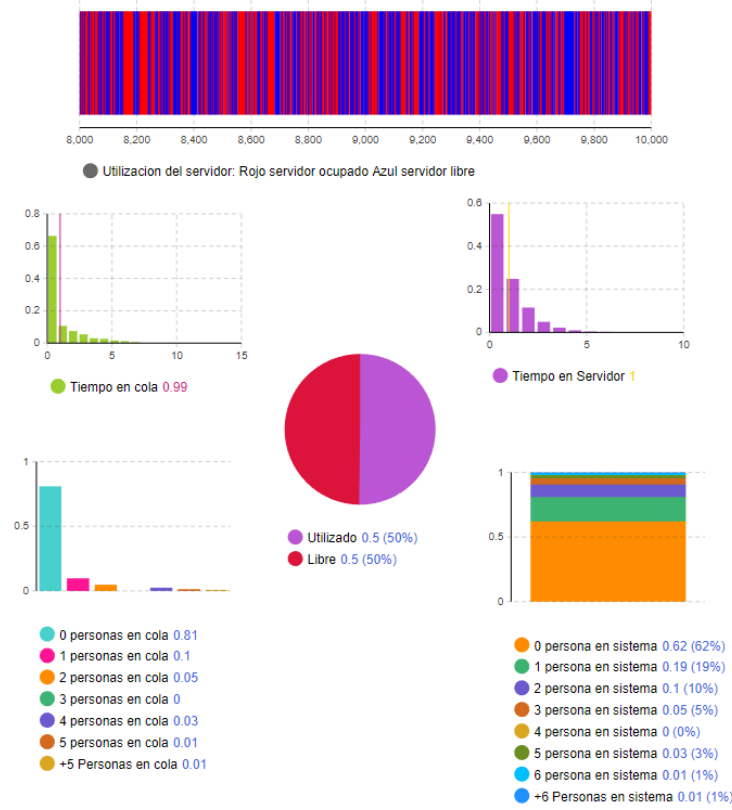


Figura 3: Gráficas de Anylogic obtenidas para  $\lambda = 0.50$  y  $\mu = 1$

Utilización del servidor: Rojo servidor ocupado, Azul servidor libre: en contraste con la simulación previa en esta gráfica podemos observar que existen mas periodos donde el servidor esta siendo utilizado, lo cual tiene sentido ya que hubo un aumento en la tasa entre arribos y la tasa de servicio permaneció igual.

En las gráficas de tiempo podemos observamos el mismo comportamiento que en la simulación previa, pero la media de el tiempo que los clientes pasan en cola aumento y el tiempo que pasan en el servidor se mantuvo igual, el cliente va a pasar aproximadamente el mismo tiempo haciendo la cola que siendo atendido.

Con respecto a la gráfica de personas en cola podemos observar que disminuyo el porcentaje de tiempo donde no había nadie en la cola en un 17 % y se puede apreciar que en momentos de la simulación la cola contó con una mayor cantidad de clientes en cola con respecto a la gráfica anterior.

En comparativa con la gráfica anterior al aumentar la tasa entre arribos se incrementa el tiempo en el cual el sistema se encuentra sin personas.

En el gráfico podemos ver reflejado el cambio de la tasa entre arribos manteniendo la tasa de servicio, ya que hay un cambio directamente proporcional a la variación de la tasa entre arribos.

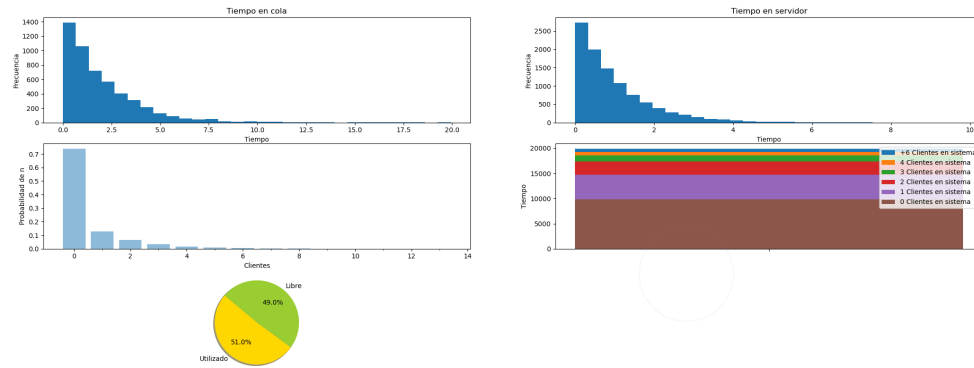


Figura 4: Gráficas de Python obtenidas para  $\lambda = 0.50$  y  $\mu = 1$

Las gráficas realizadas con Python tienen un comportamiento similar a las realizadas con Anylogic, por lo cual no es necesario hacer comentarios adicionales con respecto a lo dicho sobre las gráficas anteriores.

- Para los parámetros  $\lambda = 0.75$  y  $\mu = 1$ , con 10.000 clientes pasando por el sistema y 10 corridas obtenemos
  - Promedio de clientes en el sistema
    - Teórico: 3
    - Anylogic : 3.078
    - Python: 2.93
  - Promedio de clientes en cola
    - Teórico: 2.25
    - Anylogic : 2.318
    - Python: 2.18
  - Tiempo promedio en sistema.
    - Teórico: 4
    - Anylogic: 4.04s
    - Python: 3.91s
  - Tiempo promedio en cola
    - Teórico: 3
    - Anylogic: 3.04s
    - Python: 2.92s
  - Utilización del servidor
    - Teórico: 0.75
    - Anylogic: 0.76
    - Python: 0.74
  - Probabilidad de n clientes en cola (tomando n como 1)
    - Teórico: 0.1406
    - Python: 0.14

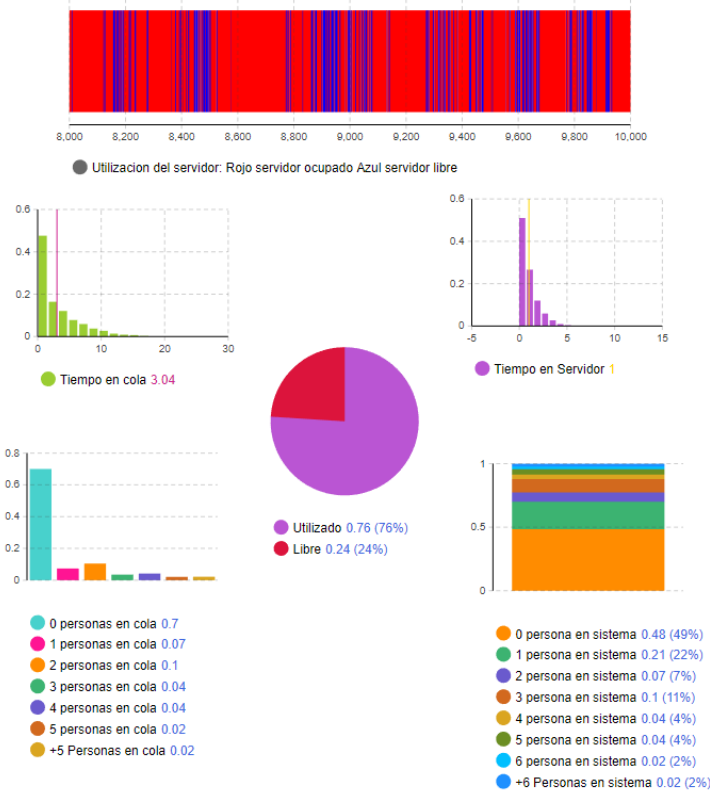


Figura 5: Gráficas de Anylogic obtenidas para  $\lambda = 0.75$  y  $\mu = 1$

En la gráfica de utilización podemos apreciar claramente la prolongación de los periodos donde el servidor se encuentra ocupado, estos periodos son picos y pueden ser de mucha utilidad a la hora de intentar de predecir el comportamiento de un futuro negocio, en base a datos como estos se podría plantear el incremento de la cantidad de servidores en los periodos picos para proveer un servicio mas eficiente.

La diferencia que observamos con respecto a las previas gráficas de los tiempos que los clientes pasan tanto en cola como en servidor, es que el tiempo en servidor se mantiene mientras que el tiempo en cola se triplico con respecto al anterior, el tiempo en cola en este caso representa un 75.24 % de el tiempo que el cliente pasa en el servidor.

Con respecto a las gráficas sobre personas en el sistema podemos observar que hubo un decremento con el porcentaje del tiempo que el sistema cuenta con 0 personas así como un incremento en los porcentajes donde el sistema cuenta con muchas personas a la vez.



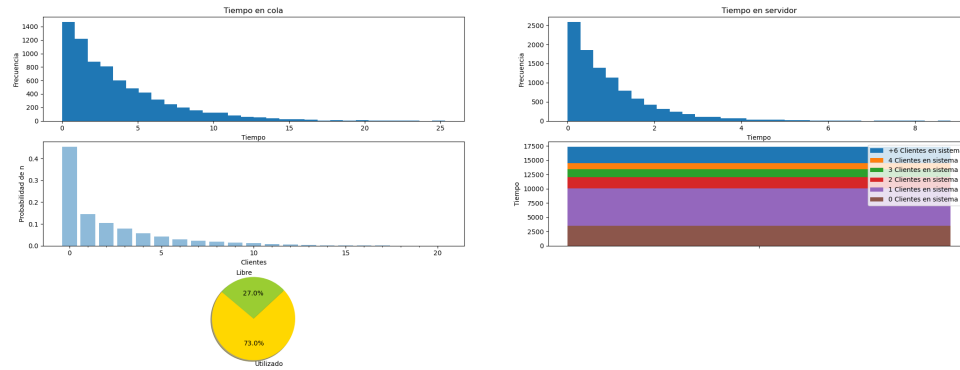


Figura 6: Gráficas de Python obtenidas para  $\lambda = 0.75$  y  $\mu = 1$

Las gráficas realizadas con Python tienen un comportamiento similar a las realizadas con Anylogic, por lo cual no es necesario hacer comentarios adicionales con respecto a lo dicho sobre las gráficas anteriores.

- Para los parámetros  $\lambda = 1$  y  $\mu = 1$ , con 10.000 clientes pasando por el sistema y 10 corridas obtenemos
  - Promedio de clientes en el sistema
    - Teórico: No se puede calcular
    - Anylogic : 54.05
    - Python: 75.17
  - Promedio de clientes en cola
    - Teórico: No se puede calcular
    - Anylogic : 53.072
    - Python: 74.19
  - Tiempo promedio en sistema.
    - Teórico: No se puede calcular
    - Anylogic: 54.45s
    - Python: 75.35s
  - Tiempo promedio en cola
    - Teórico: No se puede calcular
    - Anylogic: 53.46s
    - Python: 74.35s
  - Utilización del servidor
    - Teórico: 1
    - Anylogic: 0.98
    - Python: 1
  - Probabilidad de n clientes en cola (tomando n como 1)
    - Teórico: 0
    - Python: 0.01

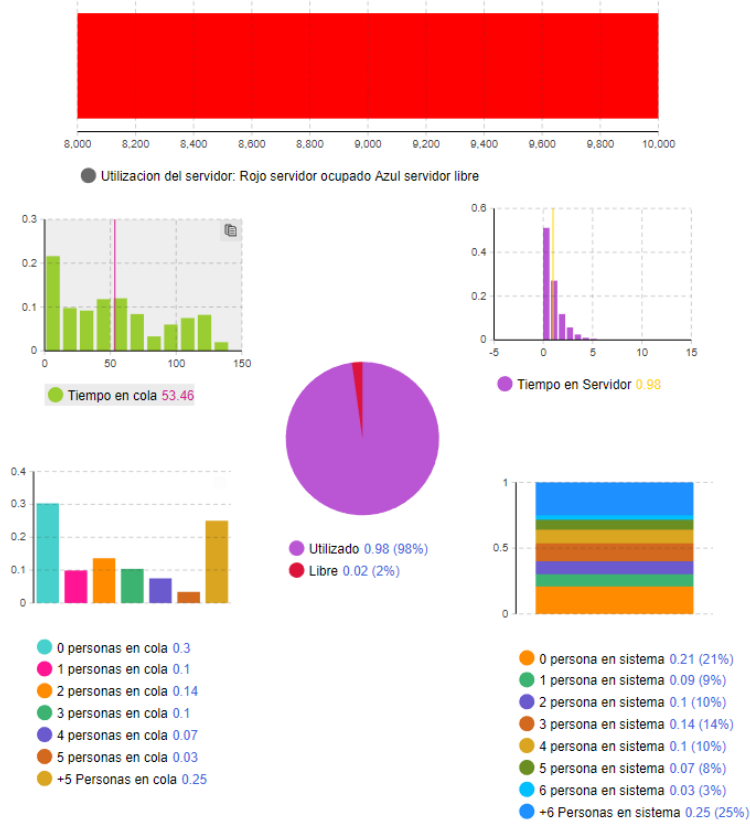


Figura 7: Gráficas de Anylogic obtenidas para  $\lambda = 1$  y  $\mu = 1$

Con respecto a la gráfica de utilización de el servidor podemos ver que el servidor se encuentra utilizado a pleno, mostrando un periodo pico durante casi toda la simulación, en un escenario real esto nos llevaría a considerar que seria necesario aumentar la cantidad de servidores para eficientizar el negocio.

En las gráficas de tiempos podemos notar el mayor incremento hasta ahora de el tiempo en la cola, un cliente pasa el 98.16 % de el tiempo que esta en el sistema esperando a ser atendido.

En las gráficas sobre cantidad de clientes pudimos observar un incremento muy notorio en la cantidad de personas que se encuentran en la cola y en el sistema, así como un aumento en la frecuencia donde el sistema cuenta con una gran cantidad de clientes a la vez.

En la gráfica de torta es evidente que la utilización de el servidor aumento hasta el punto de que el servidor paso casi toda la simulación atendiendo personas.

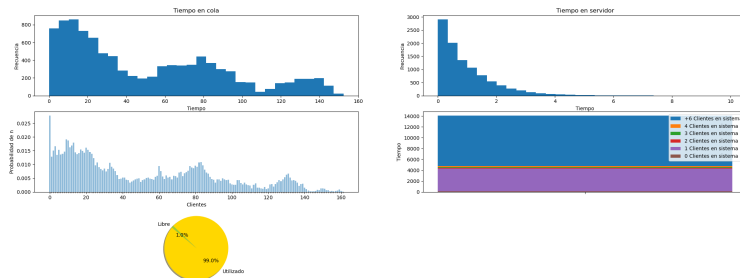


Figura 8: Gráficas de Anylogic obtenidas para  $\lambda = 1$  y  $\mu = 1$

En la simulación con estos parámetros pudimos observar que el comportamiento de la gráfica de el tiempo que pasa una persona en la cola comienza a diferir de el comportamiento de una gráfica exponencial. Otro de los puntos que pudimos observar es que existe cierta correlación entre las gráficas de el tiempo que pasa una persona en la cola y la probabilidad de encontrar  $n$  clientes en la cola, este es un comportamiento que se presenta para otros parámetros pero creemos que para esta simulación es mas fácil de apreciarlo.

- Para los parámetros  $\lambda = 1.25$  y  $\mu = 1$ , con 10.000 clientes pasando por el sistema y 10 corridas obtenemos

- Promedio de clientes en el sistema
  - Teórico: No se puede calcular
  - Anylogic : 1126.013
  - Python: 1263.85
- Promedio de clientes en cola
  - Teórico: No se puede calcular
  - Anylogic : 1125.013
  - Python: 1262.85
- Tiempo promedio en sistema.
  - Teórico: No se puede calcular
  - Anylogic: 909.28s
  - Python: 1008.42
- Tiempo promedio en cola
  - Teórico: No se puede calcular
  - Anylogic: 908.38s
  - Python: 1007.42
- Utilización del servidor
  - Teórico: No se puede calcular
  - Anylogic: 1
  - Python: 1
- Probabilidad de  $n$  clientes en cola(tomando  $n$  como 1)
  - Teórico: 0
  - Python: 0.0

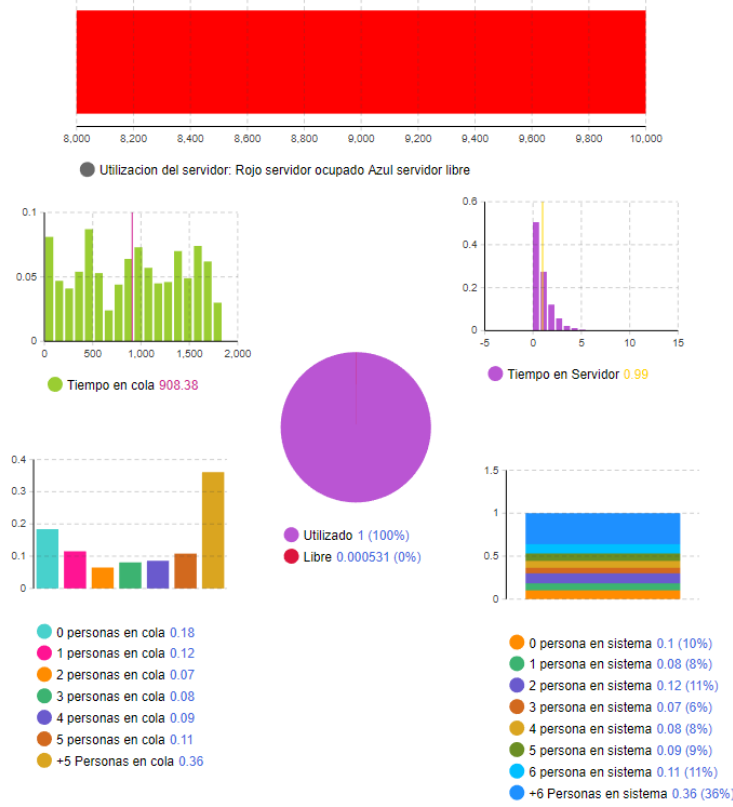


Figura 9: Gráficas de Anylogic obtenidas para  $\lambda = 1.25$  y  $\mu = 1$

En la gráfica de utilización del servidor no podemos notar una diferencia con respecto a la simulación previa debido a que el servidor ya se encontraba utilizado a lo largo de toda la simulación, este comportamiento se repite aquí.

En estas gráficas de el tiempo podemos observar que, la gráfica de el tiempo de cola deja de presentar un comportamiento de una distribución exponencial y que la media de el tiempo que una persona pasa en cola se incrementa sustancialmente, representando el 99.7 % de el tiempo que una persona pasa en el servidor. A nuestro criterio en esta simulación ya podemos ver que el sistema se encuentra saturado.

En las gráficas de cantidad de personas en sistema y en cola lo que podemos apreciar es que es mas frecuente que hayan mas de 5 personas tanto en la cola como en el sistema a que el sistema se encuentre libre, esto no lleva a repetir la misma reflexión recién comentada de que el sistema no puede saciar la demanda que se le presenta.

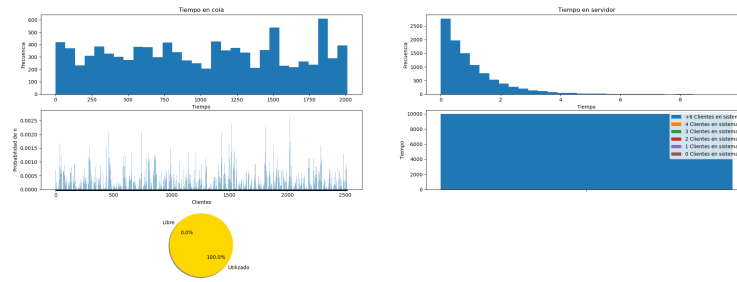


Figura 10: Gráficas de Anylogic obtenidas para  $\lambda = 1.25$  y  $\mu = 1$

Las gráficas realizadas con Python tienen un comportamiento similar a las realizadas con Anylogic, por lo cual no es necesario hacer comentarios adicionales con respecto a lo dicho sobre las gráficas anteriores.

#### 4.1. Probabilidad de denegación de servicio

Para poder simular la probabilidad de denegación de servicio fue necesario modificar nuestro modelo de cola a un modelo de cola finita. En un modelo de cola con capacidad  $n$  de clientes, lo que sucede es que cuando arriba el cliente  $n + 1$  se le deniega el servicio. Realizaremos simulaciones de 100 iteraciones con 10000 clientes que atravesaran el sistema en cada una de ellas para calcular el valor de la probabilidad de denegación de servicio tomando  $n$  como :0 ,2 ,5 ,10 , 50. Presentaremos los valores obtenidos en la siguiente tabla:

Tamaño de cola	0	2	5	10	50
$\lambda = 0,25\mu = 1$	1	1	0.3	0	0
$\lambda = 0,5\mu = 1$	1	1	1	0.36	0
$\lambda = 0,75\mu = 1$	1	1	1	1	0
$\lambda = 1\mu = 1$	1	1	1	1	1
$\lambda = 1,25\mu = 1$	1	1	1	1	1

Lo que podemos observar en la tabla es que a medida de que la tasa entre arribos se acerca a la tasa de servicio es mas probable que se deniegue el servicio al cliente.

## 5. Conclusión

Lo que pudimos concluir de nuestro estudio fue que a medida que la tasa entre arribos ( $\lambda$ ) se acerca a la tasa de servicio ( $\mu$ ) el sistema comienza a perder rendimiento con relación a la velocidad de atención que se le proporciona a un cliente, podemos observar que el tiempo que un cliente pasa en el sistema se ve representado mayormente por el tiempo que el mismo pasa esperando a recibir el servicio, también notamos que la cantidad de clientes en cola aumenta sustancialmente, estos valores a su vez tienden a infinito cuando  $\lambda$  es mayor que  $\mu$ , en otras palabras se satura el sistema. Otra observación que pudimos hacer al realizar el trabajo fue que gráficas como la de el estudio de la utilización de el servidor a lo largo de el tiempo aportan mucha información útil a la hora de realizar un estudio para mejorar o abrir un negocio.

## Referencias

- [1] SIMULATION MODELING AND ANALYSIS - Averill M. Law and W. David Kelton
- [2] THE ART OF PROCES - CENTERIN MODELING whit ANY LOGIC