# SIMULACIÓN DE UN MODELO M/M/1 E INVENTARIO

**Abud Santiago Elias**
Legajo 47015
sabudvicco@gmail.com

**Buchhamer Ariel**
Legajo 46217
arielbuchhamer1@outlook.com

**Castellano Marcelo**
Legajo 39028
marce.geek22@gmail.com

**Dolan Guillermo Patricio**
Legajo 46101
guillermo230899@gmail.com

**Navarro Franco**
Legajo 46387
franconavarro1889@gmail.com

3 de junio de 2022

**ABSTRACT**

En este trabajo estudiaremos dos modelos de simulación de eventos discretos, el primero es el comportamiento de líneas de espera. El cual es de gran ayuda para predecir el comportamiento de dichas lineas en situaciones del mundo real, desde la entrada y salida de autos de un estacionamiento hasta la utilización de una red distribuida de servidores a lo largo del mundo que alojan una página web para miles de usuarios. El siguiente será un modelo de inventario que nos permitirá saber los costos correspondientes al mantenimiento o de compras necesarias y será de gran ayuda para un inventario en la vida real

## 1. Introducción

En el siguiente trabajo simularemos mediante un programa desarrollado en lenguaje Python, una cola simple, en la que solamente habrá llegada de clientes, atendidos por un único servidor luego de haber realizado una espera determinada en la cola y posteriormente partirán. Adicionalmente se simulará una cola con una cantidad de servidores predeterminada. El modelo de inventario contará con el mismo formato y consistirá en un proceso por mes en el que se realiza una orden de compra a los proveedores, mientras se trata de alcanzar la demanda de los clientes a través de distintos indicadores de escacez y existencia. Finalmente se realizará una evaluación y se calcularán las estadísticas correspondientes. A su vez estudiaremos la eficacia de dichos programas al comparar ciertas medidas de rendimiento observadas con las teóricas computadas y a su vez con las obtenidas mediante una calculadora web, y de otra simulación desarrollada en el aplicativo Anylogic.

## 2. Marco teórico

### 2.1. Proceso Estocástico

En la teoría de la probabilidad, un proceso estocástico es un concepto matemático que sirve para representar magnitudes aleatorias que varían con el tiempo o para caracterizar una sucesión de variables aleatorias (estocásticas) que evolucionan en función de otra variable, generalmente el tiempo. Cada una de las variables aleatorias del proceso tiene su propia función de distribución de probabilidad y pueden o no estar correlacionadas entre sí.

Cada variable o conjunto de variables sometidas a influencias o efectos aleatorios constituye un proceso estocástico. Un proceso estocástico Xt puede entenderse como una familia uniparamétrica de variables aleatorias indexadas mediante el tiempo t. Los procesos estocásticos permiten tratar procesos dinámicos en los que hay cierta aleatoriedad.

## 3. Modelo M/M/1

Un sistema de espera M/M/1 es aquel que considera un servidor, con tiempos exponenciales de servicio y entre llegadas de clientes. La implicancia que los tiempos de servicio se distribuyan exponencial es que existe una preponderancia de tiempos de servicio menores al promedio combinados con algunos pocos tiempos extensos. Un ejemplo de ello es lo que sucede en las cajas de los bancos donde la mayoría de las transacciones requieren poco tiempo de proceso por parte del cajero, no obstante algunas transacciones más complejas consumen bastante tiempo. Por otra parte afirmar que los tiempos entre llegadas se distribuyen exponencial implica una preponderancia de tiempos entre llegadas menores que el promedio en combinación con algunos tiempos más extensos. Lo anterior tiene relación con la aleatoriedad del proceso de llegada de clientes que permite establecer la Propiedad de Falta de Memoria o Amnesia de la Distribución Exponencial y con los conceptos presentados en el artículo Qué son las Líneas de Espera (Teoría de Colas), donde queda en evidencia que la formación de las colas o filas esta asociada a la variabilidad del sistema.

En este contexto consideremos la siguiente notación, donde valores usuales para A y B son M (distribución exponencial) y G (distribución general).

En el estudio con colas simples, cuya notación de Kendalls son M/M/1, para la primera, donde se posee un sólo serivor, las distribuciones de arribo y de servicio proceden de manera Markoviana y su cola es infinita, y en el caso de la siguientes M/M/C dónde solamente se añadiran mas servidores. Para la distribución de arribos, se tiene un parámetro , la tasa de arribos de clientes, de él obtenemos la media de clientes entrantes por unidad de tiempo 1/. Por otro lado, la tasa de tiempo de serivicios µ y la media de tiempos de servicios 1/ µ. A partir de estos dos parámetros, obtenemos la tasa de utilización del servidor  = 1/sµ, la relación entre los clientes arribando y los clientes partiendo, siendo s el número de servidores, en nuestro caso s = 1 (luego se permitira variarlo). Como la llegada de clientes y su tiempo en ser servidos son variables aleatorias, se pueden considerar nuevas variables aleatorias, que son combinaciones lineales de las anteriores, por ejemplo:

Considere un sistema de colas de un solo servidor para el cual el intervalo entre llegadas son A1, A2, . . . veces con variables aleatorias independientes e idénticamente distribuidas.

A partir de una única ejecución de la simulación con retrasos de los clientes D1, D2, . . . , Dn, es obvio que el estimador de d(n) es:

$$\hat{d}(n) = \frac{\sum_{i=1}^{n} D_i}{n} \tag{1}$$

que es solo el promedio de los Di's que se observaron en la simulación.

Una de las medidas para nuestro modelo simple es el número promedio esperado de clientes en la cola (pero sin ser servida), denotada por q(n), donde la n es necesaria en la notación para indicar que este promedio se toma durante el período de tiempo necesario para observar los n retrasos que definen nuestra regla de parada. Este es un tipo diferente de "promedio.ª comparación del retraso promedio en cola, porque se toma el tiempo (continuo), en lugar de los clientes (siendo este discreto). Por lo tanto, necesitamos definir qué significa este número promedio de tiempo de clientes en cola. Para hacer esto, sea Q(t) el número de clientes en cola en el tiempo t, para cualquier número real t>=0 , y sea T(n) el tiempo requerido para observar nuestros n retrasos en la cola. Entonces para cualquier tiempo t entre 0 y T(n), Q(t) es un entero no negativo. Además, si dejamos que pi sea la proporción esperada (que estará entre 0 y 1) del tiempo que Q(t) es igual a i, entonces una definición razonable de q(n) sería:

$$q(n) = \sum_{i=0}^{\infty} i p_i \tag{2}$$

Por lo tanto, q(n) es un promedio ponderado de los posibles valores de i para la longitud de la cola Q(t), siendo las áreas la proporción esperada de tiempo que la cola pasa en cada uno de sus posibles longitudes. Para estimar q(n) a partir de una simulación, simplemente reemplazamos los pi con estimaciones de ellos y obtenemos:

$$\hat{q}(n) = \sum_{i=0}^{\infty} i \hat{p}_i \tag{3}$$

donde i es la proporción observada (en lugar de la esperada) del tiempo durante la simulación donde que había i clientes en la cola. Computacionalmente, sin embargo, es más fácil reescribir (n) usando algunas consideraciones geométricas. Si dejamos que Ti sea el total de tiempo durante la simulación donde la cola tiene una longitud i, entonces T(n) = T0 + T1 + T2 +  y i = Ti/T(n), por lo que podemos reescribir la ecuación 1.1 como:

$$\hat{q}(n) = \frac{\sum_{i=0}^{\infty} i T_i}{T(n)} \tag{4}$$

El numerador de la ecuación anterior, representa el área bajo la función Q(t):

$$\sum_{i=0}^{k-1} i T_i = \int_0^{T(n)} Q(t)dt \tag{5}$$

Reemplazándolo en la ecuación anterior, nos queda:

$$\hat{q}(n) = \frac{\int_0^{T(n)} Q(t)dt}{T(n)} \tag{6}$$

Esta integral puede ser calculada como la suma de rectángulos formados por la base tiempo de cierta cantidad de clientes en cola por la altura dicha cantidad de clientes en cola.

La proporción esperada de tiempo del servidor en estado ocupado u(n), deviene de la probabilidad de que el servidor no esté vacío, pN>0 = 1  p0. Para calcular su estimador u(n), primero se define la "función ocupada":

$$B(t) \begin{cases} 1 & si \quad Nt > 0 \\ 0 & si \quad Nt = 0 \end{cases} \tag{7}$$

Entonces u(n) es la porción del tiempo total en la que B(t) = 1. Al igual que a la medida anterior, podemos considerarlo como el área bajo B(t), así,

$$\hat{u}(n) = \frac{\int_0^{T(n)} B(t)}{T(n)} \tag{8}$$

En definitiva, u(n) es la sumatoria de áreas rectangulares, donde la base es el tiempo en el que el servidor está en un estado específico, y la altura es dicho estado, 0 o 1.

## 4. Modelo de Inventario

Los inventarios están presentes en todas las compañías que tratan con productos físicos, tales como fabricantes, distribuidores, comerciantes, etc. Las empresas necesitan inventarios de materias primas para la manufactura de productos y a su vez deben almacenar productos terminados en el almacén a la espera de ser vendidos. De manera similar, los distribuidores deben mantener inventarios de bienes que deberán estar disponibles cuando los consumidores los necesiten. Dado que los stocks representan una cantidad de dinero inmovilizada muy importante dentro de una empresa, la reducción de los costos de almacenamiento (evitando inventarios innecesariamente grandes) podría mejorar la competitividad de cualquier sistema productivo.

Cuando hay que analizar los inventarios con una demanda independiente los modelos de gestión de stocks que se utilizan son: el modelo de cantidad fija del pedido (EOQ) y el modelo de periodo de tiempo fijo (también llamado de revisión periódica, modelo P) [1-2]. En el modelo de cantidad fija de pedido se coloca un pedido cuando el inventario restante cae a un punto de pedido y se revisa el nivel de inventario continuamente. De esta manera, el modelo de cantidad fija de pedido es un sistema perpetuo que requiere que cada vez que se haga un retiro o una adición al inventario, los registros deban actualizarse para asegurar que el punto del nuevo pedido se ha alcanzado o no. En cambio en el modelo de periodo de tiempo fijo, el conteo tiene lugar solo durante el periodo de revisión.

Suponiendo una empresa que vende un solo producto le gustaría decidir cuántos artículos debería tener en inventario para cada uno de los siguientes n meses (n es un parámetro de entrada fijo). Los tiempos entre demandas son variables aleatorias exponenciales con una media de 0,1 mes. Los tamaños de las demandas, D, son variables aleatorias IID (independientes de cuando se presentan las demandas), con:

$$D = \begin{cases} 1 & c.p. \quad 1/6 \\ 2 & c.p. \quad 1/3 \\ 3 & c.p. \quad 1/3 \\ 4 & c.p. \quad 1/6 \end{cases} \tag{9}$$

donde c.p. significa Çon Probabilidad de".

Al comienzo de cada mes, la empresa revisa el nivel de inventario y decide cuántos artículos pedir a su proveedor. Si la empresa ordena Z artículos, incurre en un costo de K + iZ, donde K = $32 es el costo de preparación e i = $3 es el incremento coste incremental por artículo pedido. (Si Z = 0, no incurre ningún costo). Cuando se realiza un pedido, el tiempo requerido para que llegue (llamado retraso en la entrega o tiempo de entrega) es una variable aleatoria que se distribuye uniformemente entre 0,5 y 1 mes.

La empresa utiliza una política estacionaria (s, S) para decidir cuánto pedir, es decir,

$$Z = \begin{cases} S - I & if I < s \\ 0 & if I \leq s \end{cases} \tag{10}$$

donde I es el nivel de inventario al comienzo del mes.

Para nuestro modelo, supondremos que la empresa incurre en un costo de mantenimiento de h = $1 por artículo por mes mantenido en el inventario (positivo). El costo de mantenimiento incluye costos tales como alquiler de almacenes, seguros, impuestos y mantenimiento, así como el costo de oportunidad de tener capital inmovilizado en el inventario en lugar de invertirlo en otra parte. Hemos ignorado en nuestra formulación el hecho de que todavía se incurre en algunos costos de mantenimiento cuando I = 0. Sin embargo, dado que nuestro objetivo es comparar las políticas de pedido, ignorar este factor, que después de todo es independiente de la política utilizada, no afectará nuestra evaluación de qué política es la mejor. Ahora, dado que I es el número de elementos retenidos en el inventario en el momento t, la cantidad promedio de tiempo (por mes) de artículos mantenidos en el inventario durante el período de n meses es

$$\bar{I}^{+} = \frac{\int_0^n I^{+}(t)dt}{n} \tag{11}$$

De manera similar, suponga que la compañía incurre en un costo de trabajo pendiente de p = $5 por artículo por mes en reserva; esto representa el costo del mantenimiento adicional de activos cuando existe acumulación de pedidos, así como pérdida de la buena voluntad de los clientes. El número promedio de tiempo de los elementos en la reserva es:

$$\bar{I}^{-} = \frac{\int_0^n I^{-}(t)dt}{n} \tag{12}$$

por lo que el costo promedio de la cartera de pedidos por mes es $\pi I$.

## 5. Análisis de resultados

### 5.1. Modelo MM1 en AnyLogic

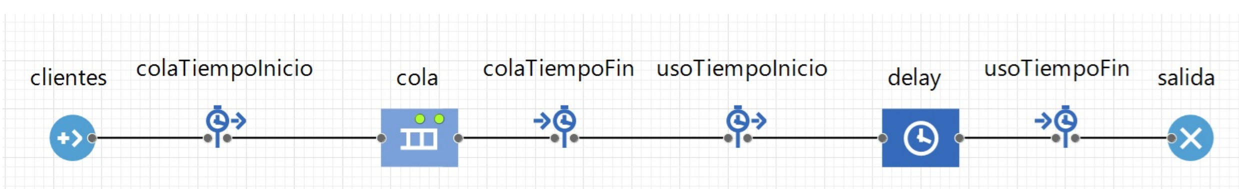


Figura 1: Bloques del modelo MM1 en Anylogic.

Para analizar el rendimiento del modelo, variamos la tasa de arribos ($T_a$) en base a la tasa de servicio ($T_s$) y la capacidad de la cola ($cap$).

Realizaremos 10 simulaciones de 1000 clientes cada una y promediamos los siguientes estadísticos:

- Demora promedio esperada en cola ($d(n)$)
- Cantidad de clientes en cola en promedio ($q(n)$)
- Ocupación del servidor ($u(n) * 100\%$)
- Promedio de clientes en el sistema ($q(n) + u(n)$)

- Probabilidad de denegación del servicio. ($p(den)$)
- Probabilidad de encontrar n clientes en cola. ($p(Q(t) = n) \times 100\,\%$)

Los primeros 5 fueron tabulados y el último fue graficado.

### 5.1.1. $T_a = 25\,\% * T_s$

| $cap$ | $d(n)$ | $q(n)$ | $u(n) \times 100\,\%$ | $q(n) + u(n)$ | $p(den)$ |
|---|---|---|---|---|---|
| 0 | N/A | N/A | $19,8$ | $1,98$ | $20,16$ |
| 2 | $0,5771$ | $0,0718$ | $24,8$ | $2,5728$ | $1,21$ |
| 5 | | | | | |
| 10 | | | | | |
| 50 | | | | | |
| $\infty$ | | | | | N/A |

## 6. Conclusiones

## 7. Referencias

https://es.wikipedia.org/wiki/Proceso$_estoc$

Arash Mahdavi, Simulation Modeling Consultant, The AnyLogic Company M/M/1// (single-server queues). En The Art of Process-Centric Modeling with AnyLogic, pages 81–103. 2020. Disponible en https://www.anylogic.com/resources/books/the-art-of-process-centric-modeling-with-anylogic/

Averill M. Law, W. David. Kelton Simulation of a single-server queueing system. En Simulation Modeling Analysis, pages 13–74. 199