



Benemérita Universidad Autónoma de Puebla

---

Facultad de Ciencias Físico Matemáticas

---

Análisis del modelo de colas  $M/M/1/\infty$  para la modelación y optimización en calles unicarriles

Tesis presentada al

**Academia de Actuaría**

como requisito parcial para la obtención del grado de

**LICENCIADO EN ACTUARÍA**

por

Melanie Aidee López Romero

Asesorado por

Dr. Francisco Solano Tajonar Sanabria

Puebla Pue.  
30 de agosto de 2025



**Título:** Análisis del modelo de colas M/M/1/∞ para la modelación y optimización en calles unicarriles

**Estudiante:** MELANIE AIDEE LÓPEZ ROMERO

COMITÉ

---

Act. Brenda Zavala López  
Presidente

---

Dr.Fernando Velazco  
Secretario

---

Dr.José Asunción Hernández  
Vocal

---

Vocal

---

Dr.Francisco Solano Tajonar Sanabria  
Asesor



# Agradecimientos

Con profundo amor y gratitud, quiero comenzar agradeciendo a las personas más importantes en mi vida, a mi pilar fundamental y mi fuente inagotable de inspiración, mi familia.

A mi mamá, por ser mi más grande inspiración, por enseñarme día a día con su ejemplo a ser una mujer fuerte, inteligente, auténtica y amorosa, gracias por todo el apoyo que me has dado, por ser mi confidente y mi soporte emocional, por todas y cada una de tus acciones que sin saberlo se convierten en mi motivación diaria. Gracias por todos los esfuerzos y sacrificios que has hecho para darme la oportunidad de estar aquí, hoy cumplo esta meta gracias a ti, agradezco la dicha de ser tu hija, las palabras nunca serán suficientes para expresarte la gran admiración y agradecimiento que te tengo.

A mi papá, por ser el hombre mas importante de mi vida y mi mas grande mentor, estas presente en todos y cada uno de mis logros, metas y aciertos, porque soy un conjunto de tu esfuerzo, dedicación y amor, admiro tu fortaleza, decisión, valentía y todas aquellas veces que fuiste duro pero justo. Gracias por siempre creer en mi e impulsarme a llegar más lejos, por hacerme ver que ningún sueño es demasiado grande si te tengo conmigo. Gracias a tu apoyo incondicional es que hoy estoy aquí.

A mi hermano, Jorge, gracias por ser el optimista de los dos, por todas las cosas que me has enseñado a valorar, por siempre estar a mi lado en los momentos difíciles y mejorar los momentos buenos, gracias por darme esperanza porque en ti veo un potencial infinito que se que aprovecharas, por iluminar mi vida con tu brillo y ser el mejor regalo que la vida me pudo dar.

A mis abuelos, abuelo Chuy y abuela Chayo, hoy quiero agradecerles tantas cosas, aunque se que todas mis palabras no bastarían para llenar todo lo que ustedes han escrito en mi vida gracias por siempre creer en mi, por cada consejo, regaño y felicitación, por cada domingo de café acompañado de lecciones e historias invaluable. Gracias a su apoyo y amor es que hoy estoy cumpliendo esta meta.

Los amo.

También quiero agradecer a las personas que me acompañaron en todo este proceso, a mis amigos, por ser mi red de apoyo en los momentos más difíciles y por celebrar conmigo cada pequeño triunfo. Especialmente quiero agradecer a Brayam por siempre estar dispuesto a escucharme, apoyarme y creer en mi aun cuando yo no lo hacia, por ser el amigo más leal y mostrarme cuan pura puede ser tu amistad, porque se que siempre serás el que decida sentarse a mi lado.

A Elizabeth por siempre aplaudir y acompañarme en cada meta cumplida, por hacer mi vida mas ligera con tu compañía y por darme fuerzas cuando más lo necesitaba, tu amistad ha sido pieza clave en este logro.

A Martin por impulsarme a tomar nuevos retos, creer en mi y siempre propiciar mi crecimiento, estoy muy agradecida de haberme cruzado contigo en la carrera y tener la dicha de haber compartido platicas, dolores y panes durante las clases.

Por ultimo pero no menos importante, agradezco inmensamente a el Dr. Francisco Solano Tajonar Sanabria, no solo por ser el director de mi tesis, si no también por encaminarme en el estudio de la probabilidad, así como también por creer en mi e impulsarme a salir de mi zona de confort, arriesgarme a tomar retos complejos y ser una guía tanto profesional como personalmente.

Así mismo, extiendo mi gratitud a la Benemérita Universidad Autónoma de Puebla por ser el cimiento de mi formación académica. Especialmente agradezco a la facultad de ciencias físico matemáticas por la calidad de sus docentes, así como por brindarme las herramientas y el entorno necesario para concluir mi formación. Confío en que la educación no debe ser un privilegio, sino un derecho universal y esta institución es prueba de esto.

# Índice general

<b>Introducción</b>	<b>VII</b>
<b>1. Preliminares</b>	<b>1</b>
1.1. Espacio de probabilidad . . . . .	1
1.2. Variables aleatorias . . . . .	2
1.3. Función de distribución . . . . .	3
1.4. Función de densidad . . . . .	5
<b>2. Cadenas de Markov</b>	<b>7</b>
2.1. Propiedad de Markov . . . . .	7
2.2. Probabilidades de transición . . . . .	8
2.3. Clasificación de estados . . . . .	13
2.4. Tiempos de alcance . . . . .	14
2.5. Propiedad fuerte de Markov . . . . .	14
<b>3. Modelo de colas M/M/1/∞</b>	<b>15</b>
3.1. Distribución exponencial . . . . .	15
3.2. Proceso de Poisson . . . . .	18
3.3. Cadena de Markov a tiempo continuo . . . . .	20
<b>4. Índices de productividad y aplicación</b>	<b>23</b>
4.1. Índices de productividad del sistema . . . . .	24
4.1.1. Ecuaciones diferenciales de Kolmogorov . . . . .	24
4.1.2. Sistema de equilibrio . . . . .	26
4.1.3. Valor medio estacionario del número de demandas en el sistema de colas . . . . .	26
4.1.4. Longitud media estacionaria de la cola . . . . .	27
4.1.5. Distribución estacionaria del tiempo de espera de inicio del servicio . . . . .	27
4.1.6. Distribución estacionaria del tiempo de estancia de una demanda en el sistema . . . . .	28
4.1.7. Tiempo medio estacionario de espera de inicio del servicio . . . . .	28
4.1.8. Tiempo medio de estancia en el sistema . . . . .	28
4.1.9. Formulas de Little . . . . .	28
4.2. Aplicación del modelo . . . . .	29
4.2.1. Calculo de los índices de productividad . . . . .	31
4.2.2. Calculo de los índices de productividad con $\mu = 9,75$ . . . . .	32
4.2.3. Calculo de los índices de productividad con una fase de avance mas larga . . . . .	33
4.2.4. Conclusiones de la aplicación . . . . .	34
<b>5. Conclusión</b>	<b>35</b>
<b>Anexos</b>	<b>37</b>
<b>Bibliografía</b>	<b>39</b>



# Introducción

Día a día se presentan constantemente situaciones en las que es necesario la formación de una línea de espera o cola, la cual puede ser definida como la línea en la que los clientes se acumulan en espera a ser atendidos por uno o varios servidores. Para este trabajo en particular se estudiara solo el modelo de colas el cual consiste en un servidor con capacidad infinita, el cual es mejor conocido como modelo de colas  $M/M/1/\infty$ .

La Teoría de Colas tiene sus inicios gracias a los aportes del ingeniero Agner Krarup Erlang quien en 1909 fue el encargado de analizar y buscar una solución a la congestión de tráfico telefónico en un hotel de Copenhague, una vez que Erlang encontró una forma de modelar este problema, este método fue utilizado en el sistema telefónico de Copenhague con el objetivo de cumplir la demanda incierta de servicios, presentando así las bases para describir el comportamiento matemático de las líneas de espera, así como sus características. Actualmente la teoría de colas es una rama de las matemáticas aplicadas que se encarga de investigar los índices de productividad de sistemas de diferentes tipos a los que se denominan Sistemas de Colas, cuya función es procesar demandas que entran en ellos para obtener un servicio.

El objetivo principal de este trabajo es dar a conocer mediante la teoría matemática correspondiente la utilidad de el estudio de las líneas de espera en especial el modelo antes mencionado, así como también incentivar el uso y aplicación de estos modelos para lograr la optimización y mejora de los mismos. Además, también se incluye en ejemplo practico que refleja la utilidad e importancia de hacer uso de esta teoría matemática.

La tesis esta estructurada de la siguiente manera.

En el capítulo 1 se hace una breve recopilación de los conceptos básicos que se necesitan para poder abordar los elementos de nuestro modelo de colas, en este capítulo se encontraran definiciones tales como el de variable aleatoria, la función de distribución y la función de densidad entre otros conceptos que serán claves para el desarrollo de este trabajo.

La teoría de colas tiene sus fundamentos en los procesos estocásticos, esencialmente en las cadenas de Markov, es por eso que el capítulo 2 se basa en desarrollar los conceptos básicos para el entendimiento de estos modelos matemáticos, entre los cuales se encuentran la definición de proceso estocástico, propiedad de Markov, tiempos de alcance y más conceptos indispensables.

Posteriormente en el capítulo 3 se comenzara con la construcción del modelo de colas  $M/M/1/\infty$  mediante la definición de distribución exponencial y proceso de Poisson, así como algunas propiedades importantes.

Por ultimo en el capítulo 4 se tratarán los índices de productividad del sistema, a través de los cuales se puede analizar el comportamiento de el modelo estudiado. Además, en este mismo capítulo se presenta un ejemplo practico en el que se analiza una calle unicarril ubicada en el centro histórico de Puebla y se dan sugerencias para optimizar el flujo mediante los índices de productividad.



# Capítulo 1

## Preliminares

Los problemas de líneas de espera se presentan constantemente en la vida diaria, tal es el caso, de ir al supermercado, banco, trámites, en el semáforo o en cualquier situación en la que sea necesario realizar una fila. Para poder estudiar y desarrollar este tipo de sistemas de líneas de espera necesitamos bases sólidas de probabilidad y más aun de procesos estocásticos, es por eso que iniciaremos con algunos conceptos importantes para este estudio.

### 1.1. Espacio de probabilidad

La teoría de la probabilidad es el estudio de aquellos modelos matemáticos de fenómenos aleatorios, entendiendo por fenómeno aleatorio a aquel que se caracteriza por la propiedad de que al ser observado bajo un determinado conjunto de condiciones, no siempre se obtiene el mismo resultado. De aquí que, al conjunto que contiene todos los resultados posibles del experimento se le conoce como espacio muestral y se denota por  $\Omega$ . Un ejemplo claro de esto es tomar como experimento aleatorio el resultado al lanzar un dado, para este caso el espacio muestral es;  $\Omega = \{1, 2, 3, 4, 5, 6\}$ . A partir de este concepto podemos definir entonces a un evento como un subconjunto del espacio muestral  $\Omega$ . A los eventos se les denota a través de letras mayúsculas del abecedario: A, B, C, ... Retomando el ejemplo que utilizamos para el espacio muestral, en este caso, definiendo nuestro evento de interés a los números impares, entonces el evento será;  $A = \{1, 3, 5\}$ .

Una vez que tenemos estas dos definiciones, se procede a definir un conjunto que agrupa a todos los eventos de un mismo experimento aleatorio y para los cuales se puede asignar sus probabilidades.

**Definición 1.1.1.** Sea  $\Omega$  un conjunto arbitrario a cuyos elementos se les nombra sucesos o eventos. Una clase no vacía  $\mathcal{F}$  de subconjuntos de  $\Omega$  es una  $\sigma$ -álgebra si cumple:

- (i)  $\Omega \in \mathcal{F}$ .
- (ii) Si  $A \in \mathcal{F}$  entonces  $A^c \in \mathcal{F}$ .
- (iii) Si  $\{A_n\}_{n \geq 1}$  es una sucesión numerable de elementos en  $\mathcal{F}$  entonces;

$$\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}. \quad (1.1)$$

Notemos que la  $\sigma$ -álgebra asociada con un espacio muestral particular describe qué subconjuntos de ese espacio muestral son eventos medibles. Siguiendo el ejemplo del lanzamiento de un dado equilibrado de seis caras, el espacio muestral consiste en los resultados posibles, que son los números del 1 al 6, por lo que la  $\sigma$ -álgebra asociada es:

$$\mathcal{F} = \{\emptyset, \{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}, \{1, 2\}, \{1, 3\}, \dots, \{1, 2, 3\}, \dots, \{1, 2, 3, 4, 5, 6\}\}. \quad (1.2)$$

**Definición 1.1.2.** *Sigma-álgebra generada.*

Sea  $\mathcal{C}$  una colección no vacía de subconjuntos de  $\Omega$ . La  $\sigma$ -álgebra generada por  $\mathcal{C}$  es denotada con  $\sigma(\mathcal{C})$  y se define como,

$$\sigma(\mathcal{C}) = \bigcap \{ \mathcal{F} : \mathcal{F} \text{ es sigma-álgebra y } \mathcal{C} \subseteq \mathcal{F} \}. \quad (1.3)$$

Note que la  $\sigma$ -álgebra generada es la  $\sigma$ -álgebra más pequeña que contiene a todos los subconjuntos medibles en  $\mathcal{C}$ . En otras palabras, es la mínima  $\sigma$ -álgebra que hace que los conjuntos dados sean eventos medibles. A partir de esta definición podemos llegar a un concepto esencial para la construcción y definición de las variables aleatorias.

**Definición 1.1.3.** *Sigma-álgebra de Borel.*

Se denota a la  $\sigma$ -álgebra de Borel en  $\mathbb{R}$  como la  $\sigma$ -álgebra generada por la clase de todos los intervalos, es decir, si tomamos a  $\mathcal{S}$  como la clase de todos los intervalos abiertos, cerrados, semiabiertos o semicerrados de  $\mathbb{R}$ , en especial esta se define como,

$$\mathcal{B}(\mathbb{R}) = \sigma\{\mathcal{S}\}.$$

A los elementos de  $\mathcal{B}(\mathbb{R})$  se les llama conjuntos de Borel o Borelianos de  $\mathbb{R}$ .

**Proposición 1.1.1.** La  $\sigma$ -álgebra generada por la clase de intervalos de la forma  $(-\infty, x]$  es  $\sigma$ -álgebra de Borel, esto es:

$$\mathcal{B}(\mathbb{R}) = \sigma\{(-\infty, x] : x \in \mathbb{R}\}.$$

*Demostración.*

$$\forall x \in \mathbb{R}, (-\infty, x] = \bigcap_{n \in \mathbb{N}} (-\infty, x + n^{-1}),$$

Puesto que  $n^{-1} \downarrow 0$ .

Debido a que es una intersección numerable de conjuntos abiertos, el conjunto  $(-\infty, x]$  está en  $\mathcal{B}$ . Resta probar que cualquier subconjunto abierto  $A \subseteq \mathbb{R}$  está en  $\mathcal{B}(\mathbb{R})$ . Como  $A$  es abierto, es una unión numerable de intervalos abiertos, por lo que ahora debe mostrarse que para  $a, b \in \mathbb{R}, a < b$ ,

$$(a, b) \in \mathcal{B}(\mathbb{R}).$$

Pero para cualquier  $u > a$ :

$$(a, u] = (-\infty, u] \cap (-\infty, a]^c \in \mathcal{B}(\mathbb{R}).$$

Y como para  $\epsilon = \frac{1}{2}(b - a)$ :

$$(a, b) = \bigcup_n (a, b - \epsilon n^{-1}).$$

Puede verse que:  $(a, b) \in \mathcal{B}(\mathbb{R})$ . □

## 1.2. Variables aleatorias

**Definición 1.2.1.** Considérese un espacio medible  $(\Omega, \mathcal{F})$ . Una **medida de probabilidad** se define como una función  $P : \mathcal{F} \rightarrow [0, 1]$  tal que:

- (i)  $P(\Omega) = 1$ .
- (ii)  $\forall A \in \mathcal{F}; P(A) \geq 0$ .
- (iii) Sea  $\{A_n\}_{n \geq 1}$  una sucesión de eventos ajenos, entonces.

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n). \quad (1.4)$$

**Definición 1.2.2.** Se denomina **espacio de probabilidad** a la terna  $(\Omega, \mathcal{F}, P)$ . En donde  $\Omega$  es el espacio muestral de nuestro experimento aleatorio,  $\mathcal{F}$  es la sigma-álgebra asociada y  $P$  una medida de probabilidad sobre  $\mathcal{F}$  [Beichelt,2016].

Ahora que tenemos todos estos conceptos básicos de probabilidad, es momento de conocer el concepto de variable aleatoria, una herramienta fundamental en la teoría de la probabilidad y la estadística.

**Definición 1.2.3.** (*Variable aleatoria.*)

Sea  $(\Omega, \mathcal{F}, P)$  un espacio de probabilidad y  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  un espacio medible, una variable aleatoria es una función o mapeo:

$$X : (\Omega, \mathcal{F}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R})).$$

Esto es;

$$X^{-1}(B) \in \mathcal{F}, \quad \forall B \in (\mathcal{B}(\mathbb{R})).$$

La definición anterior se puede encontrar en [UDP,2023].

A partir de esta definición notemos que una variable aleatoria proporciona una manera de cuantificar los resultados de un fenómeno aleatorio mediante la transformación de los resultados del espacio muestral a un valor dentro de los números reales. Las variables aleatorias se clasifican en dos tipos, los cuales son:

- **Variables aleatorias discretas:** Este tipo de variables toman un número finito o infinito numerable de posibles valores.

Un ejemplo de este tipo de variables es el número de "3" que obtenemos al lanzar dos dados.

$$X : \Omega \rightarrow \mathcal{R},$$

$$\Omega = \{(1, 1), (1, 2), \dots, (3, 1), (3, 2), (3, 3), \dots, (6, 6)\}.$$

*w = número de "3" que ocurren en cada posible resultado.*

$$R_x = \{0, 1, 2\}. \text{ (Llamado recorrido o rango de } x)$$

X es una variable aleatoria que toma valores 0,1,2, y cada uno de ellos lo tomará con una probabilidad de  $\frac{25}{36}, \frac{10}{36}, \frac{1}{36}$  respectivamente.

- **Variables aleatorias continuas:** Estas variables pueden tomar un conjunto infinito y no numerable, incluso cualquier valor de  $\mathbb{R}$ .

Un ejemplo de esto es el tiempo de fallo de un foco.

A partir de estos dos tipos de variables aleatorias también tendremos dos definiciones de suma importancia para nuestro estudio.

Hay dos tipos principales de funciones asociadas a toda variable aleatoria X: la función de distribución acumulativa (F.D.A) y la función de densidad de probabilidad (fdp), dependiendo de si la variable aleatoria es discreta o continua. Estas funciones son fundamentales en estadística y probabilidad, ya que permiten describir y analizar el comportamiento de las variables aleatorias, proporcionando herramientas para calcular probabilidades y realizar inferencias sobre la distribución de los datos.

### 1.3. Función de distribución

**Proposición 1.3.1.** Sea  $(\Omega, \mathcal{F}, P)$  un espacio de probabilidad y  $X : \Omega \rightarrow \mathbb{R}$  una variable aleatoria, la *función de distribución* (F.D. o F.D.A.) de X, denotada por  $F(x)$  y que es la función  $F(x) : \mathbb{R} \rightarrow [0, 1]$  definida por:

$$F(x) = P[\{\omega \in \Omega : X(\omega) \leq x\}] = P(X \leq x).$$

F será una función de distribución si satisface:

i)

$$\lim_{x \rightarrow -\infty} F(x) = 0, \quad \lim_{x \rightarrow \infty} F(x) = 1.$$

ii) F es continua por la derecha y tiene límites por la izquierda.

iii) F es monótona no decreciente.

La demostración de esta proposición se encuentra en [Ross,2019] y [Seactuario,2023].

**Ejemplo 1.3.1.** Sea  $X$  una variable aleatoria con función de distribución especificada abajo. Grafique  $F(x)$  y demuestre que es efectivamente una función de distribución.

$$F(x) = \begin{cases} 0 & \text{si } x < 2 \\ 1 - \frac{4}{x^2} & \text{si } x \geq 2 \end{cases}$$

Solución:

Tenemos que,

$$\lim_{x \rightarrow -\infty} F(x) = \lim_{x \rightarrow -\infty} 0 = 0.$$

Además,

$$\lim_{x \rightarrow +\infty} F(x) = \lim_{x \rightarrow +\infty} 1 - \frac{4}{x^2} = 1 - 0 = 1.$$

Vemos también que;

$$F'(x) = -4(-2)x^{-3} = \frac{8}{x^3}, \text{ cuando } x \geq 2.$$

Luego,  $F'(x) > 0$  cuando  $x \geq 2$ , por lo que es creciente en  $[2, +\infty)$ .

Para  $x^2 \geq 4 \implies \frac{-4}{x^2} \geq -1 \implies 1 - \frac{4}{x^2} \geq 1 - 1 = 0$ . Por lo tanto,  $F(x) \geq 0$ , y por lo tanto, es creciente en todo  $\mathbb{R}$ . Probamos que  $F(x)$  es continua por la derecha  $\forall x, x \neq 2$  y que:

$$\lim_{x \rightarrow 2^+} F(x) = F(2).$$

Por lo que es continua por la derecha en todo  $\mathbb{R}$ , siempre que la variable aleatoria es continua. Ya probamos las propiedades de la función de distribución. Y su gráfica es la siguiente.

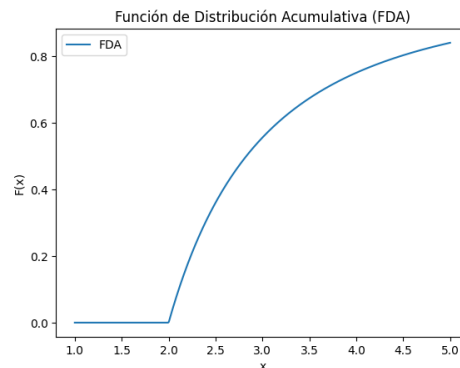


Figura 1.1: Gráfica de  $F(x)$ .

## 1.4. Función de densidad

**Definición 1.4.1.** Se dice que una variable aleatoria  $X$  es continua si su recorrido es finito y no numerable, además, existe una función  $f$  llamada *función de densidad de probabilidad* (fdp) que satisface las siguientes condiciones:

i)  $f(x) \geq 0; \forall x.$

ii)  $\int_{-\infty}^{\infty} f(x) dx = 1.$

iii) Para cualesquiera  $a, b \in \mathbb{R}$  con  $a < b$ , se tiene que  $P[a \leq X \leq b] = \int_a^b f(x) dx.$

[Meyer,1973].

La función de distribución  $F(x)$  determina de manera única a la variable aleatoria  $X$ . Además, con  $F(x)$  se pueden calcular probabilidades de eventos en el recorrido de  $X$  y que a partir de ella se pueden obtener también las probabilidades puntuales y la fdp de  $X$  como lo muestra el siguiente resultado.

**Ejemplo 1.4.1.** Grafiquemos la función de densidad de probabilidad de la variable aleatoria  $Y$  que está dada por,

$$f(y) = \begin{cases} \frac{1}{8} (y + 1), & \text{para } 2 < y < 4, \\ 0, & \text{otro caso.} \end{cases}$$

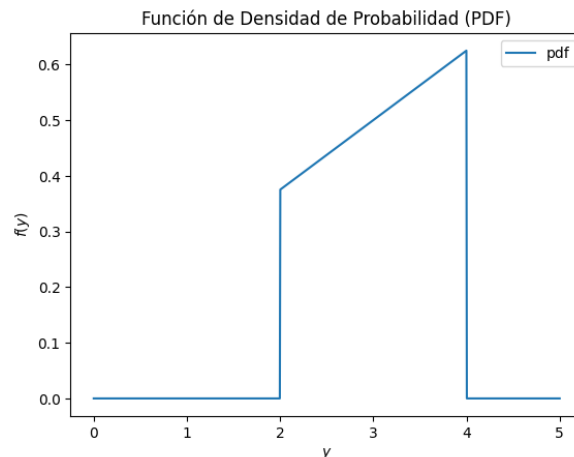


Figura 1.2: Gráfica de la función de densidad de probabilidad de  $f(y)$ .

**Teorema 1.4.1.** i) Sea  $F$  la fda de una variable aleatoria continua con fdp  $f$ . Luego,

$$f(x) = \frac{d}{dx}F(x).$$

Para toda  $x$  en la cual  $F$  es diferenciable.

ii) Sea  $X$  una variable aleatoria discreta con valores posibles  $x_1 < x_2 < \dots < x_i$ . Sea  $F$  la fda de  $X$ . Entonces,

$$p(x_i) = P[X = x_i] = F(x_i) - F(x_{i-1}).$$



## Capítulo 2

# Cadenas de Markov

Estos modelos matemáticos tienen sus inicios en 1906 después de que el matemático Andréi Markov publicó su famoso artículo *On Certain Aspects of the Behavior of Random Variables* en donde Markov introdujo el concepto de lo que ahora conocemos como *cadena de Markov*. La idea central de las cadenas de Markov es que la probabilidad de que un sistema evolucione a un estado futuro solo depende del estado presente y no de los estados pasados.

Inicialmente, las cadenas de Markov fueron utilizadas para estudiar el comportamiento de sucesiones aleatorias, pero con el tiempo se descubrió que tenían aplicaciones en una amplia gama de campos, como la economía, la física, la biología, la ingeniería, las ciencias de la computación y la inteligencia artificial. Para poder iniciar con el estudio de estas es necesario definir un concepto esencial para las cadenas de Markov, el cual es *un proceso estocástico*. Un proceso estocástico es una colección de variables aleatorias que evolucionan en función de algún parámetro, generalmente el tiempo. Matemáticamente, un proceso estocástico se define como una familia de variables aleatorias indexadas por un conjunto, que puede ser discreto o continuo.

A continuación tenemos la definición de proceso estocástico, así como algunos tipos de procesos, estas se encuentran en [Hoel, 1972] y [Rincón, 2012].

**Definición 2.0.1.** Un proceso estocástico es una colección o familia de variables aleatorias  $\{X_t : t \in T\}$ , ordenados según el subíndice  $t$ , que generalmente suele ser el tiempo. Las variables toman valores en un conjunto  $S$  llamado espacio de estados.

Es interesante notar que para cada instante  $t$  tendremos una variable aleatoria distinta representada por  $X_t$ . A los posibles valores que puede tomar la variable aleatoria se les denominara estados, por lo que se puede tener un espacio de estados discreto y un espacio de estados continuo. Además, la variable tiempo puede ser de tipo discreto o de tipo continuo.

Para este capítulo nuestro interés está en estudiar a los procesos estocásticos a tiempo discreto, en los cuales el tiempo se mueve en forma discreta y la variable aleatoria sólo toma valores discretos en el espacio de estados, mejor conocidos como **cadenas**.

### 2.1. Propiedad de Markov

Una característica importante de las cadenas es poder observar como la probabilidad de un sistema evoluciona a un estado futuro dependiendo solamente del estado presente y no de los anteriores. Esta característica es conocida como "*Propiedad de Markov*" y aquellos procesos que poseen esta propiedad son llamados cadenas de Markov. Ahora, tomando un proceso estocástico con  $\mathcal{G}$  un conjunto de estados finito, que además es un subconjunto de los enteros. Entonces a  $\mathcal{G}$  lo llamaremos *espacio de estados*, observaremos a este proceso en momentos discretos del tiempo  $n = 1, 2, \dots$ , y denotaremos con  $X_n$  el estado del proceso al tiempo  $n$ .

**Definición 2.1.1.** La propiedad de Markov se define como

$$P(X_{n+1} = x_{n+1} | X_0 = x_0, \dots, X_n = x_n) = P(X_{n+1} = x_{n+1} | X_n = x_n).$$

Para cada entero no negativo  $n$  y cualesquiera números  $x_0, \dots, x_{n+1} \in \mathcal{G}$ .

Presentamos a las cadenas de Markov como primer ejemplo de procesos estocásticos ya que estas son el modelo matemático mas simple para un fenómeno aleatorio que involucra , además de que su estructura poco compleja nos permite decir mucho a cerca de su comportamiento.

Partiendo de las definiciones anteriores, es claro que se puede definir el concepto de esperanza condicional no sólo respecto a una  $\sigma$ -álgebra, sino también con respecto a una variable aleatoria y más aún, a una sucesión posiblemente infinita de estas mismas, lo que da origen al estudio de los procesos estocásticos.

## 2.2. Probabilidades de transición

Uno de los principales intereses al estudiar cadenas de Markov es el cálculo de probabilidades de transición, las cuales indican la probabilidad de que nuestro proceso pase de un estado a otro en un paso. Estas se definen a través de las probabilidades condicionales  $P(X_{n+1} = y | X_n = x)$ .

**Definición 2.2.1.** *Probabilidad de transición en un paso.*

Suponemos que siempre que el proceso esta en el estado  $x$ , existe una probabilidad fija  $p_{xy}(n, n + 1)$  de que el proceso se encontrara en el estado  $y$  en el siguiente paso por simplicidad las probabilidades de transición en un paso se escriben como  $p_{xy}$  a partir de esto, podemos ver que se obtiene una matriz siempre que el espacio de estados sea finito, por lo que la matriz de transición en un paso es la siguiente,

$$P = \begin{bmatrix} p_{00} & p_{01} & \dots & p_{0y} & \dots \\ p_{10} & p_{11} & \dots & p_{1y} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{x0} & p_{x1} & \dots & p_{xy} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}. \quad (2.1)$$

Al representarlo en forma matricial nos facilita el calculo de las probabilidades de transición en más de un paso. Por ejemplo,  $P \times P = P^2$  corresponde a las probabilidades de transición en dos pasos, así también  $P^m$  corresponde a las probabilidades de transición en m pasos. A la matriz  $P^m$  se conoce como la matriz de transición en m pasos de la cadena de Markov, la cual sera estudiada mas adelante.

La matriz de transición  $P = (p_{xy})$  cumple con las siguientes propiedades:

- i)  $p_{xy} \geq 0, \quad x, y \in \mathcal{G}$ ,
- ii)  $\sum_{x=1} p_{xy} = 1, \quad x \in \mathcal{G}$ .

En general, toda matriz cuadrada que cumpla con estas dos propiedades se le conoce como una matriz estocástica. Esta misma nos muestra la esencia del proceso y determina el comportamiento de la cadena en cualquier tiempo futuro.

Para el caso especial de este trabajo, solo estudiaremos las cadenas de Markov con probabilidades de transición estacionarias, i.e., aquellas que  $P(X_{n+1} = y | X_n = x)$  son independientes de  $n$ .

Veamos la aplicación de estos conceptos a través del siguiente ejemplo.

**Ejemplo 2.2.1.** Una urna contiene dos bolas sin pintar. Se selecciona una bola al azar y se lanza una moneda. Si la bola elegida no está pintada y la moneda produce cara, pintamos la bola de rojo; si la moneda produce cruz, la pintamos de negro. Si la bola ya está pintada, entonces cambiamos el color de la bola de rojo a negro o de negro a rojo, independientemente de si la moneda produce cara o cruz. Modelemos el problema como una cadena de Markov y veamos su matriz de probabilidades de transición. El primer paso es identificar los estados, para esto, se escribirá en forma de vectores con

el siguiente formato: [S R N] con S el número de bolas sin pintar, R el número de bolas rojas y N el número de bolas negras que hay en la urna.

$E_0$ : [2 0 0] Inicialmente, cuando las dos bolas están sin pintar.

$E_1$ : [1 1 0] Una bola pintada de rojo.

$E_2$ : [1 0 1] Una bola pintada de negro.

$E_3$ : [0 1 1] Una bola roja y una negra.

$E_4$ : [0 2 0] Dos bolas rojas.

$E_5$ : [0 0 2] Dos bolas negras.

Ahora veamos que las probabilidades de transición estarán dadas por, de  $E_0$  sólo se puede pasar a:

- $E_1$ : si la bola que saquemos (que siempre será sin pintar  $p = 1$ ) la pintamos de rojo (cae cara  $p = 1/2$ .)
- $E_2$ : si la bola que saquemos la pintamos de negro (cae cruz  $p = 1/2$ ).

De  $E_1$  podemos pasar a:

- $E_3$ : si se saca la bola sin pintar ( $p = 1/2$ ) y se pinta de negro (cae cruz  $p = 1/2$ ).
- $E_4$ : si se saca la bola sin pintar ( $p = 1/2$ ) y se pinta de rojo (cae cara  $p = 1/2$ ).
- $E_2$ : si se saca la bola ya pintada (de rojo  $p=1/2$ ) y se le cambia color (a negro  $p = 1$ ).

De  $E_2$  podemos pasar a:

- $E_3$ : si se saca la bola sin pintar ( $p = 1/2$ ) y se pinta de rojo (cae cara  $p = 1/2$ ).
- $E_5$ : si se saca la bola sin pintar ( $p = 1/2$ ) y se pinta de negro (cae cruz  $p = 1/2$ ).
- $E_1$ : si se saca la bola ya pintada (de negro  $p = 1/2$ ) y se le cambia color (a rojo  $p = 1$ ).

De  $E_3$  podemos pasar a:

- $E_4$ : si se saca la bola negra ( $p = 1/2$ ), entonces se le cambia de color a rojo ( $p = 1$ ).
- $E_5$ : si se saca la bola roja ( $p = 1/2$ ), entonces se le cambia de color a negro ( $p = 1$ ).

De  $E_4$  sólo podemos pasar a  $E_3$  porque siempre sacaremos una bola roja ( $p = 1$ ) y le cambiaremos de color a negro ( $p = 1$ ).

De  $E_5$  sólo podemos pasar a  $E_3$  porque siempre sacaremos una bola negra ( $p = 1$ ) y le cambiaremos de color a rojo ( $p=1$ ). Por lo que la matriz de transición estará dada por:

$$P = \begin{bmatrix} 0 & 1(\frac{1}{2}) = \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 1(\frac{1}{2}) = \frac{1}{2} & (\frac{1}{2})(\frac{1}{2}) = \frac{1}{4} & (\frac{1}{2})(\frac{1}{2}) = \frac{1}{4} & 0 \\ 0 & 1(\frac{1}{2}) = \frac{1}{2} & 0 & (\frac{1}{2})(\frac{1}{2}) = \frac{1}{4} & 0 & (\frac{1}{2})(\frac{1}{2}) = \frac{1}{4} \\ 0 & 0 & 0 & 0 & 1(\frac{1}{2}) = \frac{1}{2} & 1(\frac{1}{2}) = \frac{1}{2} \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}. \quad (2.2)$$

Ahora, veamos que después de haber pintado dos bolas, ¿cuál es la probabilidad que el estado sea [0 2 0] (0 sin pintar, 2 rojas, 0 negras)? Notemos que se inicia con las dos bolas sin pintar ( $E_0$ ) y nos piden la probabilidad de que pasemos al estado [0 2 0] ( $E_4$ ) después de pintar dos bolas (a dos pasos). En resumen, nos piden  $P_{04}(2)$ , para obtener esto, elevaremos la matriz al cuadrado.

$$P = \begin{bmatrix} 0 & 0,25 & 0,25 & 0,25 & 0,125 & 0,125 \\ 0 & 0,25 & 0 & 0,375 & 0,125 & 0,25 \\ 0 & 0 & 0,25 & 0,375 & 0,25 & 0,125 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0,5 & 0,5 \\ 0 & 0 & 0 & 1 & 0,5 & 0,5 \end{bmatrix}. \quad (2.3)$$

De la matriz podemos ver que  $P_{04}(2) = 0,125$ , es decir, la probabilidad de que el estado sea [0 2 0] después de pintar dos bolas es 0.125.

Otra característica importante de las cadenas de Markov es la *distribución inicial*, ya que conocer esta distribución es esencial para el correcto análisis e interpretación de una cadena de Markov, debido a que esta proporciona información sobre el estado en el que inicia el sistema. Cada estado tiene asociada una probabilidad inicial que representa la probabilidad de que el sistema comience en ese estado específico. Además, conforme la cadena evoluciona a lo largo del tiempo, las probabilidades de transición entre los estados determinarán la distribución de probabilidad en los pasos siguientes.

**Definición 2.2.2.** *Distribución inicial.*

La función  $\pi_0(x)$ ,  $x, y \in \mathcal{G}$ , definida como:

- i)  $\pi_0(x) \geq 0$ ,  $x \in \mathcal{G}$ ,
- ii)  $\sum_{x=1} \pi_0(x) = 1$ ,  $x \in \mathcal{G}$ .

Al igual que para el instante inicial, se puede definir un vector de ocupación de estado (una distribución) para cada instante  $n$ .

**Definición 2.2.3.** A la probabilidad de ocupar cada uno de los estados en el instante  $n$  se le llamará  $\pi_n$ . Y contara con las siguientes propiedades:

- i)  $\pi_n(x) \geq 0$ ,  $x \in \mathcal{G}$ ,
- ii)  $\sum_{x=1} \pi_n(x) = 1$ ,  $x \in \mathcal{G}$ .

A partir de estas ultimas definiciones podemos obtener un resultado interesante, el cual es el siguiente.

**Proposición 2.2.1.** La distribución conjunta de  $X_0, \dots, X_n$  se puede expresar en términos de la probabilidad de transición y la distribución inicial.

*Demostración.* Veamos que,

$$\begin{aligned} P(X_0 = x_0, X_1 = x_1) &= P(X_0 = x_0)P(X_1 = x_1|X_0 = x_0) \\ &= \pi_0(x_0)P(x_0, x_1). \end{aligned}$$

También notemos que,

$$\begin{aligned} P(X_0 = x_0, X_1 = x_1, X_2 = x_2) &= P(X_0 = x_0, X_1 = x_1)P(X_2 = x_2|X_0 = x_0, X_1 = x_1) \\ &= \pi_0(x_0)P(x_0, x_1)P(X_2 = x_2|X_0 = x_0, X_1 = x_1). \end{aligned}$$

Y como  $X_n, n \geq 0$  satisface la propiedad de Markov, y tiene probabilidades estacionarias, se tiene que:

$$\begin{aligned} P(X_2 = x_2|X_0 = x_0, X_1 = x_1) &= P(X_2 = x_2|X_1 = x_1) \\ &= P(x_1, x_2). \end{aligned}$$

Entonces,

$$P(X_0 = x_0, X_1 = x_1, X_2 = x_2) = \pi_0(x_0)P(x_0, x_1)P(x_1, x_2).$$

Finalmente, por inducción, obtendremos:

$$P(X_0 = x_0, \dots, X_n = x_n) = \pi_0(x_0)P(x_0, x_1) \dots P(x_{n-1}, x_n).$$

□

Para un mejor y fácil entendimiento de los conceptos, ilustraremos estos a través de algunos ejemplos bastantes conocidos de cadenas de Markov.



*Demostración.*

$$\begin{aligned}
 p^{n+m}(x, y) &= P(X_{n+m} = y | X_0 = x) \\
 &= \sum_k P(X_{n+m} = y, X_n = k | X_0 = x) \\
 &= \sum_k P(X_n = z | X_0 = x) P(X_{n+m} = y | X_0 = x, X_n = z) \\
 &= \sum_z p^n(x, z) p(X_{n+m} = y | X_0 = x, X_n = z), \\
 &= \sum_z p^n(x, z) p^m(z, y).
 \end{aligned}$$

□

La ecuación de Chapman-Kolmogorov es una herramienta fundamental en la teoría de cadenas de Markov, ya que nos permite obtener la probabilidad de encontrar un sistema en un estado futuro dado el estado actual y un número específico de pasos en el futuro. Esto es esencial para comprender la evolución temporal de los sistemas estocásticos modelados por cadenas de Markov. Además, la ecuación de Chapman-Kolmogorov es esencial para modelar una amplia variedad de procesos estocásticos, desde sistemas físicos hasta fenómenos biológicos y económicos, ya que proporciona una estructura matemática sólida para describir la evolución probabilística de estos procesos, así como también esta presente en problemas de optimización y control donde es crucial comprender y predecir la evolución de sistemas dinámicos bajo incertidumbre. Una consecuencia importante de la ecuación de Chapman-Kolmogorov es el siguiente resultado.

**Proposición 2.2.3.** La probabilidad de transición de  $i$  a  $j$  en  $n$  pasos,  $p_{ij}(n)$ , está dada por la entrada  $(i, j)$  de la matriz de transición  $P$  elevada a la  $n$ -ésima potencia,

$$p_{ij}(n) = (P^n)_{ij}. \tag{2.4}$$

*Demostración.* Notemos que esta proposición es una aplicación directa de la ecuación de Chapman-Kolmogorov, y lo veremos a través de la suma correspondiente a la entrada

$$(i, j)$$

de la matriz resultante de multiplicar  $P$  consigo misma  $n$  veces,

$$\begin{aligned}
 p_{ij}(n) &= \sum_{i_1} p_{i, i_1}(1) p_{i_1, j}(n-1) \\
 &= \sum_{i_1, i_2} p_{i, i_1}(1) p_{i_1, i_2}(1) p_{i_2, j}(n-2) \\
 &= \sum_{i_1, i_2, i_3} p_{i, i_1}(1) p_{i_1, i_2}(1) p_{i_2, i_3}(1) p_{i_3, j}(n-3) \\
 &\quad \vdots \\
 &= \sum_{i_1, \dots, i_{n-1}} p_{i, i_1}(1) p_{i_1, i_2}(1) \dots p_{i_{n-1}, j}(1) \\
 &= (P^n)_{ij}.
 \end{aligned}$$

□

## 2.3. Clasificación de estados

La clasificación de los estados en una cadena de Markov es importante para comprender la dinámica y el comportamiento a largo plazo del sistema. En este contexto, los estados de una cadena de Markov se pueden clasificar en diferentes categorías en función de sus propiedades y comportamiento. Para poder presentar a las principales categorías de clasificación de estados, nos es necesario definir el siguiente concepto.

**Definición 2.3.1.** Sea  $X_n$ ,  $n \geq 0$  una cadena de Markov con espacios de estados  $\mathcal{G}$  y función de transición  $p$ . Denotemos con

$$\rho_{xy} = P_x(T_y < \infty).$$

A la probabilidad de que una cadena de Markov que inicia en  $x$  estará en el estado  $y$  en algún tiempo positivo y finito, en particular  $\rho_{yy}$  denota la probabilidad de que una cadena de Markov que inicie en el estado  $y$  en algún momento regresará a el estado  $y$ .

**Definición 2.3.2.** (*clasificación de estados*).

1. se dice que un estado  $y$  es recurrente si;

$$\rho_{yy} = 1.$$

Es decir, si la cadena de Markov que inicia en  $y$ , esta regresara a  $y$  con probabilidad 1.

2. Se dice que un estado  $y$  es transitorio si;

$$\rho_{yy} < 1.$$

Es decir, una cadena de Markov que inicia en el estado  $y$ , tiene una probabilidad positiva  $1 - \rho_{yy}$  de nunca regresar a  $y$ .

3. Tendremos que un estado absorbente  $y$  es aquel que satisface,

$$P_y(T_y = 1) = P(y, y) = 1.$$

Por lo tanto,  $\rho_{yy} = 1$ . Teniendo así que un estado absorbente es necesariamente recurrente.

Es momento de definir un concepto importante cuando hablamos de cadenas de Markov, el cual es la comunicación entre estados de la misma, ya que esto nos permitirá clasificar a estas. En general este concepto nos habla de la posibilidad de pasar de un estado a otro en algún número finito de transiciones.

**Definición 2.3.3.** Se dice que el estado  $j$  es accesible desde el estado  $i$  si existe un entero  $n \geq 0$  tal que  $p_{ij}(n) > 0$ , esto se escribe simplemente como  $i \rightarrow j$ . Se dice además que los estados  $i$  y  $j$  son comunicantes y se escribe  $i \leftrightarrow j$  si se cumple que  $i \rightarrow j$  y  $j \rightarrow i$ . Se puede verificar que la comunicación es una relación de equivalencia, por lo que cumple las siguientes propiedades,

- a) Es reflexiva:  $i \leftrightarrow i$ .
- b) Es simétrica: si  $i \rightarrow j$ , entonces  $j \rightarrow i$ .
- c) Es transitiva: si  $i \leftrightarrow j$  y  $j \leftrightarrow k$ , entonces  $i \leftrightarrow k$ .

Por lo que se induce una partición del espacio de estados de una cadena de Markov dada por subconjuntos de estados comunicantes, de modo que el espacio de estados se subdivide en clases de comunicación. A la clase de comunicación del estado  $i$  se le denota  $C(i)$ . Por lo tanto,  $i \leftrightarrow j$  si y sólo si  $C(i) = C(j)$ . A partir de esto podemos notar que si un estado es recurrente entonces todos los elementos en la misma clase de comunicación lo son, resultando así que a la clase se le denominara recurrente, de igual manera si el estado es transitorio, los estados en esa misma clase también lo serán y a la clase se le denominará transitoria. Entonces podemos observar que las cadenas se pueden clasificar de acuerdo a las clases de comunicación.

**Definición 2.3.4.** Se dice que una cadena es irreducible si todos sus estados se comunican entre sí, es decir, hay una única clase de comunicación.

## 2.4. Tiempos de alcance

Un concepto que también es de gran utilidad son los *tiempos de alcance o llegada*. Este tiene vínculos con conceptos clave en las cadenas de Markov, como los estados absorbentes y las probabilidades de transición. Específicamente nos indica el número esperado de pasos necesarios para que una cadena de Markov alcance un estado específico desde un estado inicial dado. Además, de tener una estrecha relación con la ecuación de Chapman-Kolmogorov, ya que esta ecuación permite calcular la probabilidad de alcanzar un estado en un número dado de pasos. Al analizar las probabilidades de transición y la estructura de la cadena de Markov, podemos determinar los tiempos de alcance esperados y su variabilidad. La siguiente definición se puede encontrar en [Hoel, 1972] y [Rincón, 2012].

**Definición 2.4.1.** Sea  $A$  un subconjunto de  $\mathcal{G}$ , es decir,  $A$  es subconjunto de estados o  $A$  en particular puede ser solo un estado de una cadena de Markov, el tiempo de alcance  $T_A$  de  $A$  está definido como,

$$T_A = \min(n > 0 : x_n \in A).$$

Esto es que  $T_A$  es el primer tiempo positivo en el que la cadena de Markov se encuentra en  $A$ .

**Proposición 2.4.1.**

$$p^n(x, y) = \sum_{m=1}^n P_x(T_y = m) p^{n-m}(y, y).$$

**Observación.** Recordemos que usamos  $P_x()$  para denotar probabilidades de varios eventos definidos en términos de una cadena de Markov que inicia en  $x$ .

*Demostración.* Notemos que los eventos,  $\{T_y = m, X_n = y\}, 1 \leq m \leq n$  son disjuntos y que,  $\{X_n = y\} = \bigcup_{m=1}^n \{T_y = m, X_n = y\}$ , es decir, hemos descompuesto el evento  $\{X_n = y\}$  de acuerdo con el tiempo de alcance de  $y$ . De ahí que,

$$\begin{aligned} p^n(x, y) &= P_x(X_n = y) \\ &= \sum_{m=1}^n P_x(T_y = m, X_n = y) \\ &= \sum_{m=1}^n P_x(T_y = m) P(X_n = y | X_0 = x, T_y = m) \\ &= \sum_{m=1}^n P_x(T_y = m) P(X_n = y | X_0 = x, X_1 \neq y, \dots, X_{m-1} \neq y, X_m = y) \\ &= \sum_{m=1}^n P_x(T_y = m) p^{n-m}(y, y). \end{aligned}$$

□

## 2.5. Propiedad fuerte de Markov

Tomando el concepto que recién se introdujo de tiempos de alcance y dado que para una cadena de Markov la distribución del siguiente estado solo depende del estado actual, específicamente podemos notar que la propiedad de Markov se cumple en tiempos de alcance. Para una visión más clara de esto veremos lo siguiente.

**Definición 2.5.1. Propiedad fuerte de Markov.** Supongamos que es  $T$  un tiempo de alcance.

Dado que  $T = n$  y  $X_T = y$ , cualquier otra información sobre  $X_0, \dots, X_T$  es irrelevante para predecir el futuro, y  $X_{T+k}, K \geq 0$  se comporta como una cadena de Markov con estado inicial  $y$ , es decir, la probabilidad

$$P(X_{T+k+1} = j | X_0 = x_0, \dots, X_{T+k-1} = x_{k-1}, X_{T+k} = i)$$

es igual a

$$P(X_{T+k+1} = j | X_{T+k} = i).$$

*Demostración:* Para su prueba puede consultar [Durrett, 2012].

## Capítulo 3

# Modelo de colas M/M/1/∞

Como se menciono anteriormente, la teoría de colas es un campo de estudio ampliamente aplicado el cual se centra en el análisis y la optimización de los sistemas de líneas de espera. Estos sistemas se pueden encontrar en una gran variedad de contextos, desde servicios al cliente y operaciones industriales hasta redes de computadoras y telecomunicaciones. La teoría de colas proporciona herramientas matemáticas y modelos para comprender y mejorar la eficiencia de estos sistemas, minimizando los tiempos de espera, optimizando los recursos y maximizando la satisfacción del cliente.

En nuestra exploración de la teoría de colas, estudiaremos conceptos clave como las tasas de llegada y servicio, las disciplinas de colas, las medidas de desempeño (como el tiempo de espera promedio y la utilización del sistema), en este trabajo analizaremos en especial el modelo de colas M/M/1/∞. Analizaremos cómo este modelo se aplica en situaciones del mundo real y cómo puede ser utilizado para tomar decisiones informadas que mejoren la eficiencia y la calidad del servicio.

A lo largo de nuestra discusión, descubriremos cómo la teoría de colas se ha convertido en una herramienta indispensable en la gestión de operaciones y la ingeniería de sistemas, permitiendo a las organizaciones diseñar procesos más eficientes, optimizar recursos, y brindar experiencias de usuario más satisfactorias.

Iniciaremos con la construcción de este modelo a partir de la estructura mas basica que se necesita para poder llegar a este modelo.

### 3.1. Distribución exponencial

La distribución exponencial proporciona una herramienta poderosa para comprender y predecir el comportamiento de eventos en el tiempo, por lo que exploraremos las características clave de esta distribución para posteriormente poder hablar del proceso de Poisson el cual es una pieza clave en el estudio de nuestro modelo. La siguiente definición, así como algunas de sus propiedades se pueden encontrar en [OpenStax,1999].

**Definición 3.1.1.** *Distribución exponencial.* Considérese un espacio de probabilidad  $(\Omega, \mathcal{F}, P)$  y sea  $X : (\Omega, \mathcal{F}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$  una variable aleatoria y  $\lambda$  un número real positivo, entonces se dice que  $X$  tiene distribución exponencial con parámetro  $\lambda$  ( $X \sim \text{exp}(\lambda)$ ) si,

$$F(x) = P(X \leq x) = 1 - e^{-\lambda x}.$$

O bien en terminos de la función de densidad,

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{si } x \geq 0, \\ 0, & \text{si } x < 0. \end{cases}$$

La distribución exponencial es una herramienta poderosa para modelar el tiempo entre eventos sucesivos en situaciones donde los eventos ocurren de manera independiente y a una tasa constante. Además de que cuenta con propiedades matemáticas sumamente importantes las cuales la hacen ampliamente utilizada en diversos campos como la estadística, probabilidad, matemáticas, física, ingeniería, biología, etc.

**Proposición 3.1.1.** Sea  $X \sim \text{exp}(\lambda)$ , entonces,

- $E(X) = \frac{1}{\lambda}$ .
- $V(X) = \frac{1}{\lambda^2}$ .

*Demostración.*   ▪ Para obtener la esperanza de  $X$  se debe integrar por partes, con lo que tenemos,

$$\begin{aligned} E(X) &= \int_0^{\infty} x f(x) dx \\ &= \int_0^{\infty} x \lambda e^{-\lambda x} dx \\ &= -x e^{-\lambda x} \Big|_0^{\infty} + \int_0^{\infty} e^{-\lambda x} dx \\ &= \left[ -x e^{-\lambda x} - \frac{1}{\lambda} e^{-\lambda x} \right]_0^{\infty} \\ &= \frac{1}{\lambda}. \end{aligned}$$

- Ahora para calcular la varianza tenemos,

$$V(X) = E(X^2) - (E(X))^2$$

En donde,

$$\begin{aligned} E(X^2) &= \int_0^{\infty} x^2 \lambda e^{-\lambda x} dx \\ &= -x^2 e^{-\lambda x} \Big|_0^{\infty} + \int_0^{\infty} 2x e^{-\lambda x} dx = \frac{2}{\lambda^2}. \end{aligned}$$

y  $(E(X))^2 = \frac{1}{\lambda^2}$ , entonces

$$\begin{aligned} V(X) &= \frac{2}{\lambda^2} - \frac{1}{\lambda^2} \\ &= \frac{1}{\lambda^2}. \end{aligned}$$

□

Una propiedad importante mencionar en cuanto hablamos de la distribución exponencial es la *perdida de memoria*, ya que sin duda, esta propiedad es una de las razones por las cuales la distribución exponencial es ampliamente utilizada para modelar tiempos entre eventos en situaciones donde los eventos ocurren de manera independiente en el tiempo.

La propiedad de pérdida de memoria se refiere a la falta de dependencia de memoria en los tiempos entre eventos sucesivos, es decir, la probabilidad de que ocurra un evento en el futuro no está influenciada por cuánto tiempo ha pasado desde el último evento. Formalmente esta propiedad se establece de la siguiente manera.

**Proposición 3.1.2.** *Propiedad de pérdida de memoria.* Sea  $X \sim \text{exp}(\lambda)$ , es decir,  $X$  es una variable aleatoria que representa el tiempo entre eventos sucesivos que siguen una distribución exponencial con parámetro  $\lambda$ , entonces para cualesquiera números positivos  $s$  y  $t$  se tiene,

$$P[X > s + t | X > s] = P[X > t].$$

Esto quiere decir que habiendo transcurrido ya un tiempo  $s$  desde la última llegada, la probabilidad de que aún falte un tiempo  $t$  hasta la siguiente llegada es independiente del valor de  $s$ . Como si la variable hubiese olvidado por completo el valor de  $s$ .

*Demostración.*

$$\begin{aligned}
 P[X > s + t | X > s] &= P[X > t] = \frac{P[\{X > s\} \cap \{X > s + t\}]}{P[X > s]} \\
 &= \frac{P[X > s + t]}{P[X > s]} \\
 &= \frac{e^{-\lambda(s+t)}}{e^{-\lambda(s)}} = e^{-\lambda t} \\
 &= P[X > t].
 \end{aligned}$$

□

Para concluir con esta sección en la que se presentaron la definición de la distribución exponencial y sus propiedades, tendremos una definición que habla sobre la suma de exponenciales.

**Definición 3.1.2.** Sea  $\lambda$  un número real positivo y  $n \in \mathbb{N}$ . Decimos que una variable aleatoria  $X$  tiene distribución Gamma con parámetros  $\lambda$  y  $n$ , es decir,  $X \sim \Gamma(n, \lambda)$  si su función de densidad está dada por:

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} \frac{(\lambda x)^{n-1}}{(n-1)!}, & \text{para } x \geq 0, \\ 0, & \text{c.o.c.} \end{cases}$$

**Proposición 3.1.3.** Sean  $\tau_1, \tau_2, \dots, \tau_n$  variables aleatorias independientes e idénticamente distribuidas exponencialmente con parámetro  $\lambda$ , entonces la suma  $T_n = \tau_1 + \tau_2 + \dots + \tau_n \sim \Gamma(n, \lambda)$ .

*Demostración.* La prueba se realizará por inducción en  $n$ . Para  $n = 1$

$$f_{T_1}(t) = \lambda e^{-\lambda t} \frac{(\lambda t)^{1-1}}{(1-1)!} = \lambda e^{-\lambda t}.$$

Por lo que para  $n = 1$  se cumple. Ahora, suponemos que es cierta para  $n$  y veamos que ocurre en  $n + 1$ . Notemos que  $T_{n+1} = T_n + \tau_{n+1}$ , entonces como para  $n$  se cumple, conocemos el valor de  $T_n$  y de acuerdo con la independencia de  $T_n$  y  $\tau_{n+1}$ , tenemos,

$$f_{T_{n+1}}(t) = \int_0^t f_{T_n}(s) f_{\tau_{n+1}}(t-s) ds.$$

Sustituyendo el valor del primer termino y el del segundo, tendremos:

$$\begin{aligned}
 &= \int_0^t \lambda e^{-\lambda s} \frac{(\lambda s)^{n-1}}{(n-1)!} \lambda e^{-\lambda(t-s)} ds \\
 &= e^{-\lambda t} \lambda^n \int_0^t \frac{s^{n-1}}{(n-1)!} ds \\
 &= \lambda e^{-\lambda t} \frac{\lambda^n t^n}{n!}.
 \end{aligned}$$

Con lo cual queda demostrado que la suma de v.a.i.i.d. exponencialmente, siguen una distribución gamma. □

Una distribución que será de utilidad debido a que la distribución exponencial y la distribución Erlang están relacionadas en el contexto de los procesos de Poisson, donde la distribución Erlang es una generalización que permite modelar el tiempo de espera para un número determinado de eventos, mientras que la distribución exponencial es un caso especial de la distribución Erlang cuando se espera un solo evento.

**Definición 3.1.3.** Sea  $X$  una Variable aleatoria, se dice que  $X$  tiene distribución de Erlang con parámetros  $(l, \mu)$  si,

$$f(x) = \frac{\mu(\mu x)^{l-1}}{(l-1)!} e^{-\mu x}, \quad x > 0. \tag{3.1}$$

Otra distribución de suma importancia para nuestro estudio es la distribución de Poisson la cual tiene una estrecha relación con la distribución exponencial.

**Definición 3.1.4.** Sea  $\lambda$  un número real positivo. Decimos que una variable aleatoria  $X$  tiene distribución de Poisson con parámetro  $\lambda$  ( $X \sim \text{Poisson}(\lambda)$ ) si su función de probabilidad está dada por:

$$f(x) = P[X = x] = \begin{cases} \frac{\lambda^x e^{-\lambda}}{x!}, & \text{si } x \geq 0, \\ 0, & \text{c.o.c.} \end{cases}$$

Notemos que la distribución exponencial y la distribución de Poisson se relacionan a través de esta fórmula, ya que el tiempo que transcurre entre dos eventos sucesivos sigue la distribución exponencial con un parámetro  $\lambda$ . Y recordemos que estos tiempos son independientes, lo que significa que el tiempo entre eventos no se ve afectado por los tiempos entre eventos anteriores. Si se cumplen estos supuestos, el número de eventos por unidad de tiempo sigue una distribución de Poisson con una media  $\lambda$ . Un ejemplo que podemos mencionar para que esta relación sea mas clara, es el numero de clientes que llegan a un banco en una hora estará modelado por una distribución Poisson, en cambio, el tiempo entre ocurrencias sigue la distribución exponencial.

## 3.2. Proceso de Poisson

Los procesos de Poisson y su distribución asociada tienen sus inicios en el trabajo del matemático francés Siméon Denis Poisson, en el siglo XIX, cuando Poisson se encontraba estudiando problemas relacionados con la teoría de la probabilidad y la estadística. Específicamente en 1837, publicó su trabajo *Recherches sur la probabilité des jugements en matière criminelle et en matière civile* ("Investigaciones sobre la probabilidad de juicios en asuntos criminales y civiles"), en este trabajo presentó la distribución de Poisson como una solución a ciertos problemas de conteo y probabilidad. Poisson realizó importantes contribuciones a diversas áreas de las matemáticas, la física y la ingeniería, pero su nombre se asocia principalmente con la distribución de Poisson y los procesos de Poisson.

Los procesos de Poisson son una parte fundamental en la teoría de probabilidad debido a su gran utilidad para modelar fenómenos aleatorios e independientes respecto a el tiempo o espacio. Específicamente, cuentan con aplicaciones en áreas como la teoría de colas, confiabilidad, la física de partículas, la biología, la economía y más. La siguiente definición se puede encontrar en [Durrett,2012].

**Definición 3.2.1. Proceso de Poisson.** Sean  $\tau_1, \tau_2, \tau_3, \dots$  v.a.i.i.d. las cuales cuentan con distribución exponencial con parámetro  $\lambda$  y sea  $T_n = \tau_1 + \tau_2 + \tau_3 + \dots + \tau_n$  para  $n \geq 1$  con  $T_0 = 0$ . Definimos el proceso de Poisson con parámetro o intensidad  $\lambda$  por,

$$N(s) = \max\{n : T_n \leq s\}$$

Veamos que las variables  $\tau_n$  representan los intervalos de tiempo entre eventos sucesivos por ejemplo, llegadas de pacientes a la sala de emergencia de un hospital y en consecuencia  $T_n = \tau_1 + \tau_2 + \tau_3 + \dots + \tau_n$  es el instante en el que ocurre el  $n$ -ésimo evento y  $N(s)$  es el número de eventos que han ocurrido hasta el instante  $s$ . Llamaremos tiempos de llegada del proceso a las variables  $\tau_n, n \geq 1$ .

Para ver por qué  $N(s), s \geq 0$ , recibe el nombre de proceso de Poisson, calculemos su distribución:  $N(s) = n$  si y sólo si  $T_n \leq s < T_{n+1}$ , es decir, el  $n$ -ésimo evento ocurre antes del instante  $s$  pero el  $(n + 1)$ -ésimo ocurre después de  $s$ .

**Proposición 3.2.1.**  $N(s)$  tiene distribución de Poisson con parámetro  $\lambda s$ .

*Demostración.* Usando la ley de la probabilidad total, condicionando respecto al instante en el cual ocurre  $T_n$ , obtendremos,

$$\begin{aligned} P(N(s) = n) &= P(T_{n+1} > s > T_n) \\ &= \int_0^s P(T_{n+1} > s | T_n = t) f_{T_n}(t) dt \\ &= \int_0^s P(\tau_{n+1} > s - t) f_{T_n}(t) dt \end{aligned}$$

Usando el resultado de la Proposición 3.1.3, tenemos,

$$\begin{aligned} &= \int_0^s \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} e^{-\lambda(s-t)} dt \\ &= \frac{\lambda^n}{(n-1)!} e^{-\lambda s} \int_0^s \tau^{n-1} dt \\ &= \frac{(\lambda s)^n}{n!} e^{-\lambda s}. \end{aligned}$$

□

Una vez que ya conocemos el proceso de Poisson, es necesario que veamos algunas propiedades de este mismo.

**Proposición 3.2.2.**  $N(t+s) - N(s), t \geq 0$  es un proceso de Poisson con parámetro  $\lambda$  y es independiente de  $N(r), 0 \leq r \leq s$ .

*Demostración.* Supongamos que  $N(s) = n$  y que el  $n$ -ésimo evento ocurrió en el instante  $T_n$ . Sabemos que el intervalo de tiempo para el siguiente evento debe satisfacer  $\tau_{n+1} > s - T_n$ , pero por la propiedad de falta de memoria de la distribución exponencial,

$$\begin{aligned} &P(\tau_{n+1} > s - T_n + t | \tau_{n+1} > s - T_n) \\ &= P(\tau_{n+1} > t) = e^{-\lambda t}. \end{aligned}$$

Lo cual demuestra que la distribución del tiempo de espera hasta el primer evento después de  $s$  es exponencial con parámetro  $\lambda$  y es independiente de  $\tau_i, 1 \leq i \leq n$ . Por otro lado,  $\tau_{n+1}, \tau_{n+2}, \dots$  son independientes de  $\tau_i, 1 \leq i \leq n$  y por lo tanto, también de  $T_i, 1 \leq i \leq n$ . Esto muestra que los intervalos entre eventos que ocurren después de  $s$  son v.a.i.i.d. con distribución exponencial con parámetro  $\lambda$ , y por lo tanto,  $N(t+s) - N(s)$  es un proceso de Poisson. □

Veamos el siguiente resultado como consecuencia de la proposición anterior.

**Proposición 3.2.3.**  $N(t)$  tiene incrementos independientes, es decir, si  $t_0 < t_1 < \dots < t_n$ , entonces,

$$N(t_1) - N(t_0), N(t_2) - N(t_1), \dots, N(t_n) - N(t_{n-1})$$

son independientes.

*Demostración.* Implica que  $N(t_n) - N(t_{n+1})$  es independiente de  $N(r), r \leq t_{n-1}$  y en consecuencia también de  $N(t_{n-1}) - N(t_{n-2}), \dots, N(t_1) - N(t_0)$ . El resultado sigue por inducción. Combinando los dos lemas anteriores tenemos la mitad del siguiente resultado, que es una caracterización fundamental del proceso de Poisson. □

Ahora veamos otra manera de definir el proceso de Poisson.

**Proposición 3.2.4.** *Segunda definición del proceso de Poisson.* Si  $N(s), s \geq 0$  es un proceso de Poisson con parámetro  $\lambda > 0$ , entonces;

1.  $N(0) = 0$ .
2.  $N(t+s) - N(s) \sim Pois(\lambda t)$ .
3.  $N(t)$  tiene incrementos independientes.

Recíprocamente, si 1, 2 y 3 se cumplen, entonces  $N(s), s \geq 0$  es un proceso de Poisson.

*Demostración.* Las Proposiciones 3.2.2 y 3.2.3 demuestran la primera afirmación. Para ver el recíproco, sea  $T_n$  el instante en el cual ocurre el  $n$ -ésimo evento. El primer evento ocurre después de  $t$  si y sólo si no ocurre ningún evento en  $[0, t]$ . Usando la fórmula para la distribución de Poisson,

$$P(T_1 > t) = P(N(t) = 0) = e^{-\lambda t}.$$

Lo cual muestra que

$$T_1 = T_1 \sim \text{Exp}(\lambda).$$

Para  $\tau_2 = T_2 - T_1$  observamos que

$$\begin{aligned} P(T_2 > t | T_1 = s) &= P(\text{no ocurre ningún evento en } (s, s+t] | T_1 = s) \\ &= P(N(t+s) - N(s) = 0 | N(r) = 0 \text{ para } r < s, N(s) = 1) \\ &= P(N(t+s) - N(s) = 0) = e^{-\lambda t}. \end{aligned}$$

Por la propiedad de incrementos independientes, de modo que  $T_2 \sim \text{Exp}(\lambda)$  y es independiente de  $T_1$ . Repitiendo este argumento vemos que  $T_1, T_2, \dots$  son i.i.d. con distribución exponencial de parámetro  $\lambda$ . □

### 3.3. Cadena de Markov a tiempo continuo

Los conceptos y definiciones que se abordaron en el Capítulo 2, son de suma importancia para poder abordar ciertas herramientas que son indispensables, sin embargo, las cadenas de Markov a tiempo discreto son solo una herramienta auxiliar en la teoría de colas, por lo que es necesario introducir las cadenas de Markov en tiempo continuo y con espacio de estados discreto ya que son el principal instrumento para el estudio de nuestro modelo de colas.

Una de las ventajas de las cadenas de Markov a tiempo continuo es su capacidad para modelar sistemas dinámicos y complejos, como la evolución de inventarios, el flujo de tráfico en redes de comunicación, la propagación de enfermedades, entre otros. Además, permiten calcular medidas de desempeño importantes, como el tiempo medio de permanencia en un estado, la probabilidad de estar en cierto estado en un momento dado, entre otras. Por lo que proporciona una visión profunda de la evolución y el comportamiento de estos sistemas en el tiempo. Las siguientes definiciones se pueden encontrar en [Peraza,2013].

**Definición 3.3.1.** Sea  $X = X(t); t \in \mathbb{R}$  Un proceso estocástico con espacio de estados discreto  $S$ ;  $X$  es una cadena de Markov en tiempo continuo si para cualquier entero  $n$  y tiempos  $t_1 < t_2 < \dots < t_n + 1$  en  $\mathbb{R}^+$ , se satisface que la distribución condicional de  $X(t_{n+1})$  dados los valores de  $X(t_1), X(t_2), \dots, X(t_n)$  solo depende de  $X(t_n)$ . Es decir, para cualesquiera  $x, x_1, \dots, x_n \in S$ ,

$$P(X(t_{n+1}) = x | X(t_1) = x_1, \dots, X(t_n) = x_n) = P(X(t_{n+1}) = x | X(t_n) = x_n).$$

Es decir, dado el presente estado, el resto del pasado es irrelevante para predecir el futuro.

**Definición 3.3.2.** Sean  $i, j \in S$  y  $0 \leq s \leq t$ . Definimos la probabilidad de transición como,

$$P_{ij}(s, t) = P(X(t) = j | X(s) = i).$$

Solo consideraremos procesos homogéneos en el tiempo, esto es, para  $i$  y  $j$  fijos, la probabilidad de transición  $P_{ij}(s, t)$  depende solo de  $t - s$ . Esto es,

$$P_{ij}(s, t) = P_{ij}(0, t - s).$$

Notemos que las probabilidades de transición solo dependen de una variable, por lo que sera equivalente a,

$$\begin{aligned} P_{ij}(t) &= P(X(t) = j | X(0) = i) \\ &= P(X(t+s) = j | X(s) = i), \forall s \geq 0. \end{aligned}$$

Al igual que en el caso discreto, se cuenta con la ecuación de Chapman-Kolmogorov.

**Proposición 3.3.1.** Sea  $X = X(t); t \geq 0$  una cadena de Markov con probabilidades de transición  $P_{ij}(t)$ : Entonces

$$P_{ij}(t+s) = \sum_{k=0}^{\infty} P_{ik}(s)P_{kj}(t) \quad \forall i, j \in S, t, s \geq 0.$$

*Demostración.* Recordemos que,

$$\begin{aligned}
 P_{ij}(t+s) &= P(X(t+s) = j | X(0) = i) \\
 &= \sum_{k=0}^{\infty} P(X(t+s) = j, X(s) = k | X(0) = i) \\
 &= \sum_{k=0}^{\infty} P(X(s) = k | X(0) = i) P(X(t+s) = j | X(0) = i, X(s) = k) \\
 &= \sum_{k=0}^{\infty} P_{ik}(s) P(X(t+s) = j | X(s) = k),
 \end{aligned}$$

por la propiedad de Markov a tiempo continuo y el hecho de que las probabilidades de transición dependen de una sola variable,

$$= \sum_{k=0}^{\infty} P_{ik}(s) P_{kj}(t).$$

□

**Proposición 3.3.2.** Un proceso de Poisson  $X = \{N(t); t \geq 0\}$  con parámetro  $\lambda > 0$  es una cadena de Markov homogénea a tiempo continuo.

*Demostración.* Veamos que,

$$\begin{aligned}
 P[N(t_{n+1}) = j | N(t_1) = x_1, \dots, N(t_n) = i] &= \frac{P[N(t_1) = x_1, \dots, N(t_n) = i, N(t_{n+1}) = j]}{P[N(t_1) = x_1, \dots, N(t_n) = i]} \\
 &= \frac{P[N(t_1) = x_1, N(t_2) - N(t_1) = x_2 - x_1, \dots, N(t_{n+1}) - N(t_n) = j - i]}{P[N(t_1) = x_1, N(t_2) - N(t_1) = x_2 - x_1, \dots, N(t_n) - N(t_{n-1}) = i - x_{n-1}]} \\
 &= \frac{P[N(t_1) = x_1] P[N(t_2) - N(t_1) = x_2 - x_1], \dots, P[N(t_n) - N(t_{n-1}) = i - x_{n-1}] P[N(t_{n+1}) - N(t_n) = j - i]}{P[N(t_1) = x_1] P[N(t_2) - N(t_1) = x_2 - x_1], \dots, P[N(t_n) - N(t_{n-1}) = i - x_{n-1}]} \\
 &= P[N(t_{n+1}) - N(t_n) = j - i].
 \end{aligned}$$

Ademas,

$$\begin{aligned}
 P[N(t_{n+1}) = j | N(t_n) = i] &= \frac{P[N(t_{n+1}) = j, N(t_n) = i]}{P[N(t_n) = i]} \\
 &= \frac{P[N(t_n) = i, N(t_{n+1}) - N(t_n) = j - i]}{P[N(t_n) = i]} \\
 &= \frac{P[N(t_n) = i] P[N(t_{n+1}) - N(t_n) = j - i]}{P[N(t_n) = i]} \\
 &= P[N(t_{n+1}) - N(t_n) = j - i].
 \end{aligned}$$

Entonces, X es una cadena de Markov homogénea, la cual cuenta con probabilidades de transición

$$\begin{aligned}
 P_{ij} &= P[N(t_{n+1}) - N(t_n) = j - i] \\
 &= \frac{e^{-\lambda t} (\lambda t)^{j-i}}{(j-i)!}, \text{ para todo } t, s \geq 0, j \geq i.
 \end{aligned}$$

□

Por ultimo tendremos la definición de *proceso de nacimiento y muerte*, estos procesos son modelos matemáticos utilizados para estudiar sistemas en los que los eventos de "nacimiento" (creación, entrada, llegada) y "muerte" (salida, destrucción) ocurren de manera aleatoria y siguen ciertas tasas de llegada y salida. Estos procesos son fundamentales en la teoría de colas, la teoría de probabilidades y otros campos de la ciencia y la ingeniería.

**Definición 3.3.3.** Un *proceso de nacimiento y muerte* representa el crecimiento o extinción de una población y formalmente, es una cadena de Markov homogénea a tiempo continuo  $X = X(t); t \geq 0$  con espacio de estados  $S = 0, 1, 2, \dots$  cuya intensidad de transición, o probabilidades infinitesimales son de la forma,

$$q_{ij} = \begin{cases} \lambda_i, & \text{si } j = i + 1 (i \geq 0), \\ \mu_i, & \text{si } j = i - 1 (i \geq 1), \\ -(\lambda_i - \mu_i), & \text{si } j = i (i \geq 0), \\ 0, & \text{si } |j - i| \geq 2. \end{cases} \quad (3.2)$$

Notemos la importancia de la relación

$$q_{ij} = 0, \text{ si } |j - i| \geq 2,$$

ya que nos dice que las probabilidades infinitesimales desde un estado dado  $i$  solo pueden ocurrir hacia sus vecinos inmediatos  $(i - 1, i + 1)$ . De aquí, a los parámetros  $\lambda_i = q_i, i + 1$  se les llama tasas (o intensidades) de nacimiento, mientras que a los parámetros  $\mu_i = q_i, i - 1$  se les llaman tasas (o intensidades) de muerte.

## Capítulo 4

# Indices de productividad y aplicación

Hasta este momento contamos con todas las herramientas necesarias para poder abordar el modelo de colas objetivo, además como hemos visto el proceso de Poisson es un modelo matemático que describe la ocurrencia de eventos que ocurren de forma aleatoria en el tiempo, pero con una tasa promedio constante. Por ejemplo, podríamos estar hablando de la cantidad de llamadas que llegan a un centro de atención al cliente en un período de tiempo dado. Este proceso se caracteriza por tener intervalos de tiempo entre eventos que siguen una distribución exponencial. En especial este proceso es una pieza clave para poder hablar de nuestro modelo de colas de interés, el modelo  $M/M/1/\infty$ , así que a continuación presentaremos tal modelo, sus parámetros, así como algunos ejemplos en donde este se puede presentar.

Iniciaremos el estudio de este modelo definiendo una cola o una línea de espera como un concepto utilizado para describir el fenómeno en el que los elementos llegan al servidor más rápido de lo que este los puede atender creando así una línea en la que se espera a ser atendido se presenta con frecuencia cuando se solicita un servicio por parte de una serie de clientes y tanto el servicio como los clientes son de tipo probabilístico. Los componentes que intervienen en un sistema de espera son los siguientes:

- **Fuente.** Población de clientes potenciales del sistema, la cual puede ser finita o infinita.
- **Procesos de arribos y de servicio.** Supondremos que los tiempos entre arribos de los clientes así como del tiempo de servicio son aleatorios.
- **Capacidad del sistema.** El número máximo de clientes en el sistema puede ser finito o infinito.
- **Número de servidores.** Puede considerarse una cantidad finita o infinita de servidores.
- **Disciplina de servicio.** Existen varias formas de seleccionar a los clientes que recibirán servicio las cuales determinan la disciplina de servicio. Entre las más comunes se encuentran:
  - **FIFO (first in first out):** Se atiende al cliente en el orden que llegan a la cola, el primero en llegar será el primero en ser atendido. En los modelos básicos de colas se supone como normal la disciplina de primero en entrar, primero en salir, a menos que se establezca de otra manera.
  - **LIFO (last in first out):** Consiste en atender primero al que ha llegado al último, también se le conoce como 'pila'.
  - **RSS (random selection of service):** Se selecciona a los clientes de una cola de forma aleatoria, con algún procedimiento de prioridad o algún otra preclasificación.
  - **Processor Sharing:** Todos los clientes experimentan con eficacia el mismo retraso, ya que comparten entre todos los clientes de la cola la capacidad del sistema atendiendo a todos por igual.
  - **PRI (Priority queue discipline):** En esta línea de espera cada miembro tiene prioridades determinadas por factores externos.

[Gnedenko,1968].

Una manera más simple de describir un sistema de colas fue propuesta en 1953 por el británico David G. Kendall en su obra "*Stochastic processes occurring in the theory of queues and their analysis by the method of imbedded Markov chains*" quien realizó importantes contribuciones a la teoría de colas en la segunda mitad del siglo XX.

En esta notación, un sistema de colas se representa con una notación alfanumérica que consta de,

$$A/S/c/K/F/d$$

En donde cada elemento representa:

- **A** = Distribución de tiempos entre llegadas. El cual a su vez se caracteriza de la siguiente manera:
  - **M**: Indica una llegada de clientes de acuerdo con un proceso de Poisson, es decir, llegadas de clientes independientes y aleatorias a una tasa constante.
  - **D**: Indica una llegada determinística de clientes, es decir, llegadas regulares y predecibles en intervalos de tiempo fijos.
  - **G**: Indica una llegada general de clientes, que puede seguir un patrón de llegada más complejo que no se ajusta a los modelos M o D estándar.
- **S** = Distribución del tiempo de servicio. Estas a su vez las podemos especificar como:
  - M : distribución exponencial
  - D : los tiempos entre arribos o de servicio son constantes o determinísticos.
  - G : distribución general.
- **c** = número de servidores.
- **K** = capacidad máxima del sistema.
- **d**= número de individuos en la fuente.

En este trabajo abordaremos el modelo de colas  $M/M/1/\infty$ , el cual por la notación de Kendall codifica un sistema unilineal con acumulador de capacidad infinita, un flujo de entrada que sigue el proceso de Poisson cuya intensidad sera denotada por  $\lambda$  y los tiempo de servicio son variables aleatorias independientes y distribuidas exponencialmente con parámetro  $\mu$ . Además, este modelo sigue una disciplina FIFO.

## 4.1. Indices de productividad del sistema

Denotaremos por  $v(t)$  al número de demandas que hay en el sistema al momento  $t$ , es indispensable notar que debido a que este sigue un proceso de Poisson, el proceso de las demandas que lleguen después del tiempo  $t$ , no dependen de su comportamiento antes de este. Por otro lado, debido a que los tiempos de servicio de las demandas se distribuyen exponencialmente, entonces también contarán con la Propiedad de pérdida de memoria 3.1.2 así entonces, el proceso  $\{v(t)\}_{t \geq 0}$  posee la Propiedad de Markov 2.1.1, por lo que será un proceso Markoviano a tiempo continuo y puesto que la cantidad de demandas en el sistema es numerable, el espacio de estados será discreto ( $x = \{0, 1, 2, \dots\}$ ).

### 4.1.1. Ecuaciones diferenciales de Kolmogorov

A partir de la información anteriormente dada nos es fácil notar que las tasas de llegada serán  $\lambda_n = \lambda$  para todo  $n = 0, 1, \dots$  y las tasas de servicio serán  $\mu_n = \mu$  para todo  $n = 1, \dots$ . Como primer paso se analizara el comportamiento de  $v(t)$  en el intervalo de tiempo infinitesimal  $(t, t + \Delta)$ , en donde  $\Delta$  es un incremento pequeño de tiempo, para esto se enunciarán las siguientes probabilidades de transición:

$$P_{00}(\Delta) = 1 - \lambda\Delta + o(\Delta), \quad i = 0$$

$$p_{ii}(\Delta) = 1 - (\lambda + \mu)\Delta + o(\Delta), \quad i > 0$$

$$p_{ii+1}(\Delta) = \lambda\Delta + o(\Delta), \quad i \geq 0$$

$$p_{i+1i}(\Delta) = \mu\Delta + o(\Delta), \quad i \geq 0$$

$$p_{ij}(\Delta) = o(\Delta), \quad |j - i| \geq 2.$$

Notemos que de acuerdo con la definición de proceso de nacimiento y muerte 3.3.3 el proceso  $\{v(t)\}_{t \geq 0}$  es de este tipo.

Para obtener las probabilidades estacionarias del sistema, denotemos por  $p_i(t) = P\{v(t) = i\}$  y partimos del teorema de la probabilidad total, considerando las distribuciones condicionales para los estados 0 e  $i$  para obtener las siguientes probabilidades estacionarias en el intervalo de tiempo considerado:

$$p_0(t + \Delta) = (1 - \lambda\Delta)p_0(t) + \mu\Delta p_1(t) + o(\Delta),$$

$$p_i(t + \Delta) = [1 - (\lambda + \mu)\Delta]p_i(t) + \lambda\Delta p_{i-1}(t) + \mu\Delta p_{i+1}(t) + o(\Delta), \quad i \geq 1.$$

A partir de estos resultados, se puede obtener el sistema de ecuaciones diferenciales de Kolmogórov de la siguiente manera:

$$p_0(t + \Delta) = p_0(t) - \lambda\Delta p_0(t) + \mu\Delta p_1(t) + 0(\Delta)$$

$$p_0(t + \Delta) - p_0(t) = -\lambda\Delta p_0(t) + \mu\Delta p_1(t) + 0(\Delta)$$

$$\frac{p_0(t + \Delta) - p_0(t)}{\Delta} = -\lambda p_0(t) + \mu p_1(t) + 0(\Delta)$$

$$\lim_{\Delta \rightarrow \infty} \frac{p_0(t + \Delta) - p_0(t)}{\Delta} = -\lambda p_0(t) + \mu p_1(t)$$

$$p'_0 = -\lambda p_0(t) + \mu p_1(t).$$

Sea  $i \geq 1$ ,

$$p_i(t + \Delta) = p_i(t) - \lambda\Delta p_i(t) - \mu\Delta p_i(t) + \lambda\Delta p_{i-1}(t) + \mu\Delta p_{i+1}(t) + 0(\Delta)$$

$$p_i(t + \Delta) - p_i(t) = -\lambda\Delta p_i(t) - \mu\Delta p_i(t) + \lambda\Delta p_{i-1}(t) + \mu\Delta p_{i+1}(t) + 0(\Delta)$$

$$\frac{p_i(t + \Delta) - p_i(t)}{\Delta} = -\lambda p_i(t) - \mu p_i(t) + \lambda p_{i-1}(t) + \mu p_{i+1}(t) + 0(\Delta)$$

$$\lim_{\Delta \rightarrow \infty} \frac{p_i(t + \Delta) - p_i(t)}{\Delta} = -(\lambda + \mu)p_i(t) + \lambda p_{i-1}(t) + \mu p_{i+1}(t)$$

$$p'_i = -(\lambda + \mu)p_i(t) + \lambda p_{i-1}(t) + \mu p_{i+1}(t)$$

Por lo que el sistema de ecuaciones de Kolmogorov será el siguiente:

$$p'_0 = -\lambda p_0(t) + \mu p_1(t),$$

$$p'_i = -(\lambda + \mu)p_i(t) + \lambda p_{i-1}(t) + \mu p_{i+1}(t).$$

Notemos que dado que el proceso no se interrumpe en un tiempo finito entonces tenemos un proceso que es regular. Además, por cómo definimos el proceso podemos notar que dados los estados  $i$  y  $j$  dentro del espacio de estados y un  $t > 0$ ,  $p_{ij}, t > 0$  y  $p_{ji}, t > 0$  esto nos dice que sus estados son comunicantes, lo cual implica que también tenemos un proceso irreducible.

### 4.1.2. Sistema de equilibrio

Un resultado interesante e importante para nuestro análisis es la siguiente definición así como las propiedades de el sistema las cuales se pueden encontrar en [Albores,1997].

**Definición 4.1.1.** *Teorema ergódico para un proceso de Márkov con espacio de estados numerable.* Sea  $\{V(t)\}_{t \geq 0}$  un proceso de Márkov irreducible regular con espacio de estados  $I = 1, 2, \dots$  numerable, entonces,

$$p_{ij}(t) \rightarrow_{t \rightarrow \infty} 0, \quad i, j \geq 1. \quad (4.1)$$

Por tanto, el proceso es ergódico.

El Teorema Ergódico establece que, bajo ciertas condiciones, el comportamiento temporal promedio de un sistema se aproxima al comportamiento promedio en el largo plazo. Esto significa que si observamos el sistema durante un período de tiempo suficientemente largo, el comportamiento promedio que observamos en ese período largo se acerca al comportamiento temporal promedio que vemos a medida que observamos el sistema durante más tiempo. En el contexto de la teoría de colas, esto implica que si un sistema de colas es ergódico, su comportamiento promedio a largo plazo es estable y predecible. Esto es importante para comprender y gestionar sistemas de colas en aplicaciones prácticas.

Podemos obtener el Sistema de Ecuaciones de Equilibrio igualando las derivadas a 0. Puesto que  $V(t)_{t \geq 0}$  es un proceso ergódico, con el paso del tiempo, el funcionamiento del sistema de colas tenderá al régimen estacionario, de este modo, tendremos que las probabilidades estacionarias satisfacen el sistema de ecuaciones de equilibrio, que se obtiene de igualar las derivadas a cero en el sistema de ecuaciones diferenciales de Kolmogórov.

$$\begin{aligned} 0 &= -\lambda p_0(t) + \mu p_1(t), \\ 0 &= -(\lambda + \mu)p_i(t) + \lambda p_{i-1}(t) + \mu p_{i+1}(t), \quad i \geq 1. \end{aligned}$$

Para resolver este sistema, empleamos la ecuación de balance local, que muestra la igualdad de los flujos de probabilidades de encuentro entre los estados  $i, i + 1$ . Partiendo del sistema anterior, tendremos que:

$$\begin{aligned} \lambda p_i &= \mu p_{i+1}, \quad i \geq 0, \\ p_i &= \frac{\lambda}{\mu} p_{i-1}, \quad i \geq 1, \\ p_i &= p_0 \rho^i, \quad i \geq 0. \end{aligned}$$

donde,

$$\rho = \frac{\lambda}{\mu} = \frac{\text{Promedio de demandas que llegan por unidad de tiempo}}{\text{promedio de demandas que se pueden atender por unidad de tiempo}}.$$

A  $\rho$  se le conoce como coeficiente de carga del sistema, para que el sistema este en equilibrio, este coeficiente debe ser menor a uno, esto nos dice que las demandas que lleguen podrán ser atendidas y en caso contrario, el servidor estará desbordado por lo que no existirá el equilibrio de modo que para modelar los índices de productividad deberemos recurrir al sistema de ecuaciones diferenciales de Kolmogorov. Considerando que el sistema de colas es equilibrado, podemos calcular las siguientes expresiones para los índices de productividad.

### 4.1.3. Valor medio estacionario del número de demandas en el sistema de colas

$$N = \sum_{i=0}^{\infty} i p_i = \sum_{i=0}^{\infty} i \rho^i (1 - \rho) = (1 - \rho) \sum_{i=0}^{\infty} i \rho^i.$$

Sumando la serie, obtendremos;

$$\begin{aligned} S &= \rho + 2\rho^2 + 3\rho^3 + 4\rho^4 + \dots \\ -\rho S &= -\rho^2 - 2\rho^3 - 3\rho^4 + \dots \end{aligned}$$

$$(1 - \rho)S = \rho^2 + \rho^3 + \rho^4 + \dots = \frac{\rho}{1 - \rho}.$$

Así tendremos,

$$N = (1 - \rho) \frac{\rho}{(1 - \rho)^2} = \frac{\rho}{1 - \rho}.$$

#### 4.1.4. Longitud media estacionaria de la cola

$$Q = \sum_{i=1}^{\infty} (i - 1)p_i = (1 - \rho) \sum_{i=1}^{\infty} (i - 1)\rho^i.$$

Veamos que podemos separar esta serie en dos mas sencillas;

$$\sum_{i=1}^{\infty} (i - 1)\rho^i = \sum_{i=1}^{\infty} i\rho^i + \sum_{i=1}^{\infty} -\rho^i.$$

Notemos que estas series se resolvieron en el valor medio estacionario del número de demandas en el sistema de colas, con lo que tendremos;

$$\begin{aligned} \sum_{i=1}^{\infty} (i - 1)\rho^i &= \sum_{i=1}^{\infty} i\rho^i - \sum_{i=1}^{\infty} \rho^i \\ &= \frac{\rho}{(1 - \rho)^2} - \frac{\rho}{1 - \rho} = \frac{\rho^2}{(1 - \rho)^2}. \end{aligned}$$

Por lo tanto,

$$Q = (1 - \rho) \frac{\rho^2}{(1 - \rho)^2} = \frac{\rho^2}{1 - \rho}.$$

#### 4.1.5. Distribución estacionaria del tiempo de espera de inicio del servicio

Para determinar la FD del tiempo de espera de inicio del servicio en el régimen estacionario ( $W(x)$ ), suponemos primero que la demanda que llega al sistema se encuentra en el  $i$ -ésimo lugar. Como sabemos por la descripción del sistema cada demanda que se encuentra en el sistema, se atiende en un tiempo distribuido exponencialmente con parámetro  $\mu$ , el tiempo de espera total de inicio del servicio, para una demanda que al llegar encuentra  $i$  demandas, va a tener la distribución de Erlang  $E_i(x)$  con parámetros  $\mu$  e  $i$ . Es claro que si,  $i = 0$ , la demanda al llegar encuentra el sistema vacío e inicia su servicio inmediatamente y, por lo que su tiempo de espera es igual a 0. Ahora entonces la probabilidad de que en el régimen estacionario una demanda que llega al sistema encuentre  $i$  demandas es igual a  $p_i$ . Usaremos la fórmula de la probabilidad total, con lo que obtendremos lo siguiente,

$$W(x) = \sum_{i=0}^{\infty} E_i(x)p_i = (1 - \rho) \sum_{i=0}^{\infty} \rho^i E_i(x).$$

Utilizando la transformada de Laplace-Stieltjes (TLS), y como la TLS de la distribución de Erlang es  $\epsilon_i(s) = \left(\frac{\mu}{s + \mu}\right)^i$  y  $\epsilon_0 = 1$ , tenemos,

$$\begin{aligned} \omega(s) &= \int_0^{\infty} e^{-sx} dW(x), \\ \sum_{i=0}^{\infty} \epsilon_i(s)\rho^i &= (1 - \rho) \sum_{i=0}^{\infty} \frac{(\mu\rho)^i}{(s + \mu)^i} \\ &= 1 - \rho + \rho \frac{\mu(1 - \rho)}{s + \mu(1 - \rho)}. \end{aligned}$$

Pasando de la TLS a la función de distribución tendremos,

$$W(x) = 1 - \rho + \rho[1 - e^{-\mu(1 - \rho)x}], x > 0. \quad (4.2)$$

#### 4.1.6. Distribución estacionaria del tiempo de estancia de una demanda en el sistema

Como el tiempo de servicio de una demanda se distribuye exponencialmente con parámetro  $\mu$ , la distribución estacionaria del tiempo de estancia de una demanda en el sistema ( $V(x)$ ) tiene la siguiente TLS,

$$\begin{aligned}\varphi(s) &= \int_0^{\infty} e^{-sx} dV(x) \\ &= \frac{\mu}{s + \mu} \omega(s) = \left( \frac{\mu}{s + \mu} \right) (1 - \rho + \rho \frac{\mu(1 - \rho)}{s + \mu(1 - \rho)}) \\ &= \frac{\mu(1 - \rho)}{s + \mu(1 - \rho)}.\end{aligned}$$

Pasando de la TLS a la función de distribución tendremos,

$$V(x) = 1 - e^{-(\mu - \lambda)x}. \quad (4.3)$$

#### 4.1.7. Tiempo medio estacionario de espera de inicio del servicio

Los tiempos medios estacionarios de inicio del servicio están dados por la formula,

$$\begin{aligned}w &= \int_0^{\infty} x dW(x) \\ &= -\omega'(0) = \frac{\rho}{\mu(1 - \rho)}.\end{aligned}$$

Entonces,

$$w = \frac{\rho}{\mu(1 - \rho)}.$$

#### 4.1.8. Tiempo medio de estancia en el sistema

El tiempo medio de estancia de una demanda en el sistema esta dado por la formula.

$$\begin{aligned}v &= \int_0^{\infty} x dV(x) \\ &= -\varphi'(0) = w + \frac{1}{\mu} \\ &= \frac{1}{\mu(1 - \rho)}.\end{aligned}$$

Entonces,

$$v = \frac{1}{\mu(1 - \rho)}.$$

#### 4.1.9. Formulas de Little

Las fórmulas de Little son un conjunto de relaciones matemáticas fundamentales en la teoría de colas que fueron desarrolladas por John D. C. Little en la década de 1960. Estas fórmulas establecen relaciones importantes entre tres conceptos clave en sistemas de colas. Las fórmulas de Little son especialmente útiles para comprender y analizar el comportamiento de sistemas de colas en diversas aplicaciones. El camino para obtener estas ecuaciones es intuitivo a partir de los elementos que tene-

mos, supongamos que el sistema está funcionando en un régimen estacionario durante un intervalo de tiempo  $T$  suficientemente grande. Entonces, durante este tiempo, llegan al sistema en promedio  $\lambda T$  demandas. A su vez, cada demanda se encuentra en el sistema en promedio  $v$  unidades de tiempo, por lo tanto, el tiempo medio  $V$  que pasan en el sistema las  $\lambda T$  demandas es igual a  $\lambda v T$  y, en cada instante en el sistema se encuentran  $N = \frac{V}{T} = \lambda v$  demandas.

$$N = \lambda v. \quad (4.4)$$

Y la igualdad análoga es,

$$Q = \lambda w. \quad (4.5)$$

La última característica de interés en nuestro modelo es el flujo de salida. Esto es, el flujo de demandas servidas que abandonan el sistema. Dado que  $V(t)_{t \geq 0}$  es proceso de nacimientos y muertes, se sigue que el flujo de salida del sistema  $M/M/1/\infty$  funcionando en equilibrio será de Poisson con intensidad  $\lambda$ .

## 4.2. Aplicación del modelo

Por ultimo, se realizara la aplicación de este modelo en un caso practico para el cual se hará uso de los datos obtenidos a través de la observación de 60 ciclos de servicio en el semáforo ubicado en la calle 16 de septiembre en esquina con la calle 15 oriente frente al parque del Carmen. Esta calle es una calle unicarril hace poco mas de un año debido a reestructuraciones viales, debido a que es un punto muy cercano al zócalo de la ciudad de Puebla, además de estar rodeada de diversas escuelas es una calle con mucha afluencia vehicular a pesar de ser muy pequeña. En la siguiente imagen se muestra la circulación vial de la calle ya mencionada.



Figura 4.1: Mapa de las calles.

Los datos obtenidos a partir de la observación del funcionamiento de dicho semáforo durante una hora, se presentan en la primera tabla. Mientras que en la segunda tabla se muestra los tiempos observados para cada fase del semaforo, además se pudo observar que durante la fase amarilla los conductores seguían avanzando, por lo que se podrían reducir simplemente a dos fases.

N° de Observación	Vehiculos en espera	Vehiculos servidos	Total de llegadas	N° de Observación	Vehiculos en espera	Vehiculos servidos
1	8	7	401	31	11	5
2	7	5		32	10	7
3	9	6		33	8	3
4	1	1		34	12	6
5	6	3		35	9	8
6	9	7		36	9	9
7	10	5		37	8	4
8	9	6		38	11	9
9	8	5		39	11	8
10	12	8		40	10	7
11	7	3		41	12	9
12	10	7		42	11	9
13	11	7		43	7	7
14	7	6		44	8	5
15	10	9		45	11	7
16	8	4		46	7	3
17	8	6		47	13	10
18	10	9		48	7	3
19	8	7		49	7	4
20	11	9		50	11	9
21	12	9		51	12	10
22	11	7		52	10	8
23	13	8		53	8	6
24	7	4		54	9	7
25	9	9		55	11	10
26	8	7		56	8	7
27	9	7		57	11	7
28	10	8		58	11	8
29	7	5		59	10	7
30	10	8		60	7	6

Figura 4.2: Tabla con los datos obtenidos.

	Segundos
Tiempo en Rojo	25
Tiempo en Amarillo	5
Tiempo en Verde	30
Tiempo de Circulación	35

Figura 4.3: Duración de las fases del semáforo.

### 4.2.1. Cálculo de los índices de productividad

Debido a las características de nuestro sistema y a la observación de la cantidad promedio de demandas que podían servirse por ciclo de servicio, los tiempos de servicio son variables aleatorias independientes idénticamente distribuidas exponenciales con parámetro  $\mu = 9,25$  ya que un auto tarda aproximadamente 4.59 segundos en atravesar el semáforo. Por otro lado, para calcular la intensidad del proceso

de arribos, calculamos el número promedio de demandas que llegan por unidad de tiempo, es decir, será un proceso de Poisson con  $\lambda = 7,61$ , como consecuencia se tiene que el coeficiente de carga será el cociente de ambos parámetros i.e.,  $\rho = \frac{\lambda}{\mu} = 0,823423423$ . A partir de estos datos es posible calcular los

índices de productividad del sistema. Iniciaremos con el cálculo de las **probabilidades estacionarias de que el Sistema de colas se encuentre vacío en un instante cualquiera**

$$p_0 = 1 - \rho = 1 - 0,823423423 = 0,176576577.$$

**Probabilidades estacionarias de que el sistema de colas tenga  $i$  demandas en cualquier instante.**

$$p_i = (1 - \rho)\rho^i = (1 - 0,823423423)(0,823423423)^i$$

$$p_i = 0,176576577(0,823423423)^i.$$

Conforme a las expresiones obtenidas, podemos calcular también el **valor medio estacionario del número de demandas en el sistema de colas** y el resto de índices de productividad:

$$N = \frac{\rho}{1 - \rho} = \frac{0,823423423}{(1 - 0,823423423)} = 4,663265306.$$

**Longitud media estacionaria de la cola**

$$Q = \frac{\rho^2}{1 - \rho} = \frac{(0,823423423)^2}{(1 - 0,823423423)} = 3,839841883.$$

**Distribución estacionaria del tiempo de espera de inicio del servicio.**

$$w(x) = 1 - \rho + \rho[1 - e^{-\mu(1-\rho)x}] = 1 - 0,823423423 + 0,823423423[1 - e^{-9,25(1-0,823423423)x}]$$

$$w(x) = 0,999993423 - 0,823423423e^{-(1,63334)x}.$$

**Distribución estacionaria del tiempo de estancia de una demanda en el sistema.**

$$v(x) = 1 - e^{-(\mu - \lambda)x} = 1 - e^{-(9,25 - 7,616666667)x} = 1 - e^{-(1,633333)x}.$$

**Tiempo medio estacionario de espera de inicio del servicio.**

$$w = \frac{\rho}{(\mu(1 - \rho))} = \frac{0,823423423}{9,25(1 - 0,823423423)} = 0,5041367883.$$

**Tiempo de estancia en el sistema.**

$$v = \frac{1}{\mu(1 - \rho)} = \frac{1}{9,25(1 - 0,823423423)} = 0,6122448965.$$

**Fórmulas de Little.**

$$N = \lambda v = (7,616666667)0,6122448965 = 4,663265306.$$

$$Q = \lambda w = (7,616666667)0,5041367883 = 3,839841883.$$

Notemos que el coeficiente de carga es muy cercano a 1 en la observación realizada, esto quiere decir que, al presentarse retrasos en los tiempos de servicio, la cola corre el riesgo de desbordarse.

### 4.2.2. Cálculo de los índices de productividad con $\mu = 9,75$

Para conocer mejor el comportamiento de nuestro modelo, veamos que ocurre si en vez de servir 9.25 demandas por unidad de tiempo, el servidor pudiera atender 9.75, es decir, que los conductores cruzan ligeramente más rápido el semáforo (3.58 segundos). Como ya sabemos debido a las características del modelo, los tiempos de servicio son variables aleatorias independientes idénticamente distribuidas exponenciales con parámetro  $\mu = 9,75$  y  $\lambda = 7,616666667$  y el coeficiente de carga para el sistema de colas será de  $\rho = \frac{\lambda}{\mu} = 0,781196581$ .

Como en el caso anterior, procedemos a calcular los índices de productividad del sistema.

**Probabilidades estacionarias de que el Sistema de colas se encuentre vacío en un instante cualquiera.**

$$p_0 = 1 - \rho = 1 - 0,781196581 = 0,218803419$$

**Probabilidades estacionarias de que el sistema de colas tenga  $i$  demandas en cualquier instante.**

$$p_i = (1 - \rho)\rho^i = (1 - 0,781196581)(0,781196581)^i.$$

$$p_i = 0,218803419(0,781196581)^i.$$

Conforme a las expresiones obtenidas, podemos calcular también el **valor medio estacionario del número de demandas en el sistema de colas** y el resto de índices de productividad:

$$N = \frac{\rho}{1 - \rho} = \frac{0,781196581}{(1 - 0,781196581)} = 3,5703125.$$

**Longitud media estacionaria de la cola**

$$Q = \frac{\rho^2}{1 - \rho} = \frac{(0,781196581)^2}{(1 - 0,781196581)} = 2,789115919.$$

**Distribución estacionaria del tiempo de espera de inicio del servicio.**

$$w(x) = 1 - \rho + \rho[1 - e^{-\mu(1-\rho)x}] = 1 - 0,781196581 + 0,781196581[1 - e^{-9,75(1-0,781196581)x}]$$

$$[w(x) = 1 - 0,781196581e^{-(2,1333)x}.$$

**Distribución estacionaria del tiempo de estancia de una demanda en el sistema.**

$$v(x) = 1 - e^{-(\mu - \lambda)x} = 1 - e^{-(9,75 - 7,616666667)x} = 1 - e^{-(2,13333)x}.$$

**Tiempo medio estacionario de espera de inicio del servicio.**

$$w = \frac{\rho}{(\mu(1 - \rho))} = \frac{0,781196581}{9,75(1 - 0,781196581)} = 0,366185897.$$

**Tiempo de estancia en el sistema.**

$$v = \frac{1}{\mu(1 - \rho)} = \frac{1}{9,75(1 - 0,781196581)} = 0,46875.$$

**Fórmulas de Little.**

$$N = \lambda v = (7,616666667)0,46875 = 3,5703125.$$

$$Q = \lambda w = (7,616666667)0,366185897 = 2,789115919.$$

Como se menciona al inicio para este caso particular proponemos que los conductores tengan una reacción más rápida cuando es su turno de avanzar, apelando así a su educación vial. Al realizar este pequeño cambio podemos observar que el  $\rho$  disminuye considerablemente, alejándose de la posibilidad de que la cola se desborde, con lo que se puede deducir que el sistema es muy sensible a cambios.

### 4.2.3. Cálculo de los índices de productividad con una fase de avance mas larga

En busca de poder encontrar un método para que nuestro sistema no corra el riesgo de desbordarse, se propone un último arreglo, en el que se modifique el tiempo de la luz verde y se le agreguen 5 segundos más para que los vehículos sigan fluyendo. Con este ajuste y debido a la naturaleza de nuestro sistema tendríamos que los tiempos serían exponenciales independientes e idénticamente distribuidos con parámetro  $\mu = 10,09$ . El promedio de visitas sigue siendo el mismo,  $\lambda = 7,3666$ . Y el cociente de ambos parámetros es  $\rho = 0,754804805$ . Como en los casos anteriores, procedemos a calcular los índices de productividad del sistema.

**Probabilidades estacionarias de que el Sistema de colas se encuentre vacío en un instante cualquiera.**

$$p_0 = 1 - \rho = 1 - 0,754804805 = 0,245195195$$

**Probabilidades estacionarias de que el sistema de colas tenga  $i$  demandas en cualquier instante.**

$$p_i = (1 - \rho)\rho^i = (1 - 0,754804805)(0,754804805)^i.$$

$$p_i = 0,245195195(0,754804805)^i.$$

Conforme a las expresiones obtenidas, podemos calcular también el **valor medio estacionario del número de demandas en el sistema de colas** y el resto de índices de productividad:

$$N = \frac{\rho}{1 - \rho} = \frac{0,754804805}{(1 - 0,754804805)} = 3,078383344.$$

**Longitud media estacionaria de la cola**

$$Q = \frac{\rho^2}{1 - \rho} = \frac{(0,754804805)^2}{(1 - 0,754804805)} = 2,323578539.$$

**Distribución estacionaria del tiempo de espera de inicio del servicio.**

$$w(x) = 1 - \rho + \rho[1 - e^{-\mu(1-\rho)x}] = 1 - 0,754804805 + 0,754804805[1 - e^{-10,09(1-0,754804805)x}]$$

$$w(x) = 1 - 0,754804805 - e^{-(2,47401)x}.$$

**Distribución estacionaria del tiempo de estancia de una demanda en el sistema.**

$$v(x) = 1 - e^{-(\mu - \lambda)x} = 1 - e^{-(10,09 - 7,616666667)x} = 1 - e^{-(2,47333)x}.$$

**Tiempo medio estacionario de espera de inicio del servicio.**

$$w = \frac{\rho}{(\mu(1 - \rho))} = \frac{0,754804805}{10,09(1 - 0,754804805)} = 0,305065016.$$

**Tiempo de estancia en el sistema.**

$$v = \frac{1}{\mu(1 - \rho)} = \frac{1}{10,09(1 - 0,754804805)} = 0,404164115.$$

**Fórmulas de Little.**

$$N = \lambda v = (7,616666667)0,404164115 = 3,078383344.$$

$$Q = \lambda w = (7,616666667)0,305065016 = 2,323578539.$$

En este supuesto si cambiamos los tiempos de los ciclos del semáforo. Al realizar este pequeño cambio podemos observar que el  $\rho$  disminuye considerablemente respecto a nuestros primeros resultados, alejándose así de la posibilidad de que la cola se desborde, con lo que también aumenta la posibilidad de que la cola se quede vacía, disminuyendo así el valor medio estacionario del número de demandas, la longitud media estacionaria de la cola, tiempo medio estacionario de inicio del servicio y tiempos de espera. Es decir, el sistema es muy sensible a cambios en los tiempos de servicio.

Para poder analizar las 3 propuestas juntas y así poder ver sus diferencias de una manera mas simple, se presenta un gráfico comparativo.

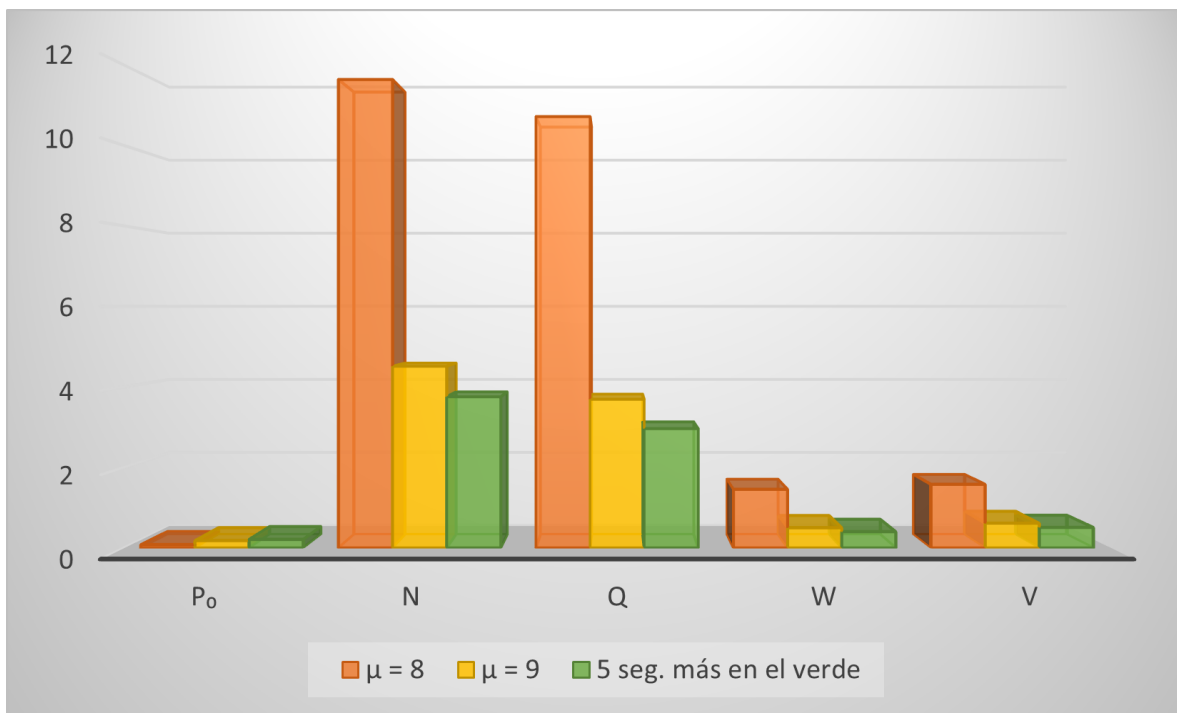


Figura 4.4: Comparación de resultados.

#### 4.2.4. Conclusiones de la aplicación

Tras analizar la teoría correspondiente y aplicarla a la observación realizada notamos que el sistema es equilibrado, aunque con un coeficiente de carga muy cercano a 1. Esto nos indica que con facilidad ante retrasos en los tiempos de servicio la cola podrá desbordarse. Este hecho es algo que nos preocupa, ya que al desbordarse la cola Significa que el sistema es ineficiente, específicamente en nuestra aplicación esto resultara en la creación de caos vehicular, lo cual incrementa la probabilidad de que algún accidente se presente. Por lo anterior, se repitió la prueba suponiendo ahora que el servidor es capaz de atender

9.75 demandas por unidad de tiempo. Al realizarlo obtuvimos que la carga del sistema disminuye seriamente, así como los índices de productividad también disminuyeron, lo cual nos dice que el sistema es sensible a los cambios. Como adicional repetimos la prueba, pero ahora cambiando los tiempos de servicio, obteniendo resultados aun mas bajos que en el segundo supuesto, llegando a una optimización bastante considerable puesto que con este pequeño cambio se reducen tiempos de espera, longitud de la cola y otros índices de productividad que nos llevan a reducir gastos en el combustible de los conductores, además, de agilizar el tránsito evitando congestionamientos y accidentes en la cola.

Aunque el tercer supuesto tiene índices de productividad mas bajos, no hace un cambio tan drástico como lo hace el segundo modelo con respecto al primero que es como actualmente funciona el sistema, por lo que nos refleja que si los conductores eliminaran distracciones y mejoraran con respecto a su educación vial el sistema funcionaria de la mejor manera evitando así el congestionamiento de esta calle.

# Capítulo 5

## Conclusión

Este trabajo se basó en el desarrollo y análisis del modelo de colas  $M/M/1/\infty$  el cual describe un modelo con un solo servidor, demandas que siguen un proceso de Poisson, tiempo de servicio exponenciales y la cola cuenta con capacidad infinita, así como los parámetros y propiedades matemáticas del mismo para posteriormente ver de una manera mas clara todos estos conceptos a través de una aplicación del modelo. Es imprescindible mencionar que los índices de productividad del modelo, como lo notamos en el desarrollo del modelo, nos describen el estado o evolución del sistema.

De manera resumida podemos notar que a través de la tasa de llegada ( $\lambda$ ) y la tasa de servicio ( $\mu$ ) se puede determinar la eficiencia del sistema, específicamente mediante la carga del sistema ( $\rho = \frac{\lambda}{\mu}$ ) la cual nos da una medida directa de la eficiencia del sistema y su estabilidad.

También hemos examinado las distribuciones de probabilidad de los tiempos de espera y el número de clientes en el sistema, obteniendo a través de este análisis un marco para predecir el comportamiento del sistema bajo diversas condiciones.

Es importante notar que el modelo  $M/M/1/\infty$  es uno de los modelos fundamentales en la teoría de colas debido a su simplicidad matemática y la obtención de un análisis exacto en su aplicación, esto es gracias a que al basarse en distribuciones Poisson para arribos y en exponenciales para los inter-arribos, el modelo permite obtener expresiones matemáticas para métricas clave como el tiempo medio en el sistema, el número esperado de clientes y la probabilidad de que la cola esté vacía. Además, de su gran utilidad para evaluar el impacto de diferentes tasas de llegada y servicio en la congestión del sistema ayudándonos así en la asignación eficiente de recursos brindándonos una optimización de los mismos. Adicional a esto es importante mencionar que su formulación permite modelar fenómenos estocásticos en diversas disciplinas, como telecomunicaciones, informática y logística, ya que cuenta con capacidad para aproximar una variedad de sistemas reales.

Aunque el modelo que analizamos en este trabajo es uno de los más básicos y sencillos dentro de la teoría, su estudio ha demostrado una complejidad considerable. Esta dificultad se incrementa, en gran medida, debido a que no tuve la oportunidad de cursar las asignaturas directamente relacionadas con este tema durante mi formación académica. Esto me obligó a dedicar un esfuerzo adicional para comprender los conceptos fundamentales, lo que resalta la complejidad del modelo, incluso en sus formas más simplificadas.

A pesar de que el modelo  $M/M/1/\infty$  es sumamente útil, suponer llegadas y tiempos de servicio exponenciales y una capacidad infinita puede ser una limitante para modelar algunos fenómenos o variantes en algunos contextos, por lo cual proponemos la presente tesis como una base para futuros trabajos que pueden enfocarse en extensiones del modelo para incluir distribuciones de tiempo de llegada y servicio más complejas, análisis de sistemas con múltiples servidores ( $M/M/c$ ) y colas con capacidad limitada. Específicamente hablando sobre el caso tratado en el trabajo, se sugiere como base para una línea de futuras investigaciones para el análisis del flujo y optimización del mismo en calles con mas de un carril o incluso la modelación matemática de intersecciones tales como boulevares de la ciudad de Puebla en donde se presenta una gran trafico de automóviles.



# Anexos

## Anexo: Códigos de Python

A continuación se muestran los códigos en lenguaje Python utilizado para la creación de las gráficas mostradas en los capítulos 1 y 2.

Código correspondiente a la gráfica de la función de distribución presentada en el capítulo 1.

```
import numpy as np
import matplotlib.pyplot as plt

# Definir la función de distribución acumulativa (CDF)
def cdf(x):
    return np.where(x < 2, 0, 1 - 4 / x**2)

# Generar datos para el eje x
x_values = np.linspace(1, 5, 1000)

# Calcular los valores de la función de distribución acumulativa (CDF) para cada punto en x
cdf_values = cdf(x_values)

# Crear la gráfica
# Use raw strings (r'') for the LaTeX code and escape backslashes with double backslashes (\\\\)
plt.plot(x_values, cdf_values, label='F(x)')

# Configurar etiquetas y leyenda
plt.xlabel('$x$')
plt.ylabel('$F(x)$')
plt.title('Función de Distribución Acumulativa (CDF)')
plt.legend()

# Mostrar la gráfica
plt.show()
```

Código correspondiente a la gráfica de la función de densidad presentada en el capítulo 1.

```
import numpy as np
import matplotlib.pyplot as plt

# Definir la función de densidad de probabilidad (PDF)
def pdf(y):
    return np.where((2 < y) & (y < 4), 1/8 * (y + 1), 0)

# Generar datos para el eje y
y_values = np.linspace(0, 5, 1000)

# Calcular los valores de la función de densidad de probabilidad (PDF) para cada punto en y
pdf_values = pdf(y_values)

# Crear la gráfica
plt.plot(y_values, pdf_values, label='f(y)')

# Configurar etiquetas y leyenda
plt.xlabel('$y$')
plt.ylabel('$f(y)$')
plt.title('Función de Densidad de Probabilidad (PDF)')
plt.legend()

# Mostrar la gráfica

plt.show()
```

---

Código correspondiente a la gráfica de la Probabilidad de transición en  $n$  pasos presentada en el capítulo 2.

```
import matplotlib.pyplot as plt

# Coordenadas de los tres puntos
x = [1, 2, 3]
y = [3, 1, 2]
z = [1, 2, 3]
w = [3, 2, 2]
i = [1, 2, 3]
j = [3, 3, 2]
m = [1, 2, 3]
a = [3, 4, 2]
d = [1, 2, 3]
s = [4, 4, 4]
labels = ['x', ' ', 'y']
lbl= ['Inicio ', 'n','n+m']
# Graficar los puntos y conectarlos con líneas
plt.plot(x, y, marker='o')
plt.plot(z, w, marker='o')
plt.plot(i, j, marker='o')
plt.plot(m, a, marker='o')
# Agregar etiquetas a los puntos
for i, label in enumerate(labels):
    plt.text(x[i], y[i], label, fontsize=14, ha='center', va='bottom')

for i, label in enumerate(lbl):
    plt.text(d[i], s[i], label, fontsize=16, ha='center', va='bottom')
# Etiquetas y título del gráfico
plt.axis('off')
plt.title('Chapman-Kolmogorov')

# Mostrar el gráfico
plt.show()
```

# Bibliografía

- [Albores,1997] Albores F.J., Bocharov P.P., Pechinkin A.V.(1997) *Teoria de Colas*. Universidad Autónoma de Tlaxcala.
- [Beichelt,2016] Beichelt Frank.(2016) *Applied Probability and Stochastic Processes*. Taylor & Francis Group.
- [Durrett,2012] Durrett Richard.(2012) *Essentials of stochastic processes*. Springer.
- [Gnedenko,1968] Gnedenko B.V., Kovalenko I.N.(1968) *Introduction to queueing theory*.Israel Program for Scientific Translations.
- [Hoel,1972] Hoel G. Paul, Port C. Sidney, Stone J. Charles.(1972) *Introduction to stochastic processes*.HOUGHTON MIFFLIN COMPANY.
- [Meyer,1973] Meyer L. Paul.(1973) *Introductory Probability and Statistical Applications*.Addison-Wesley Publishing Company.
- [OpenStax,1999] OpenStax.(1999) *La distribución exponencial*. Último acceso: 15 de Febrero de 2025. URL: <https://openstax.org/books/introducci%C3%B3n-estad%C3%ADstica-empresarial/pages/5-3-la-distribucion-exponencial>
- [Rincon,2012] Rincón Luis.(2012) *Introducción a los procesos estocásticos*. Departamento de Matemáticas, Facultad de Ciencias UNAM.
- [Ross,2019] Ross M. Sheldon.(2019) *Introduction to Probability Models*. Elsevier.
- [seactuario,2023] seactuario.(2023) *FUNCIÓN DE DISTRIBUCIÓN ACUMULATIVA*. Último acceso: 10 de Diciembre de 2024. URL: <https://seactuario.com/ContMatematicas/ProbEstadistica/PROBABILIDAD12.html>
- [Peraza,2013] Siqueiros Peraza Gerardo Fabian.(2013) *INTRODUCCIÓN A LA TEORÍA DE COLAS Y SU SIMULACIÓN*. División de Ciencias Exactas y Naturales, Departamento de Matemáticas, Universidad de Sonora.
- [UDP,2023] Universidad de las palmas de gran canaria. Variables aleatorias.(2023) *Estadística y Procesos Estocásticos Tema 2: Variables Aleatorias* Último acceso: 18 de Enero de 2025. URL: [https://estadistica-dma.ulpgc.es/EyPE/tema2-4\\_DistribucionesEspecialesContinuas.html#1](https://estadistica-dma.ulpgc.es/EyPE/tema2-4_DistribucionesEspecialesContinuas.html#1)

