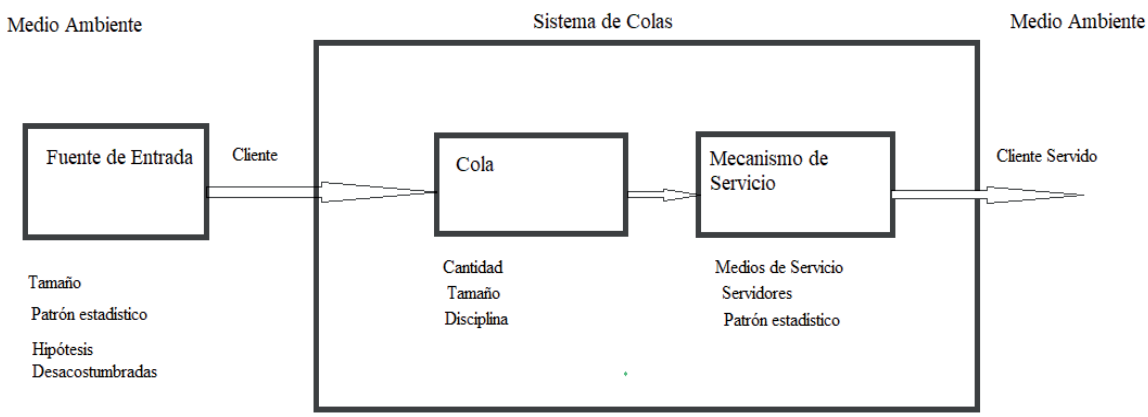
III- Teoría de Colas

III.1 Sistemas de Espera

La teoría de colas estudia y modela los sistemas de espera. Se entiende por tales sistemas a aquellos en los que un cliente, término técnico que pude referirse a personas u objetos según el contexto, ingresa desde el medio ambiente en una suerte de cola para esperar la realización de un servicio. Por ejemplo; una persona llega a una peluquería y espera hasta que le toca su turno para cortarse el pelo, luego se lo cortan y finalmente paga y sale del sistema. Una situación similar ocurre cuando un automóvil llega a un lavadero, cuando una trama de bits llega a un nodo de ruteo en una red y así, en un número grande de sistemas posibles.

El objetivo de la teoría es descriptivo e intenta evaluar algunos parámetros de comportamiento de un sistema de espera partiendo de la idea de que las actividades conformantes se representan por variables aleatorias. Estas variables adquieren sus valores con distinta probabilidad y poseen esperanzas matemática y varianzas. Las distribuciones de cada una de las variables se integran entonces para formar una variable de estado del sistema con su ley de probabilidad. Es a partir de esta nueva distribución que pueden calcularse los referidos parámetros tales como la cantidad esperada de clientes en el sistema o en la cola, o el tiempo esperado de permanencia en el sistema o en la cola, además de las propias probabilidades de que el sistema esté vacío, o de que haya en él una cantidad determinada de clientes.

Comenzaremos entonces por una descripción general de las entidades, atributos y actividades que integran el sistema y de aquellas que interactúan con él desde medio ambiente.



La figura muestra los elementos de interés. En el medio ambiente se ubica la Fuente de Entrada, que emite los clientes hacia el sistema de colas. Esta fuente puede ser de tamaño infinito, lo que quiere decir que tiene la capacidad de emitir clientes sin que estos se agoten o, por el contrario, puede ser finita. En tal caso la cantidad de clientes que debieran pasar por el sistema en algún momento se agotaría. Por supuesto, se trata de un esquema teórico por lo cual, si el número potencial de clientes fuese muy grande, la fuente podría pensarse o modelarse como de tamaño infinito, aunque esto no fuera estrictamente real. La fuente de entradas emite los clientes con un patrón estadístico. Esto quiere decir que se observa una estadística de la entrada de clientes en una dada unidad de tiempo y se utiliza su distribución de frecuencias como un patrón. Por ejemplo, se observa durante un mes la cantidad de compradores que entra por cada hora, de las 12 diarias, en que está abierto un supermercado. Se cuenta entonces la cantidad de horas en las que entró 0 clientes, la cantidad de horas en las que entró 1 cliente, la cantidad de horas en las que entraron 2 clientes y así. Se obtiene entonces una distribución de frecuencias que detalla lo ocurrido.

X = cantidad de clientes que entraron al supermercado en una hora. Variable estadística	F= cantidad de veces que entraron X clientes en una hora. Frecuencia Absoluta	f= cantidad de veces que entraron X clientes en una hora / total de horas de observación. Frecuencia relativa
0	F_0	$f_0 = \frac{F_0}{\sum_{i=0}^{k} F_i}$
1	F_1	$f_1 = \frac{F_1}{\sum_{i=0}^{k} F_i}$
2	F_2	$f_2 = \frac{F_2}{\sum_{i=0}^{k} F_i}$
…	…	…
k	F_k	$f_k = \frac{F_k}{\sum_{i=0}^{k} F_i}$

Esta es claramente una función dada por una tabla. Pero además es de gran importancia pues al ser todas las frecuencias absolutas $F_i \geq 0\ \forall i$, se verifica que $f_i \geq 0$ y también que $\sum_1^k f_i = 1$. De tal modo la tabla puede interpretarse también como una función de probabilidad. Es decir; en base a la estadística registrada se puede calcular la probabilidad de que en la próxima hora lleguen 3 clientes al supermercado, por ejemplo. Por supuesto una función dada por una tabla puede reemplazarse por una fórmula mediante un ajuste adecuado y por lo tanto este patrón estadístico de la fuente de entrada podrá expresarse mediante distribuciones discretas muy conocidas como la de Poisson, por ejemplo, en el caso en que esto corresponda. Esta variable aleatoria que representa la cantidad de clientes en la unidad de tiempo, está asociada a su recíproca que representa, en este caso, el tiempo que transcurre entre la llegada de dos clientes. Así la variable aleatoria tiempo tiene una distribución que también suele ser expresada por fórmulas, por ejemplo, la exponencial. Observemos que si decimos que se producen 30 llegadas por hora, $\frac{1}{30}$ horas/llegadas $=\frac{60}{30}$ minutos/llegadas= 2 minutos entre clientes. En suma, el patrón estadístico es en realidad doble. Por un lado proporciona la probabilidad de que arriben x clientes en una unidad de tiempo, y por otro da la probabilidad de que el tiempo que pase entre la llegada de dos clientes sea $\frac{1}{x}$. Digamos además que, a veces, puede ser necesario considerar alguna hipótesis no muy usual. Si se tratara de analizar un sistema de transporte público, podría muy bien ocurrir que un cliente, potencial pasajero de una línea de colectivos, se contrariara al llegar a la cola porque hay demasiada gente y renunciara a entrar en el sistema. Es decir, renunciara a viajar por ese medio. Si ese fuera un fenómeno frecuente, debiéramos tenerlo en cuenta como hipótesis posible al elaborar un modelo.

Ya dentro del sistema, el cliente, persona u objeto, entra en la cola. Cola es, en el marco de esta teoría, una palabra general que designa tanto la existencia de una como de varias colas. Por ejemplo, en un banco hay tres cajeros automáticos. Puede hacerse una sola cola para dirigirse al primero de ellos que se desocupe, como también podría hacerse una cola delante de cada cajero. En un caso habrá una única cola y en otro la cantidad de colas será tres. Ambas situaciones se modelan distintamente. Por otra parte, el tamaño de la cola se dirá infinito, no estrictamente porque lo sea, sino porque la cola admitirá tantos clientes como le lleguen. Al contrario, cuando la cola admite solo un número de clientes, como por ejemplo en una situación en que se dan 20 números diarios para acceder a un servicio, la cola se dirá finita. Por último, hay que explicar el tema de la disciplina. Una cola puede funcionar bajo la idea que el primero que llega será el primero en ser atendido y así. Es decir, una disciplina que se denomina FIFO (First Input First Output). También puede ocurrir que el último que llega sea el primero en atenderse (LIFO). Y también puede establecerse un sistema de prioridades asignando una valoración en un rango numérico distinto a los clientes, por ejemplo si fueran personas, teniendo en cuenta su edad.

Al pasar al mecanismo de servicio el cliente puede encontrarse con distintos medios de servicio. Por ejemplo, en una gran estación de servicio hay un medio para el despacho de la nafta, otro para poner aire a las cubiertas, otro para cambiar aceite y otro para pagar y comprar golosinas. En cada uno de esos distintos medios del mecanismo hay a su vez servidores que son los surtidores donde se expende la nafta, las fosas donde se cambia el aceite o las distintas cajas donde pagar. De tal modo hay en el mecanismo de servicio distintos medios de servicio y dentro de estos hay servidores. A su vez, se puede considerar para cada servidor un patrón estadístico de servicio cuya variable será, o bien el tiempo que se tarda en concluir un servicio, o bien la cantidad de clientes que son atendidos en una unidad de tiempo. Habrá entonces dos distribuciones de frecuencias asociadas con la actividad de cada servidor en las cuales una variable será la recíproca de la otra como en el caso de las llegadas. Por supuesto también aquí esas distribuciones pueden ser vistas como de probabilidad y permitirnos calcular la probabilidad de concluir un servicio en un dado tiempo o la probabilidad de atender una cantidad de clientes en una unidad temporal.

Esta exposición nos revela el alto grado de variación que pueden tener los supuestos de un sistema de espera. Pueden ser distintos los tamaños y los patrones estadísticos de la fuente de entrada. La cola puede ser única o varias, puede admitir tantos clientes como se quiera o no y puede tener distintas disciplinas de atención. El mecanismo de servicio puede tener uno o varios medios, cada uno con uno o varios servidores, con iguales o distintos patrones estadísticos.

Cualesquiera sean las hipótesis iniciales, el modelado apunta a combinarlas de modo de establecer la forma de medir la cantidad esperada de clientes en el sistema, la cantidad esperada de clientes en la cola, el tiempo esperado de permanencia en el sistema y en la cola y la ley de probabilidad de la variable aleatoria de estado del sistema, que nos permite calcular todos los parámetros antedichos que no son otra cosa que esperanzas matemáticas.

III.2 Proceso de Nacimiento y Muerte

Existe un modelo para representar la vida de las personas según la forma en que se producen los nacimientos y las muertes en la sociedad. La observación estadística de vieja data permite dar por sentado que el número de nacimientos en un período de tiempo fijado es una variable aleatoria que se distribuye Poisson. Lo mismo se ha observado respecto de las muertes producidas en un lapso de tiempo determinado. Recordemos que la distribución de Poisson tiene la fórmula:

$$p(k) = \frac{\lambda^k e^{-\lambda}}{k!} \quad k = 0,1,2,\dots$$

dónde el parámetro λ es la esperanza de la variable aleatoria $E(k) = \lambda$. Las unidades de la variable son, en este modelo, personas/unidad de tiempo. Llamaremos entonces λ a la cantidad esperada de nacimientos en la unidad de tiempo y μ a la cantidad esperada de muertes en la misma unidad temporal. Ahora bien; en términos del modelo que se plantea, la recíproca de la variable k, tanto si representa nacimientos como muertes en el tiempo fijado, modela el tiempo entre eventos. Es decir $t = \frac{1}{k}$ unidades de tiempo/personas es la variable aleatoria que simboliza el tiempo entre nacimientos o muertes según corresponda. Esta variable es continua y puede probarse que su distribución es exponencial. Dicho de otra forma, su función de densidad de probabilidad es:

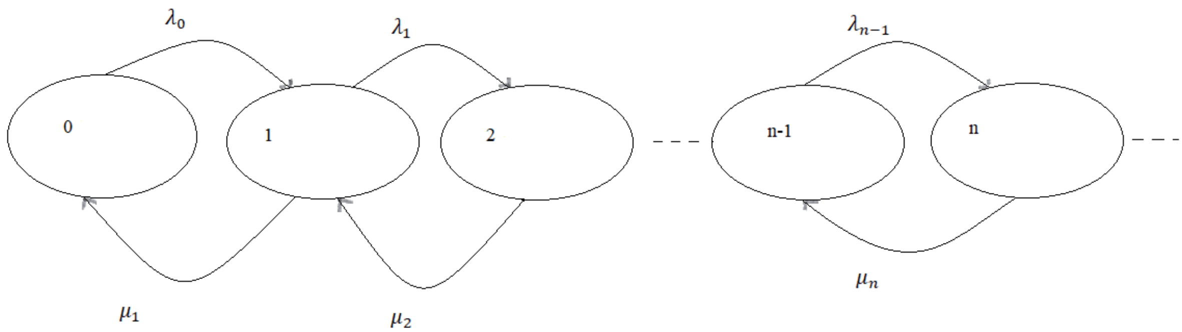
$$f(t) = \begin{cases} \alpha e^{-\alpha t} & t \geq 0 \; y \; \alpha > 0 \\ 0 & en \; otra \; parte \end{cases}$$

La esperanza de tal variable aleatoria es $E(t) = \frac{1}{\alpha}$ y existe una relación recíproca también entre la esperanza de Poisson y la de la exponencial, habida cuenta que ambas son valores esperados de las recíprocas variables aleatorias k y t. Así resulta:

$$\lambda = E(k) = \frac{1}{E(t)} = \frac{1}{\frac{1}{\alpha}} = \alpha.$$

En definitiva, el parámetro de la exponencial es la esperanza de la variable Poisson asociada. Así si, por ejemplo, en una determinada comunidad nacen en promedio 6 personas por día, teniendo en cuenta la distribución de Poisson de los nacimientos podremos decir que $\lambda = 6 \; \frac{nacimientos}{día}$. También podremos afirmar que el tiempo promedio que transcurre entre nacimientos en un día es $E(t) = \frac{1}{\alpha} = \frac{1}{\lambda} = \frac{1}{6} \; \frac{dias}{nacimiento} = \frac{24}{6} \; \frac{horas}{nacimiento} = 4 \; \frac{horas}{nacimiento}$
Debe leerse entonces que, en promedio, al día transcurren 4 horas entre cada nacimiento. A partir de lo expuesto se observa que el patrón estadístico es doble pues contiene a la distribución de Poisson de la variable aleatoria que representa los nacimientos (o las muertes) en la unidad de tiempo, y, también, a la variable aleatoria que simboliza al tiempo que transcurre entre nacimientos (o muertes). En adelante hablaremos entonces de patrón estadístico Poisson-Exponencial.

Ahora debemos considerar el tamaño de la población. Si hay 1000 personas vivas podemos pensar que al año habrá una cierta cantidad esperada de nacimientos. Para fijar ideas supongamos que esa cantidad promedio fuera $\lambda = 200\ \frac{nacimientos}{año}$. Pero si hubiera 1000000 de personas vivas los nacimientos tendrían que ser, en promedio, más. Es decir que en nuestro modelo de nacimiento y muerte debemos considerar también como variable la cantidad de personas vivas para registrar el hecho que los parámetros de las distribuciones de nacimientos y muertes deben variar según ella. Así las cosas, en términos generales diremos que si el estado del sistema es n, n personas vivas, habrá un parámetro λ_n para los nacimientos y un parámetro μ_n para las muertes. Si el estado n se considera también una variable aleatoria que depende a su vez de las variables aleatorias que representan los nacimientos y las muertes, en cada valor de la variable aleatoria n habrá distribuciones poisson-exponencial distintas, con diferentes parámetros λ_ny μ_n. Una sucesión de distintas variables aleatorias a través del tiempo se denomina proceso estocástico y entonces este que aquí presentamos será llamado el proceso de nacimiento y muerte. El diagrama de tasas que se muestra a continuación expresa esta idea.



Las flechas indican la dirección de cambio de estado según se trate de nacimientos o muertes. Es claro que para que el sistema permanezca en estado 0 se hace necesario que las personas que nacen y cambian el estado hacia 1 se equilibren con las que mueren y cambian el estado hacia 0. Pero los estados 0 y 1 no son igualmente probables. Es decir hay una p_0 de que el sistema esté en estado 0 y una p_1 de que el sistema esté en estado 1. De modo que el equilibrio para el estado 0 se alcanza cuando $\lambda_0 p_0 = \mu_1 p_1$. De similar forma el equilibrio en estado 1 se alcanza cuando $\lambda_0 p_0 + \mu_2 p_2 = \lambda_1 p_1 + \mu_1 p_1$. En forma general este balance se expresa por una lista de ecuaciones:

$\lambda_0 p_0 = \mu_1 p_1$ $\qquad$ $p_1 = \frac{\lambda_0}{\mu_1} p_0$

$\lambda_0 p_0 + \mu_2 p_2 = \lambda_1 p_1 + \mu_1 p_1$ $\qquad$ $\lambda_0 p_0 + \mu_2 p_2 = \lambda_1 \frac{\lambda_0}{\mu_1} p_0 + \cancel{\mu_1} \frac{\lambda_0}{\cancel{\mu_1}} p_0$ $\qquad$ $p_2 = \frac{\lambda_0 \lambda_1}{\mu_1 \mu_2} p_0$

$\lambda_1 p_1 + \mu_3 p_3 = \lambda_2 p_2 + \mu_2 p_2$ $\qquad$ $p_3 = \frac{\lambda_0 \lambda_1 \lambda_2}{\mu_1 \mu_2 \mu_3} p_0$

… $\qquad$ …

$\lambda_{n-2} p_{n-2} + \mu_n p_n = \lambda_{n-1} p_{n-1} + \mu_{n-1} p_{n-1}$ $\qquad$ $p_n = \frac{\lambda_0 \lambda_1 \ldots \lambda_{n-1}}{\mu_1 \mu_2 \ldots \mu_n} p_0$

$\lambda_{n-1} p_{n-1} + \mu_{n+1} p_{n+1} = \lambda_n p_n + \mu_n p_n$ $\qquad$ …

…

En suma, las probabilidades de los distintos estados quedan en función de los parámetros de cada una de las distribuciones de nacimiento y muerte en cada estado y de la probabilidad de estado 0. Puede hacerse ahora el reemplazo: $p_n = C_n p_0$ donde $C_n = \frac{\lambda_0 \lambda_1 \ldots \lambda_{n-1}}{\mu_1 \mu_2 \ldots \mu_n}$. La sucesión $p_0, p_1, \ldots p_{n-1}, p_n$ ...expresa la distintas probabilidades de estado, o sea las probabilidades de que haya un número n de personas en el sistema de nacimiento y muerte. Como se trata de una función de probabilidad se cumple que $\sum_0^\infty p_n = 1$ y habrá parámetros de la distribución tales como la cantidad esperada de personas en el sistema que sean de interés.

Para conocer un tanto mejor esta distribución de la variable aleatoria de estado se pueden hacer algunas cuentas y reemplazos.

$$1 = \sum_0^\infty p_n = p_0 + \sum_{n=1}^\infty p_n = p_0 + \sum_{n=1}^\infty C_n p_0 = p_0 (1 + \sum_{n=1}^\infty C_n)$$

Es decir; $p_0 = \frac{1}{1 + \sum_{n=1}^\infty C_n}$

lo que significa que la probabilidad de estado 0 también queda en función de los parámetros λ_n y $\mu_n \quad \forall\, n = 0, 1, 2, \ldots$

El proceso de nacimiento y muerte puede ser visto como un sistema de colas en el cual los nacimientos se interpretan como emisiones de la fuente de entrada, la vida como el tiempo que la persona permanece en la cola y la muerte como el servicio. Según la cantidad de personas vivas, es decir el estado del sistema, serán las tasas de nacimiento y muerte de las respectivas distribuciones exponenciales. Además, el proceso es una cadena de Markov pues, en cada estado, la probabilidad de un evento, se trate de un nacimiento o una muerte, depende solo de ese estado actual. En efecto, el modelado exponencial de los tiempos entre nacimientos y los de muertes, surge en cada estado con independencia de cualquier otro estado anterior o futuro. Esto es lo que define su propiedad markoviana.

III.3 El modelo M/M/1

El nombre M/M/1 se refiere a una modificación simplificadora en las hipótesis que condujeron a establecer el proceso de nacimiento y muerte. Aquí se supondrá que, no importa cuál sea el valor de la variable de estado, las distribuciones de tiempos entre llegadas y de tiempos de servicio tendrán siempre iguales valores en sus parámetros. O sea, suponemos que $\lambda = \lambda_n$ es una constante cualquiera sea el valor de *n*, y que, de la misma forma, $\mu = \mu_n$ es también una constante. Se observa que, bajo este supuesto, la propiedad markoviana en las llegadas al sistema y en los servicios se continúa manteniendo. La nomenclatura M/M/1 nombra entonces a un sistema que tenga estas características para entradas y salidas con un solo servidor. El resumen de los supuestos que realizamos para este modelo M/M/1 es el siguiente: la fuente de entrada tiene tamaño infinito con patrón estadístico Poisson-exponencial y sin hipótesis desacostumbradas; la cola será única, de tamaño potencialmente infinito y disciplina FIFO; y el mecanismo de servicio tendrá un solo medio de servicio con un solo servidor cuyo patrón será Poisson-exponencial.

Como ya se ha mencionado, el interés de la teoría está en conocer la ley de probabilidad de estado para calcular efectivamente la probabilidad de que haya n clientes en el sistema. A su vez, interesa también establecer la cantidad esperada de clientes en el sistema y en la cola. Otros parámetros de interés son el tiempo esperado total de espera en la cola y en el sistema. A partir del proceso de nacimiento y muerte y de todas las suposiciones hechas, en particular las de constancia de λ y μ, trataremos ahora de deducir las fórmulas necesarias.

La constancia de λ y μ permite definir el llamado factor de tráfico del sistema $\rho = \frac{\lambda}{\mu}$. Si se considera que los clientes deben entrar al sistema a una velocidad promedio menor de la que salen, pues de lo contrario el sistema se saturaría al existir una cola cada vez más larga, se ve que $\lambda < \mu$. Por lo tanto, habida cuenta además, que $\lambda > 0$ y $\mu > 0$, debe resultar que $0 < \rho < 1$. Si las velocidades de entrada y salida son similares el parámetro estará cercano a 1, mientras que si son muy distintas el parámetro se acercará al valor 0. Obsérvese ahora que $C_n = \frac{\lambda_0\lambda_1\ldots\lambda_{n-1}}{\mu_1\mu_2\ldots\mu_n} = \frac{\lambda\lambda\ldots\lambda}{\mu\mu\ldots\mu} = \frac{\lambda^n}{\mu^n} = \rho^n$. Por lo tanto, como $\rho^0 = 1$, resulta

$$1 = \sum_0^\infty p_n = p_0 + \sum_{n=1}^\infty p_n = p_0 + \sum_{n=1}^\infty C_n p_0 = p_0(1 + \sum_{n=1}^\infty \rho^n) = p_0 \sum_{n=0}^\infty \rho^n$$

Dado que $0 < \rho < 1$, la última suma es la de los infinitos términos de una progresión geométrica de razón ρ, que converge a $\frac{1}{1-\rho}$ por lo que resulta $1 = p_0(\frac{1}{1-\rho})$ y de aquí:

$p_0 = 1 - \rho$. Además como $p_n = C_n p_0$ se tiene que $p_n = \rho^n(1-\rho)$. Ahora contamos entonces con una expresión para la probabilidad de estado n con $n = 0,1,2,\ldots$

La cantidad esperada de clientes en el sistema nos es otra cosa que la esperanza matemática de la variable aleatoria de estado. Por lo tanto, llamándola L tenemos:

$$L = \sum_{n=0}^\infty n p_n = \sum_{n=0}^\infty n\,\rho^n(1-\rho) = (1-\rho)\sum_{n=0}^\infty n\rho\rho^{n-1} = (1-\rho)\rho\sum_{n=0}^\infty n\rho^{n-1}$$

El argumento bajo la última sumatoria es la derivada de ρ^n respecto de ρ, entoces :

$$L = (1-\rho)\rho\sum_{n=0}^\infty \frac{d\rho^n}{d\rho} = (1-\rho)\rho\frac{d}{d\rho}(\sum_{n=0}^\infty \rho^n) = (1-\rho)\rho\frac{d}{d\rho}\left(\frac{1}{1-\rho}\right)$$

Luego: $$L = (1-\rho)\rho\frac{1}{(1-\rho)^2} = \frac{\rho}{1-\rho} = \frac{\frac{\lambda}{\mu}}{1-\frac{\lambda}{\mu}} = \frac{\frac{\lambda}{\mu}}{\frac{\mu-\lambda}{\mu}} = \frac{\lambda}{\mu-\lambda}$$

Es decir, finalmente tenemos que la cantidad esperada de clientes en el sistema queda en función de los parámetros λ y μ de las distribuciones de entrada y salida del sistema.

Con un cálculo similar se establece que la cantidad esperada de clientes en la cola es:

$$L_q = \frac{\lambda^2}{\mu(\mu-\lambda)}$$

En 1961 el matemático John Little estableció dos relaciones importantes entre la cantidad esperada de clientes en el sistema y la cola por un lado y, por el otro, el tiempo esperado de permanencia en el sistema y en la cola.

$L = \lambda W$ y

$L_q = \lambda W_q$.

Entonces el tiempo esperado de espera en el sistema se calcula por

$$W = \frac{L}{\lambda} = \frac{1}{\lambda}\left(\frac{\lambda}{\mu-\lambda}\right) = \frac{1}{\mu-\lambda}$$

A su vez, el tiempo esperado de permanencia en la cola es

$$W_q = \frac{L_q}{\lambda} = \frac{\lambda}{\mu(\mu-\lambda)}$$

Si μ es la cantidad esperada de clientes atendidos en la unidad de tiempo entonces $\frac{1}{\mu}$ es el tiempo esperado de servicio. Al sumar el tiempo esperado de permanencia en cola y el tiempo esperado de servicio, la suma debiera dar el tiempo esperado de estancia en el sistema. Esto efectivamente puede comprobarse según:

$$W_q + \frac{1}{\mu} = \frac{\lambda}{\mu(\mu-\lambda)} + \frac{1}{\mu} = \frac{\lambda + (\mu-\lambda)}{\mu(\mu-\lambda)} = \frac{\mu}{\mu(\mu-\lambda)} = \frac{1}{\mu-\lambda} = W$$

Finalmente hemos obtenido las fórmulas esenciales del modelo que corresponden a la probabilidad de la variable de estado y a las esperanzas matemáticas de cantidad de clientes en sistema y cola, y a los tiempos esperados de permanencia en el sistema y la cola. Una aplicación de todas estas cantidades puede verse en el ejemplo siguiente.

Ejemplo 1: Un lavadero de automóviles lava a razón de 15 autos/hora los coches que llegan a razón media de uno cada 6 minutos. Asumiendo las hipótesis adecuadas, se desea calcular:

a) La cantidad esperada de autos en la cola.
b) El tiempo esperado de permanencia de cada auto en el lavadero
c) La probabilidad de que un auto al llegar tenga que esperar para ser lavado

Cuando se habla de hipótesis adecuadas en realidad se está pidiendo que se piense si la forma de trabajo en el lavadero se ajusta a los supuestos de un determinado modelo de colas. Es decir, comprobar que, en este caso, se cumplen las hipótesis del modelo M/M/1 en cuanto a la fuente de entrada, la cola y el mecanismo de atención. En particular, deberán encontrarse los parámetros λ y μ de los patrones Poisson-exponencial de llegadas y servicios.

Asumiendo entonces el comportamiento de Poisson para la cantidad de automóviles lavados en una hora, establecemos $\mu = 15\ automóviles/hora$ como su media. Para las llegadas observamos que lo que en realidad tenemos es el tiempo medio entre llegadas. Como sabemos que este valor es el recíproco del parámetro λ de la distribución de Poisson que las modela hacemos $6\frac{minutos}{automóvil} = \frac{1}{\lambda}$ de donde obtenemos $\lambda = \frac{1}{6}\ automóvil\ /minuto = \frac{60}{6} automóvil/hora$

Es decir; finalmente $\lambda = 10\ automóviles/hora$. Podemos ahora calcular:

a) $L_q = \frac{\lambda^2}{\mu(\mu-\lambda)} = \frac{10^2}{15(15-10)} = 1{,}33\ automóviles$ Hay que ver que esta es una cantidad teórica, que permitirá calcular costos, por ejemplo, pero que no tendrá realización práctica pues difícilmente puedan hallarse un auto y la tercera parte de otro realizando una cola.

b) $W = \frac{1}{\mu-\lambda} = \frac{1}{15-10} = 0{,}2\ horas = 0{,}2 \times 60\ minutos = 12\ minutos$

c) Si un auto tiene que esperar al llegar es porque en el sistema ya hay al menos uno. En el caso de haber exactamente uno estará en el servicio y el nuevo auto que llega será el primero de la cola. Si ya hubiera más de un auto, la cola ya estará formada y el nuevo automóvil se incorporará a ella. En cualquier caso, la probabilidad de tener que esperar se podrá calcular como:

   a. $P(esperar) = P(n \geq 1) = 1 - p_0 = 1 - (1-\rho) = \rho = \frac{\lambda}{\mu} = \frac{10}{15} = 0{,}33$

Ejercicios resueltos

Ejercicio N° 1- En un lugar de atención con un único lugar de despacho y cola simple se conoce el arribo de clientes que asciende a 15 unidades por hora, siendo la velocidad media de servicio de 25 unidades por hora. Calcular todos los elementos característicos de la cola y del sistema, suponiendo que la ley de arribos y servicios es Poisson. Calcular los mismos parámetros suponiendo λ constante y valores de μ menor y mayor al dato original. Resumir en un cuadro los resultados obtenidos y enunciar las conclusiones referentes a la variación obtenida en los parámetros.

λ = (velocidad de arribos)
μ= (velocidad de servicio)

	λ= 15u/h μ= 25u/h	λ= 15u/h μ=20u/h	λ= 15u/h μ=30u/h	λ= 20u/h μ=25u/h	λ=10u/h μ=25u/h
L	1,5	3	1	4	0,67
L_q	0,9	2,25	0,5	3,2	0,267
W_s	0,1	0,2	0,067	0,2	0,067
W_q	0,06	0,15	0,033	0,16	0,0267
ρ	0,6	0,75	0,5	0,8	0,4
p_0	0,4	0,25	0,5	0,2	0,6

Los cálculos revelan como el modelo acierta a describir las variaciones del sistema. En la medida en que el factor de tráfico ρ se acrecienta la cantidad esperada de clientes en el sistema y en la cola. Se acrecientan también en ese caso los tiempos esperados de permanencia de un cliente en el sistema y en la cola. Finalmente, la probabilidad p_0 de que el sistema este ocioso se hace cada vez menor cuando crece el factor de tráfico.

Ejercicio N° 2- Las llegadas de personas a una cabina telefónica so del tipo Poisson con un promedio de 6 minutos entre dos consecutivas. La duración de una llamada telefónica tiene una distribución exponencial con una media de 4 minutos. Calcular:

a) La probabilidad de que una persona que llegue a la cabina tenga que esperar.
b) La probabilidad de que halle más de 4 personas en la cola
c) La longitud promedio de la cola y del sistema

d) La compañía está dispuesta a instalar otra cabina si se comprueba que el tiempo medio de espera asciende a 10 minutos. Determinar cuál debe ser el flujo de arribos para justificar esa instalación.

λ=1/6 cliente/min μ=1/4 cliente/ min
λ=10 clientes/ hora μ= 15clientes/ hora

a) Una persona que llegue a la cabina tendrá que esperar siempre que haya al menos un cliente en el sistema. Es decir, la probabilidad de esperar será el complemento de la probabilidad de que el sistema esté vacío

$$p_0 = 1 - \rho = 1 - \frac{10}{15} = 0.33$$
$$P(esperar) = 1 - p_0 = 1 - 0.33 = 0{,}667$$

b) Si hay más de 4 personas en la cola es porque hay más de 5 en el sistema ya que si hay cola es porque hay una persona hablando en la cabina. Luego la probabilidad de que haya más de 4 personas en la cola es la probabilidad de que haya más de 5 en el sistema.

$$P(n > 5) = P(n \geq 6) = 1 - [\,P_0 + P_1 + P_2 + P_3 + P_4 + P_5] = \rho^6 = (2/3)^6$$
$$P(n \leq 6) = 0{,}0878$$

La probabilidad de hallar más de 4 personas en la cola es del 0.0878

c) $L_q = \frac{\lambda^2}{\mu(\mu-\lambda)} = \frac{(10^2}{15(15-10)} = \frac{4}{3}$

$$L = \frac{\lambda}{\mu(\mu-\lambda)} = \frac{10clientes/h}{(\frac{15clientes}{h} - \frac{10clientes}{h})} = 2$$

d) $W_q = \frac{1}{6}$

$$\frac{\lambda}{\mu(\mu-\lambda)} = \frac{\lambda}{15(15-\lambda)} = \frac{1}{6}$$

$$21\lambda = 225$$

$$\lambda = 10{,}71\ \frac{\text{cl}}{\text{h}}$$

Ejercicio N° 3- En una fábrica un mecánico destinado al mantenimiento de las maquinas atiende todos los desperfectos que en ellas se presentan. Se ha observado que la demanda de servicios sigue la ley de Poisson con un parámetro λ=2.5 y que el mecánico atiende los pedidos con una velocidad promedio μ= 4.6, según una rigurosa disciplina F I FO.
Determinar:

a) Número promedio de máquinas en espera de ser atendidas.
b) Tiempo promedio de atención.
c) Tiempo promedio entre desperfectos.
d) Tiempo promedio de espera de la maquina hasta comenzar su atención y hasta que se reintegre a la producción.
e) Determinar hasta que suma podría pagarse un incentivo para que el mecánico eleve su rendimiento de 90% (rendimiento actual) al 120% si la hora-hombre cuesta $200 y la hora-máquina cuesta $500.

λ=2.5 media de desperfectos en máquinas/hora con patrón Poisson exponencial.
μ=4.6 media de atención en máquinas/hora suponiendo patrón Poisson exponencial.

a) Número promedio de máquinas en espera de ser atendidas.

$$L_q = \frac{2.5^2}{4.6(4.6-2.5)} = 0.647$$

b) Tiempo promedio de atención

$$\frac{1}{\mu} = \frac{1}{4.6} = 0.22\,\frac{horas}{máquina} = 0.22\,\frac{60\ minutos}{máquina} \approx 13 minutos/máquina$$

En promedio se atiende una máquina en 13 minutos

c) Tiempo promedio entre desperfectos.

$$\frac{1}{\lambda} = \frac{1}{2.5} = 0.4\,\frac{horas}{máquina} = 0.4\,\frac{60\ minutos}{máquina} = 24 minutos/máquina$$

En promedio se descompone una máquina cada 24 minutos.

d) Tiempo promedio de espera de la maquina hasta comenzar su atención y hasta que se reintegre a la producción.

$$W_q = \frac{2.5}{4.6(4.6-2.5)} = 0.259\,\frac{horas}{máquina} \approx 16 minutos/máquina$$

El tiempo promedio de espera de una máquina en ser atendida es de 16 minutos aproximadamente.

Teniendo en cuenta este resultado y el del punto b) hacemos:

$$W = W_q + \frac{1}{\mu} = 16 + 13 = 29\ minutos/máquina$$

El tiempo promedio que permanece una máquina fuera de servicio es de 29 minutos

e) Determinar hasta que suma podría pagarse un incentivo para que el mecánico eleve su rendimiento de 90% (rendimiento actual) al 120% si la hora-hombre cuesta $200 y la hora-máquina cuesta $500.

$\acute{\mu} = \frac{120}{90} 4.6 = 6.13$ es la velocidad de atención a la que el mecánico trabajaría si eleva su rendimiento al 120%.

El costo de mantenimiento comprende el costo de la espera de la máquina hasta ser atendida y el costo mientras se la atiende y sigue detenida más el trabajo del mecánico. Es decir:

$$Costo(\mu) = 500W_q + (500 + 200)\frac{1}{\mu}$$

Actualmente ese costo es:

$$Costo(4.6) = 500\frac{2.5}{4.6(4.6-2.5)} + (500 + 200)\frac{1}{4.6} = 281.75\$$$

Si el mecánico trabaja a velocidad $\acute{\mu} = 6.13$ máquinas/hora y se le paga un incentivo Δ el costo de esa operatoria será:

$$Costo(6.13) = 500\frac{2.5}{6.13(6.13-2.5)} + (500 + 200 + \Delta)\frac{1}{6.13} = 56.17 + 114.19 + \frac{\Delta}{6.13}$$

O sea $Costo(6.13) = 170.36 + \frac{\Delta}{6.13}$ de donde para que toda la operatoria a la nueva velocidad sea más barata debe ocurrir que:

$281.75 \geq 170.36 + \frac{\Delta}{6.13}$ De aquí $\Delta \leq 682.82\$$. Entonces el incentivo podría ser de hasta 682.82 $ sin que pagarlo incremente los costos actuales.

Ejercicios propuestos

Ejercicio N° 4- Frente a la ventanilla de franqueo de una oficina de correos se presentan 70 personas por día (jornada de 10 hs.). Se atiende en promedio a 10 personas por hora. Asumiendo las hipótesis convenientes, determinar:

a) Longitud media de la cola frente a la ventanilla.
b) Probabilidad de que exista una fila de más de 2 personas.
c) Tiempo promedio de espera en el sistema y en la fila.
d) Probabilidad de tener que esperar.

Ejercicio N° 5- En una estación de abastecimiento de oxigeno liquido de una planta química en la que se utiliza transporte por tanques semiremolque, se sabe que el costo del tiempo de espera de cada tanque para ser abastecido es de 40.000 $/hora, mientras que el costo horario de llenado de cada tanque es de $10.000. Si los tanques arriban a razón de 2 por hora, ¿cuál es la velocidad de despacho que hace mínimo el costo operativo total? (suponemos arribos Poisson y tiempos de servicio exponenciales)

Ejercicio N° 6- Una planta textil posee 30 telares, se requiere la presencia de un operario para la atención permanente de las maquinas (arreglo de desperfectos). Se supone que las roturas verifican las hipótesis de Poisson con una media de 3 roturas por día. La empresa debe elegir entre contratar al operario A o al operario B. A cobra $400 por día y arregla un promedio de 5 máquinas / día (servicio exponencial), mientras que B cobra $900 por 7 máquinas / día.

Una máquina detenida (tiempo no productivo) origina a la empresa un costo de $700 diarios. ¿Cuál será el operario elegido?
Hipótesis simplificadora: $N > 20$ población infinita.

Ejercicio N°7- En una gran empresa se examina el departamento de despacho ya que muchas veces se producen retrasos en las entregas. Los datos evidencian que el tiempo empleado en despachar los envíos constituye el 51% del tiempo. Se supone que las órdenes de envío llegan a la sección a razón de 5,75 por hora (y satisface Poisson).

a) ¿Cuál es la probabilidad de que la sección esté inactiva?
b) ¿Cuál es la longitud media de la cola?
c) ¿Cuál debería ser aproximadamente la tasa de servicio para que la longitud media de la cola fuera $\approx 0{,}12$ (punto doce)?

Si se sustituyeran las mecanógrafas actuales que escriben 75 palabras por minuto, por otras que escribieran 100 palabras. ¿Cuál sería el número medio de mensajes esperando a ser mecanografiados?

Ejercicio N° 8- Sea un modelo de colas (Poisson- exponencial) con λ=3 cl/m y μ=5 cl/m. ¿Cuál es la probabilidad de que no haya ningún cliente en la cola?

Ejercicio N° 9- Suponga una máquina automática expendedora de bebidas gaseosas. El tiempo que tarda en satisfacer el pedido de los clientes es de 68 segundos (determinístico). Si la llegada de los clientes puede considerarse Poisson con λ=0,3 cl/minuto. ¿Cuál será el tiempo medio de permanencia del cliente en el sistema?

Ejercicio N° 10- Un cerrajero hace llaves "en el acto". Los clientes llegan con una tasa media de arribos igual a 0,08 clientes por minuto, siguiendo aproximadamente una ley de Poisson. Como en el vecindario hay otras cerrajerías, el cerrajero desea que por lo menos la mitad de sus clientes no tengan que esperar. En tal caso, ¿qué velocidad debe imprimir a su trabajo, o sea como debe ser, en promedio, el tiempo de servicio? (Asumimos distribución exponencial).

Ejercicio N° 11- Modelo de una sola cola (Poisson-exponencial). Si λ clientes / minuto mide la tasa media de arribos. ¿Cómo debe ser el tiempo medio de servicio para que por lo menos la mitad de los clientes no tenga que esperar?

Ejercicio N° 12- Se sabe que $P(n \leq 1) = 0{,}36$ ¿Cuál es la probabilidad de canal ocioso?

Ejercicio N°13- En un servicio de mecánica ligera al paso los automóviles llegan, en promedio, para un cambio de aceite cada 15 minutos. El servicio trabaja con una atención media de 48 automóviles en la jornada de 8 hs.

a) asuma las hipótesis convenientes para modelar este servicio y explicítelas.
b) Halle la cantidad esperada de automóviles en el sistema suponiendo que hay un solo servidor y halle el tiempo esperado de espera para ser atendido.
c) ¿Cuál es la probabilidad de que haya cola frente al servidor?

d) Si los automóviles llegaran en promedio cada 10 minutos en que porcentaje tendría que aumentar la velocidad de servicio para mantener igual el tiempo medio de la espera para ser atendido.

Ejercicio N° 14- En un sistema de atención de una sucursal bancaria, los clientes llegan a razón media de 10 por hora. Los tiempos de atención son, en promedio, de 2 minutos por cliente. Asumiendo las hipótesis convenientes se desea establecer:

a) Probabilidad de que un cliente no tenga que esperar al llegar al banco.
b) Probabilidad de que haya cola frente a la caja.
c) Si se calcula que la hora promedio del cliente vale 50$ y que cada hora de atención del cajero vale 10$ ¿cuál es el costo económico promedio de la transacción por cliente?
d) ¿En que porcentaje debería aumentar la velocidad de atención del cajero para que dicho costo disminuyera un 10%?

Ejercicio N° 15- El 40% del tiempo que un automóvil taxímetro permanece en un servicio de lavado corresponde a la espera en ser atendido. Si los automóviles se atienden a razón media de 3 por hora se desea establecer el costo promedio que representa para cada taxi su paso por el servicio. El lavado de cada vehículo cuesta 30 pesos y el costo de no estar trabajando se estima en 50$ por hora.

Ejercicio N° 16- Un operario repara equipos electrógenos a razón media de 4 por día en su jornada laboral de 8 hs . En el mismo lapso horario se descomponen 2 equipos en promedio cuyo costo por hora de inactividad es 500 $. La hora del operario tiene un valor de 50 $. Asumiendo las hipótesis necesarias calcular el costo promedio de tener un equipo fuera de servicio.

Ejercicio N° 17- A una cabina de peaje llegan vehículos según un proceso de Poisson, con una media de 90 vehículos por hora. El tiempo en pasar por la cabina se distribuye exponencialmente con un promedio de 38 segundos. . El concesionario se plantea reducir este promedio a 30 segundos instalando dispositivos de cobro automático siempre que se satisfagan dos condiciones: que la cantidad promedio actual de vehículos en el sistema sea mayor que 5 y que el porcentaje de tiempo ocioso de la cabina con el nuevo dispositivo sea menor que el 10%. Con estas condiciones; ¿se puede justificar la instalación del dispositivo automático?

Ejercicio N° 18- Un cajero automático ubicado en un aeropuerto recibe clientes a razón media de 10 por hora durante todo el día. Con el software actual el cajero atiende a cada cliente en un promedio de 3 minutos a un costo estimado de 10 pesos por minuto. Un nuevo software podría atender a cada cliente en un promedio de dos minutos, pero su costo estimado sería de 12 pesos por minuto de atención. Por otra parte, el banco ha evaluado que el tiempo que emplea el cliente para toda la operación, en relación con la imagen de eficiencia del banco, en 1 peso por cada minuto. Se desea entonces saber si en términos de costos para el banco es conveniente renovar el software.