Di Marco, Silvia
Zárate, Melina
Quaglino, Marta
Instituto de Investigaciones Teóricas y Aplicadas de la Escuela de Estadística

UNA APLICACIÓN DE TEORÍA DE COLAS EN EL ÁREA ADMINISTRATIVA

1. INTRODUCCIÓN

Todos hemos experimentado en alguna ocasión la sensación de estar perdiendo el tiempo al esperar en una cola. Ejemplos reales de esa situación son: los cruces de dos vías de circulación, los semáforos, el peaje de una autopista, los cajeros automáticos, la atención a clientes en un establecimiento comercial, la avería de electrodomésticos u otro tipo de aparatos que deben ser reparados por un servicio técnico, etc.

En el contexto de la informática y de las nuevas tecnologías estas situaciones de espera son más frecuentes. Así, por ejemplo, los procesos enviados a un servidor para su ejecución forman colas de espera mientras no son atendidos; la información solicitada, a través de Internet, a un servidor Web puede recibirse con demora debido a la congestión en la red; también se puede recibir la señal de línea de la que depende nuestro teléfono móvil, ocupada si la central está colapsada en ese momento, etc.

Como clientes no queremos esperar *¿por qué hay que esperar?*. La respuesta es casi siempre simple, en algún momento la capacidad de servicio es menor que la capacidad demandada. La Teoría de Colas es la rama de la investigación operativa que estudia las listas de espera. El estudio de las colas es importante porque proporciona tanto una base teórica del tipo de servicio que podemos esperar de un determinado recurso, como la forma en la cual dicho recurso puede ser diseñado para proporcionar un determinado grado de servicio a sus clientes. Por ello se plantea como algo muy útil el desarrollo de una herramienta que sea capaz de dar una respuesta sobre las características que tiene un determinado modelo de colas. Algunos de los campos de utilización más conocidos son: logística de procesos industriales de producción, ingeniería de redes y servicios e ingeniería de sistemas informáticos.

El origen de la Teoría de Colas se remonta a los estudios que realizó Agner Kraup Erlang (Dinamarca, 1878 - 1929) en 1909 para analizar la congestión de tráfico telefónico con el objetivo de cumplir la demanda incierta de servicios en el sistema telefónico de Copenhague. Sus investigaciones acabaron en una nueva teoría denominada **Teoría de Colas o de Líneas de Espera**. Esta teoría es ahora una herramienta de valor en negocios debido a que un gran número de problemas pueden caracterizarse, como problemas de congestión llegada-salida.

En la teoría de la formación de colas, generalmente se llama *sistema* a un grupo de unidades físicas, integradas de tal modo que pueden operar al unísono con una serie de operaciones organizadas. La teoría de la formación de colas busca una solución al problema de la espera prediciendo primero el comportamiento del sistema. Pero la solución al problema de la espera consiste no solo en minimizar el tiempo que los clientes pasan en el sistema, sino también en minimizar los costos totales de aquellos que solicitan el servicio y de quienes lo prestan.

En este trabajo se aborda un ejemplo del uso de Teoría de Colas como una herramienta de análisis de los tiempos administrativos de la justicia, introducido por Miró y Depaoli (2006), que mostrarán cómo puede utilizarse esta herramienta matemática para estimar el tiempo medio de resolución de una causa y algunas probabilidades particulares, planteándose la posibilidad de incorporar el enfoque de simulación del sistema, que permitiría comprender su dinámica y abordar su estudio desde otra perspectiva.

2. METODOLOGÍA

Un sistema de colas se puede describir como: *clientes* que llegan a un "lugar" demandando un servicio a un "servidor", el cual tiene una cierta capacidad de atención. Si el servidor no está disponible inmediatamente y el cliente decide esperar, entonces se forma la línea de espera. En algunos casos se puede admitir que los clientes abandonan el sistema si se cansan de esperar. El término *cliente* se utiliza en sentido general y no implica que sea un ser humano. Puede significar, entre otros, piezas esperando su turno para ser procesadas, una lista de trabajos esperando para imprimir en una impresora de red, etc.

La Teoría de Colas es una colección de modelos matemáticos que describen sistemas de línea de espera particulares y pueden representarse como colas sencillas o como un sistema de colas interconectadas formando una *red de colas*. Los modelos sirven para encontrar el equilibrio que debe haber entre el costo de proporcionar un buen servicio y el costo del tiempo de espera del cliente. Es claro que una representación detallada exige definir un número elevado de parámetros y funciones. En la Figura 1 podemos ver un ejemplo de modelo de colas sencillo.

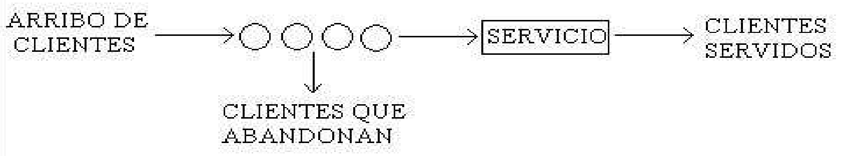

FIGURA 1. UN SISTEMA DE COLA TÍPICO

El problema es determinar qué capacidad o tasa de servicio proporciona el balance correcto. Esto no es sencillo, ya que un cliente no llega a un horario fijo ni el tiempo de servicio es constante. Las características básicas que se deben utilizar para describir un sistema de colas son las siguientes:

a) Régimen de llegada de los clientes

b) Régimen de servicio de los servidores

c) Disciplina de cola

d) Capacidad del sistema

e) Número de canales de servicio
f) Número de etapas de servicio

Algunos autores incluyen una séptima característica que es la población posible de clientes. El régimen de llegada de los clientes (a) habitualmente es estocástica, es decir, depende de una cierta variable aleatoria y será necesario conocer la distribución probabilística entre dos llegadas de clientes sucesivas. También es posible que los clientes arriben de una manera programada (un paciente cada 15 minutos) o que sean impacientes, es decir, que lleguen a la cola y si es demasiado larga se vayan, o que tras esperar un rato en la cola decidan abandonar. Por último, es posible que el patrón de llegada varíe con el tiempo. Si se mantiene constante se denomina estacionario. El número de arribos por unidad de tiempo frecuentemente se estima por medio de la Distribución de Poisson.

Los servidores pueden tener un tiempo de servicio (b) variable, en cuyo caso hay que asociarle, para definirlo, una función de probabilidad. Pueden atender en lotes o de modo individual. Los tiempos de servicio pueden modelarse como variables aleatorias idénticamente distribuidas, cuando todos los clientes son tratados de igual modo y, en este caso, el tiempo es independiente del número de clientes en la cola o del horario. Si varían con el número de clientes en la cola, trabajando más rápido o más lento, se denominan patrones de servicio dependientes, pudiendo ser también no-estacionarios.

La disciplina de la cola (c) es la manera en que los clientes se ordenan en el momento de ser servidos de entre los de la cola. Lo habitual es atender por orden de llegada (FIFO), sin embargo, pueden atenderse a la inversa, primero al último en llegar (LIFO). En cualquier caso si un cliente llega a la cola con una orden de prioridad superior al cliente que esta siendo atendido, éste se retira dando paso al más importante. Y es aquí donde aparecen dos nuevos subcasos: el cliente retirado vuelve a empezar o retorna donde se había quedado.

La capacidad del sistema (d) es el máximo número de clientes que pueden estar haciendo cola. Es finita cuando en el sistema existe una limitación respecto al número de clientes que pueden esperar en la cola. Esta limitación puede ser considerada como una simplificación en la modelización de la impaciencia de los clientes.

El número de canales de servicio (e) es el número de servidores que hay en el sistema. Puede haber una cola por cada servidor, sistema multicanal, o bien una única cola que alimenta a varios servidores, sistema mono-canal (Figura 2).

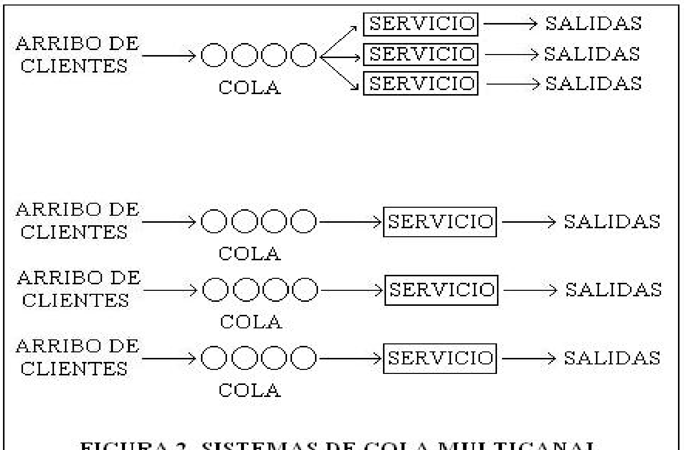

FIGURA 2. SISTEMAS DE COLA MULTICANAL

Un sistema de colas (f) puede ser unietapa o multietapa. En los sistemas unietapa, el cliente pasa por una fila de espera, es servido y luego abandona el sistema. En los sistemas multietapa el cliente pasa por varias filas de espera antes de retirarse del sistema. En algunos sistemas multietapa se puede admitir la *vuelta atrás* o *reciclado*, esto es habitual en sistemas productivos como controles de calidad y procesos (Figura 3).

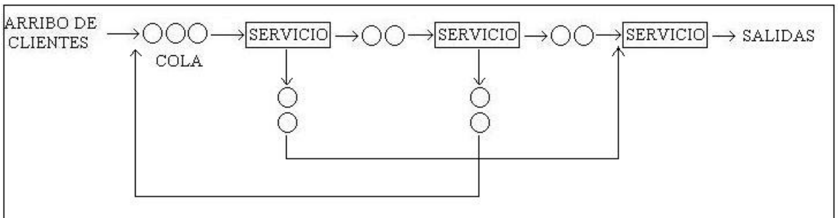

FIGURA 3. SISTEMA MULTIETAPA CON RETROALIMENTACIÓN.

Así, en general, los sistemas para el servicio son clasificados en función del número de canales (servidores) y el número de *fases* o *etapas* (número de paradas que deben hacerse durante el servicio), como se representan en la Figura 4.

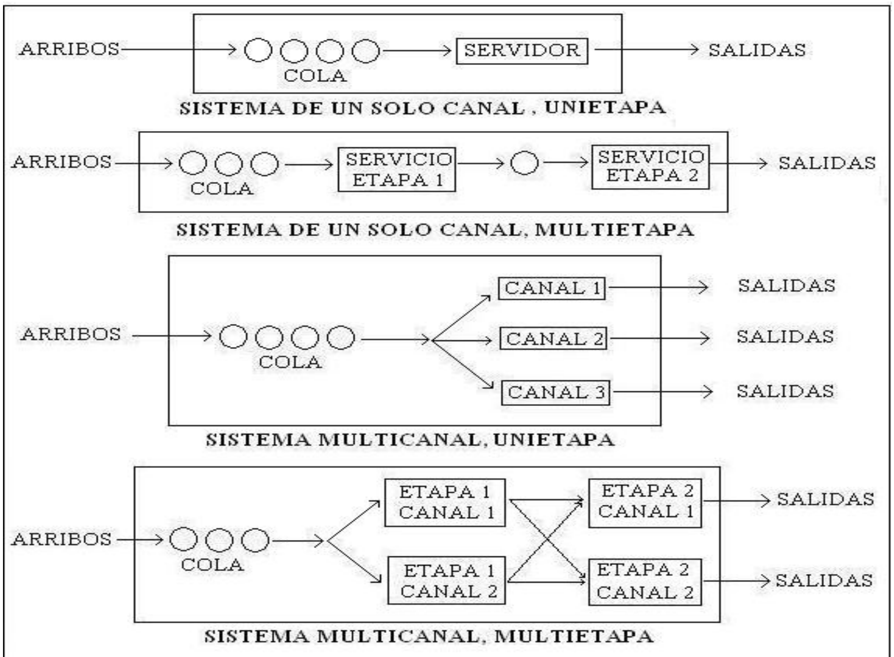

FIGURA 4

2.1 Simbología usual para representar un sistema de colas

Reconociendo la diversidad de los sistemas de colas, Kendall (1953) propuso un sistema de notación que consta de cinco símbolos separados por barras: A/B/X/Y/Z, donde:

A: Indica la distribución del tiempo entre llegadas consecutivas

B: Alude al patrón de servicio de los servidores

X: Es el número de canales de servicio

Y: Es la restricción en la capacidad del sistema

Z: Es la disciplina de cola

En la Tabla 1 se mencionan los símbolos con los que se representan las variantes clásicas utilizadas para cada integrante del sistema.

Si no existe restricción en la capacidad del sistema (Y = ∞) y la política de servicio es FIFO, no se suelen incorporar dichos símbolos en la notación, así, M/D/3 es equivalente a M/D/3/∞/FIFO y significa que los clientes entran según una distribución exponencial, se sirven de manera determinística con 3 servidores sin limitación de capacidad en el sistema y siguiendo una estrategia FIFO de servicio.

Tabla 1. Simbología de la notación

Característica	Símbolo	Explicación
	M	Exponencial
	D	Determinista
Distribución de los tiempos de llegada (A)	Ek	Erlang tipo k
o de servicio(B)	Hk	Mezcla de k exponenciales
	PH	Tipo Fase
	G	General (cualquiera)
Número de servidores (X)	1,2,…, ∞	
	FIFO	Servir al primero que llega
	LIFO	El último que llega se sirve primero
Disciplina de la cola (Z)	RSS	Selección aleatoria de servicio
	PR	Prioridad
	GD	Disciplina general

2.2 Cómo medir el rendimiento de un sistema

A partir de los datos que nos suministra la Teoría de Colas se puede obtener la información necesaria para tomar decisiones. Los tres aspectos que pueden resolverse son:

a) Tiempo medio de espera (en el total del sistema o en la cola)

b) Cantidad de clientes esperando (en el total del sistema o en la cola)

c) Tiempo ocioso de los servidores (total o particular de cada servicio)

Estos aspectos dependen de **parámetros del sistema.** Algunos de ellos deben ser estimados por observación del fenómeno y otros se deducen a partir de ellos:

λ: Nº de llegadas p/unidad de tiempo. Tasa de arribos.

μ: Nº servicios p/unidad de tiempo. Tasa de servicios.

c: Número de servidores en paralelo.

$\rho = \dfrac{\lambda}{c\mu}$: Congestión de un sistema con parámetros (λ, μ, c).

$N(t)$: Nº de clientes en el sistema en el instante t.

$N_q(t)$: Nº de clientes en la cola en el instante t.

$N_s(t)$: Nº de clientes en servicio en el instante t.

$N(t) = N_q(t) + N_s(t)$

N : Nº de clientes en el sistema en estado estable.

L : Nº medio de clientes en el sistema.

L_q : Nº medio de clientes en la cola.

T_q : Tiempo que un cliente invierte en la cola.

S : Tiempo de servicio.

$T = T_q + S$: Tiempo total que un cliente invierte en el sistema.

$W_q = E(T_q)$: Tiempo medio de espera de los clientes en la cola.

$W = E(T)$: Tiempo medio de estancia de los clientes en el sistema.

r : Nº medio de clientes que se atienden por término medio.

A partir de ellos pueden derivarse algunas probabilidades de interés como la Probabilidad de que haya n clientes en el sistema en el instante t ($P_n(t) = P(N(t) = n)$), o la Probabilidad de que haya n clientes en estado estable ($P_n = P(N = n)$). Si $\rho > 1$ el sistema tenderá a crecer inexorablemente. En los siguientes modelos de colas consideraremos $\rho < 1$

El ***número de clientes*** en el instante t, $N(t)$, es el número de llegadas que han ocurrido hasta t menos el número de servicios completados hasta t. El número medio de clientes en el sistema y en la cola se puede calcular de diferentes maneras:

$$L = E(N) = \sum_{n=0}^{\infty} nP_n$$

$$L_q = E(N_q) = \sum_{n=c+1}^{\infty} (n-c)P_n$$

Little establece una relación entre la longitud de la longitud de la cola y el tiempo de espera:

$$L = \lambda W$$

$$L_q = \lambda W_q$$

El tiempo de estancia de un cliente en el sistema se relaciona con el tiempo de espera de un cliente en la cola,

$$W = W_q + \frac{1}{\mu}$$

El número de clientes que por término medio se están atendiendo en cualquier momento es:

$$r = L - L_q = \lambda\left(W - W_q\right) = \frac{\lambda}{\mu}$$

y, en un sistema de un único servidor:

$$r = L - L_q = \sum_{n=1}^{\infty} nP_n - \sum_{n=1}^{\infty}(n-1)P_n = \sum_{n=1}^{\infty} P_n = 1 - P_0 \quad \text{y} \quad P_0 = 1 - \rho$$

La probabilidad de que un servidor (habiendo c servidores en paralelo) esté ocupado es:

$$P_b = \rho = \frac{\lambda}{c\mu}$$

2.3 Modelo de colas simples

2.3.1. Colas con servidores en paralelo M/M/c

Un sistema con servidores en paralelo se caracteriza porque hay más de un servidor que ejecuta la misma función con la misma eficiencia. Cuando se consideran c servidores en paralelo, se define $r = \lambda/\mu$ y la tasa de ocupación del sistema es $\rho = \lambda/c\mu$. En este caso, la tasa de llegadas es $\lambda e^{-\lambda t}$ y la de servicios es $\mu_n e^{-\mu_n t}$ donde

$$\mu_n = \begin{cases} n\mu & \text{si } 1 \le n < c \\ c\mu & \text{si } n \ge c \end{cases}$$

La probabilidad de que haya n clientes en un sistema de este tipo es:

$$P_n = \begin{cases} \dfrac{\lambda^n}{n!\mu^n} P_0 & \text{si } 1 \le n < c \\ \dfrac{\lambda^n}{c!c^{n-c}\mu^n} P_0 & \text{si } n \ge c \end{cases}$$

siendo la probabilidad de que el sistema esté vacío:

$$P_0 = \left(\sum_{n=0}^{c-1} \frac{r^n}{n!} + \frac{r^c}{c!(1-\rho)} \right)^{-1} \quad \text{donde} \quad \frac{r}{c} = \rho < 1$$

El número medio de clientes en la cola es:

$$L_q = \frac{r^c \rho}{c!(1-\rho)^2} P_0$$

El tiempo medio de espera de los clientes en la cola es:

$$W_q = \frac{L_q}{\lambda} = \frac{r^c}{c!c\mu(1-\rho)^2} P_0$$

Y, por lo tanto:

$$W = W_q + \frac{1}{\mu} = \frac{r^c}{c!c\mu(1-\rho)^2} P_0 + \frac{1}{\mu} \qquad \text{y} \qquad L = r + L_q = r + \frac{r^c}{c!(1-\rho)^2} P_0$$

2.3.2. Cola con un solo servidor M/G/1

En este modelo de cola, el tiempo entre llegadas es exponencial con una tasa media de llegada λ y los tiempos de servicio son independientes con la misma distribución de probabilidad, que puede ser cualquiera (G). En este caso hay un canal de servicio y una línea de espera. Las llegadas se distribuyen de acuerdo a una distribución de Poisson, al igual que en el caso anterior, pero los tiempos de servicio no necesariamente se distribuyen de acuerdo con la distribución exponencial negativa. Si conocemos la desviación estándar y la media de la distribución de los tiempos de servicio, G, puede obtenerse una fórmula para el valor de L_q a partir de la siguiente ecuación:

$$L_q = \frac{\lambda^2\sigma^2 + \rho^2}{2(1-\rho)} \qquad \text{Fórmula de Pollaczek-Khinchin}$$

Y luego podemos calcular L por medio de la siguiente ecuación: $L = L_q + \rho$

El tiempo que un cliente espera para ser atendido es $W_q = L_q / \lambda$ y el tiempo que espera en el sistema es $W = W_q + 1/\mu$.

Por último, la probabilidad de que haya n clientes en el sistema es: $P_n = \rho$ y $P_0 = 1 - \rho$

3. APLICACIÓN

3.1. Descripción del proceso administrativo de un expediente:

A los fines de describir un problema a través de la Teoría de Colas, es importante determinar claramente cuál es el sistema entendiendo por “sistema”, sus componentes y la forma en que ellos se relacionan (clientes, forma de arribo de los mismos, número de servidores, tiempo de servicios, etc).

En este trabajo, se considera un problema administrativo sobre la organización del recorrido de los expedientes, desde que ellos ingresan a un juzgado cuando son presentados por un abogado litigante en la mesa de entradas, hasta que el mismo sale del sistema con resolución del Juez (Miró, Depaoli, 2006), agregando al análisis propuesto por los autores, una mirada desde el punto de vista de la simulación.

Cada expediente que se ingresa al Juzgado en cuestión, puede estar iniciando un nuevo proceso, o corresponder a uno ya iniciado y es recibido por un abogado de la sala de despacho, quien lo analiza jurídicamente y genera un informe que se incorpora al mismo. En este momento comienza una nueva etapa, donde el Secretario (S) revisa el estado de la causa y lo pasa al Juez (J) para que éste tome una decisión.

Es posible que por errores administrativos, tanto cuando es ingresado el expediente como cuando entra a la sala de despacho, el expediente regrese a la etapa anterior para corregir los errores producidos.

El abogado litigante, en su nueva presentación, averiguará las novedades producidas y generará acciones que inicien un nuevo ciclo dentro del juzgado. Este proceso, que involucra tres etapas o fases bien diferenciadas, continuará hasta que haya una sentencia judicial, u otra conclusión de la causa. La Figura 5, que representa esquemáticamente todo el proceso que siguen los expedientes es lo que define el "sistema" de cola.

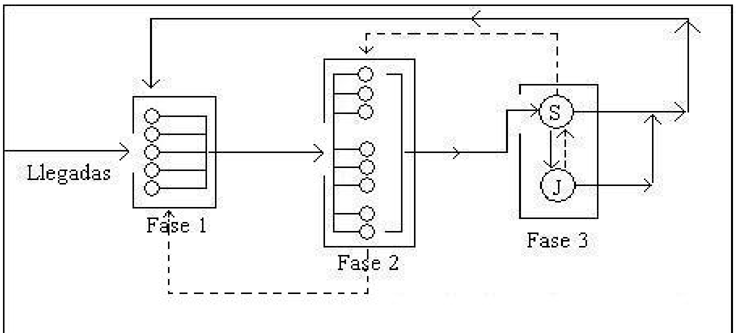

FIGURA 5

3.2 Uso de Teoría de Colas para describir el sistema administrativo del juzgado

Fase 1: Mesa de entradas

La mesa de entradas de un juzgado civil es básicamente un mostrador donde un cierto número de empleados brindan atención a los abogados litigantes, de acuerdo con el criterio FIFO, atendiendo por orden de llegada. Se considera como tiempo de servicio al intervalo de tiempo entre el instante en que el abogado ocupa un espacio libre del mostrador y el instante en que lo abandona, no necesitando la asistencia permanente de un empleado. En consecuencia, el número de empleados necesarios puede resultar menor que el número de lugares de servicio.

Sobre la base de las observaciones realizadas, esta fase puede modelarse como una cola M/M/c en donde los servidores son los c lugares de servicio (Sección 2.3.1). Los parámetros para la distribución de Poisson que deberán utilizarse para estimar las características del sistema, fueron obtenidos por los autores luego de un período de observación en un juzgado particular. Los datos obtenidos fueron:

$\lambda = 0.892$ expedientes/minutos y $\mu = 0.1953$ empleados/minutos

En base a estos valores, siguiendo los resultados de la Sección 2.3.1, se obtiene que:

- El coeficiente de ocupación es: $\rho = \dfrac{\lambda}{c\mu} = \dfrac{0.892}{0.1953\, c}$. Dado que el sistema es estable

si y sólo si $\rho < 1$, entonces el valor mínimo para c, número de servidores o empleados en el mostrador debe ser 5.

- La cantidad media de empleados en servicio es $r = \dfrac{\lambda}{\mu} = 4.567$.

- La probabilidad de que un cliente (expediente) llegue y el sistema esté vacío es:

$$P_0 = \left(\sum_{n=0}^{c-1} \frac{r^n}{n!} + \frac{r^c}{c!(1-\rho)} \right)^{-1} = 0.004$$

- El número medio de clientes en la cola es:

$$L_q = \frac{r^5 \rho}{5!(1-\rho)^2} P_0 = 8.4718$$

- El tiempo medio de los clientes en la cola es:

$$W_q = \frac{L_q}{\lambda} = \frac{r^5}{5!5\mu(1-\rho)^2} P_0 = 10.13 \text{ minutos}$$

- El tiempo medio de espera del cliente en la fase 1 del sistema es:

$$W = W_q + \frac{1}{\mu} = \frac{r^5}{5!5\mu(1-\rho)^2} P_0 + \frac{1}{\mu} = 14.62 \text{ minutos}$$

- La cantidad media de expedientes en la fase 1 del sistema es:

$$L = r + L_q = r + \frac{r^5}{5!(1-\rho)^2} P_0 = 13.04 \text{ expedientes}$$

Fase 2: Despacho

La fase 2 tiene la estructura de un sistema de una sola etapa, multicanal, con tasas de servicio diferentes para cada servidor. Esta fase no se estudia dado que en este trabajo no se detallan las colas de este tipo.

Fase 3: Toma de decisiones

Esta fase comprende el trabajo del juez y del secretario. Aquí nos limitamos a analizar la actividad de este último mediante un modelo M/G/1. El arribo de expedientes al escritorio puede aproximarse mediante un proceso de Poisson con tasa $\lambda = 1$. El secretario puede efectuar 3 tareas diferentes (ver Figura 6):

(1) tomar decisiones sobre la base del análisis efectuado en la fase 2,

(2) cuando la complejidad de la causa lo requiere, profundiza el análisis iniciado en la fase 2 o

(3) realiza una audiencia oficial con alguna o ambas partes en conflicto.

En esta última fase los parámetros para la distribución de Poisson que deberán utilizarse para estimar las características del sistema, fueron también obtenidos por los autores luego de un período de observación en ese mismo juzgado. Los datos obtenidos fueron:

- la tasa de entrada: $\lambda = 0.266019$ expedientes/minutos

- las probabilidades de ocurrencia de las tareas:

$P_1 = 0.982142$ probabilidad de ocurrencia de la tarea (1)

$P_2 = 0.008929$ probabilidad de ocurrencia de la tarea (2)

$P_3 = 0.008929$ probabilidad de ocurrencia de la tarea (3)

- las tasas parciales de servicio $\mu_i \quad (i = 1,2,3)$ de cada tarea:

$\mu_1 = 1.034482$ expedientes/minutos

$\mu_2 = 0.011111$ análisis/minutos

$\mu_3 = 0.008333$ audiencias/minutos

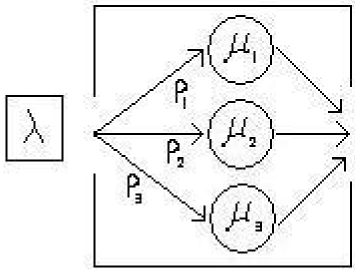

FIGURA 6. Flujo de expedientes en el escritorio del secretario.

Con estos datos se pueden calcular: la tasa global de servicio μ, el coeficiente de ocupación ρ, la cantidad media de expedientes en espera L_q, el tiempo medio de espera W_q de los expedientes y el índice de presión laboral. Suponemos que los tiempos de servicio para cada tarea del secretario tienen una distribución exponencial. Con esto, la distribución de los tiempos de servicio globales será hiper exponencial, con función densidad:

$$b(t) = \sum_{i=1}^{3} P_i \mu_i e^{-\mu_i t}$$

Dado el evento S = {El secretario comienza y termina su trabajo en un expediente antes del próximo arribo}, $P(S)$ es la probabilidad de que la pila de expedientes sobre el escritorio disminuya. Dentro del modelo M/G/1 existe una expresión cerrada para $P(S)$. Sea X la variable aleatoria que mide el tiempo de servicio del secretario, $b(t)$ su función de densidad e Y la variable aleatoria que registra el tiempo de arribo de un expediente, $P(X < Y)$ es la probabilidad $P(S)$. Como el ingreso es Poisson, $P(Y > t) = e^{-\lambda t}$. Las variables aleatorias X e Y son independientes; luego, la función de densidad conjunta de ambas será $g(u,v) = b(u)\lambda e^{-\lambda v}$. Entonces:

$$P(X < Y) = \int_{u \geq 0} \int_{v \geq 0} b(u)\lambda e^{-\lambda v} du dv = \int_0^\infty e^{-\lambda u} b(u) du = L(b)(\lambda) = P(S)$$

Esta integral es la Transformada de Laplace de la función densidad $b(t)$.

El número $I_p = 1 - P(S)$ mide la probabilidad de que la cola permanezca igual o se incremente (también denominado, por esta razón, índice de **presión laboral**) y la forma que adopta depende de la función de distribución de los tiempos de servicio de cada modelo. Por ejemplo: en una cola M/M/1, la función densidad de los tiempos de servicio es exponencial, por lo cual

$$I_p = 1 - P(S) = 1 - L(b)(\lambda) = 1 - \frac{1}{1+\rho}$$

- El tiempo medio de servicio del secretario es: $\bar{t} = \sum_{i=1}^{3} P_i \mu_i$

- El índice de presión laboral es: $I_p = 1 - L(b)(\lambda) = 1 - \sum_{i=1}^{3} \frac{P_i \mu_i}{\mu_i + \lambda}$

En el caso M/G/1, la longitud media de la cola L_q, está dada por la fórmula de Pollaczek-Khinchin: $L_q = \frac{\rho^2 \left(C^2+1\right)}{2(1-\rho)}$, donde C denota el coeficiente de variación $C^2 = \frac{\mathrm{Var}(t)}{\bar{t}^2}$. Y el tiempo medio de permanencia en la cola W_q está dado por la Ley de Little: $L_q = W_q \lambda$.

A continuación podemos observar los resultados numéricos obtenidos por el modelo para el escritorio del secretario:

$\mu = 0.354040$ tareas/minutos (μ global)

$\rho = 0.75$ (Coeficiente de ocupación)

$C^2 = 49.59$ (Coeficiente de variación)

$L_q = 57.44$ expedientes (Número medio de expedientes en la cola)

$W_q = 215.95$ minutos (Tiempo medio de espera de los expedientes en la cola)

$I_p = 0.22$ (Probabilidad de que la cola permanezca igual o se incremente)

El valor alto obtenido para el coeficiente de variación se debe a que las tareas de servicio parciales son muy diferentes entre sí. Esto significa que si el Secretario está a cargo de distintas tareas con tiempos de servicios muy dispares, la cola de expedientes en su escritorio será grande pese a su eficiencia. Entonces, a igualdad de recursos, las tareas deben distribuirse de manera que la dispersión de los tiempos de servicio sea mínima en cada servidor.

Partiendo de estos resultados, que están dados como determinísticos, se está trabajando desde el punto de vista de la simulación de los tiempos de servicio, a fin de agregar un enfoque inferencial, calculando intervalos de confianza por bootstrap, que permitan medir precisión alrededor de los estimadores obtenidos.

4. COMENTARIOS FINALES

La Teoría de Colas, cuya génesis proviene de muchos años atrás, sigue teniendo amplia vigencia para el estudio de características vitales en la organización de sistemas cuyas componentes no son determinísticas sino aleatorias. Si bien la teoría matemática subyacente no es simple, el estudio de sus propiedades permite abordar problemas complejos de una forma sistemática, con resultados muy ricos desde el punto de vista práctico. Las áreas de aplicación son muy amplias, abarcando desde aspectos comerciales, de organización en empresas, tanto en sectores productivos como administrativos o administración de servicios de salud, educativos o de gobierno, hasta en la informática u otros aspectos de la tecnología.

En particular en este trabajo, se ha retomado una aplicación que se refiere a la organización en un Juzgado, planteando la posibilidad de incorporar un enfoque diferente a partir del uso de simulación. Este procedimiento, permitiría agregar información valiosa a los resultados obtenidos por la Teoría de Colas, referidos a la precisión de los mismos, considerando la aleatoriedad de las estimaciones en base a las cuales fueron derivados. La simulación de los sistemas que representan las distintas fases administrativas del Juzgado no son triviales, pero el valor agregado de sus resultados a la información obtenida es muy útil desde el punto de vista práctico.

REFERENCIAS BIBLIOGRÁFICAS

__ Gross, Donald; Harris, Carl M. (1998) "Fundamentals of Queueing Theory".

Editorial: Wiley-Interscience, 3ª Edition.

__ Kleinrock, Leonard (1975) "Queueing Systems" Volume I: Theory. Editorial: Wiley – Interscience.

__ Miró, Ricardo; Depaoli, Roberto (2006) "La teoría de las colas en el modelado de un Juzgado Nacional Civil". Comunicación efectuada en la sesión privada extraordinaria de la Academia Nacional de Ciencias de Buenos Aires.