Temas Selectos de Probabilidad

Carlos E. Martínez-Rodríguez

Julio 2024

Contents

Part I

TERCERA PARTE: TEMAS SELECTOS

CAPÍTULO 1

CADENAS DE MARKOV

1.1 Estacionareidad

Sea $v = (v_i)_{i\in E}$ medida no negativa en E, podemos definir una nueva medida $v\mathbb{P}$ que asigna masa $\sum_{i\in E} v_i p_{ij}$ a cada estado j.

Definición 1.1.1. *La medida v es estacionaria si $v_i < \infty$ para toda i y además $v\mathbb{P} = v$.*

En el caso de que v sea distribución, independientemente de que sea estacionaria o no, se cumple con

$$\mathbb{P}_v\left[X_1 = j\right] = \sum_{i\in E} \mathbb{P}_v\left[X_0 = i\right] p_{ij} = \sum_{i\in E} v_i p_{ij} = (vP)_j$$

Teorema 1.1.1. *Supongamos que v es una distribución estacionaria. Entonces*

i) La cadena es estrictamente estacionaria con respecto a $\mathbb{P}_v$, es decir, $\mathbb{P}_v$-distribución de $\{X_n, X_{n+1}, \ldots\}$ no depende de n;

ii) Existe un aversión estrictamente estacionaria $\{X_n\}_{n\in Z}$ de la cadena con doble tiempo infinito y $\mathbb{P}(X_n = i) = v_i$ para toda $n \in Z$.

Teorema 1.1.2. *Sea i estado fijo, recurrente. Entonces una medida estacionaria v puede definirse haciendo que v_j sea el número esperado de visitas a j entre dos visitas consecutivas i,*

$$v_j = \mathbb{E}_i \sum_{n=0}^{\tau(i)-1} \mathbb{1}\left(X_n = i\right) = \sum_{n=0}^{\infty} \mathbb{P}_i\left[X_n = j, \tau(i) > n\right] \tag{1.1}$$

Teorema 1.1.3. *Si la cadena es irreducible y recurrente, entonces existe una medida estacionaria v, tal que satisface $0 < v_j < \infty$ para toda j, y es única salvo factores multiplicativos, es decir, si v, v^* son estacionarias, entonces $c = cv^*$ para alguna $c \in (0, \infty)$.*

Corolario 1.1.1. *Si la cadena es irreducible y positiva recurrente, existe una única distribución estacionaria π dada por*

$$\pi_j = \frac{1}{\mathbb{E}_i \tau_i} \mathbb{E}_i \sum_{n=0}^{\tau(i)-1} \mathbb{1}\left(X_n = j\right) = \frac{1}{\mathbb{E}_j \tau\left(j\right)}. \tag{1.2}$$

Corolario 1.1.2. *Cualquier cadena de Markov irreducible con un espacio de estados finito es positiva recurrente.*

1.2 Funciones Armónicas, Recurrencia y Transitoriedad

Definición 1.2.1. *Una función Armónica es el eigenvector derecho h de P correspondiente al eigenvalor 1.*

$$Ph = h \Leftrightarrow h(i) = \sum_{j \in E} p_{ij} h(j) = \mathbb{E}_i h(X_1) = \mathbb{E}\left[h(X_{n+1}) | X_n = i\right].$$

es decir, $\{h(X_n)\}$ es martingala.

Proposición 1.2.1. *Sea $\{X_n\}$ cadena irreducible y sea i estado fijo arbitrario. Entonces la cadena es transitoria sí y sólo si existe una función no cero, acotada $h : E - \{i\} \to \mathbb{R}$ que satisface*

$$h(j) = \sum_{k \neq i} p_{jk} h(k) \;\; \textit{para } j \neq i. \tag{1.3}$$

Proposición 1.2.2. *Suponga que la cadena es irreducible y sea E_0 un subconjunto finito de E tal que se cumple la ecuación 5.2 para alguna función h acotada que satisface $h(i) < h(j)$ para algún estado $i \notin E_0$ y todo $j \in E_0$. Entonces la cadena es transitoria.*

1.3 Teoría Ergódica

Lema 1.3.1. *Sea $\{X_n\}$ cadena irreducible y se F subconjunto finito del espacio de estados. Entonces la cadena es positiva recurrente si $\mathbb{E}_i \tau(F) < \infty$ para todo $i \in F$.*

Proposición 1.3.1. *Sea $\{X_n\}$ cadena irreducible y transiente o cero recurrente, entonces $p_{ij}^n \to 0$ conforme $n \to \infty$ para cualquier $i, j \in E$, E espacio de estados.*

Utilizando el teorema (2.2) y el corolario 1.1.1, se demuestra el siguiente resultado importante.

Teorema 1.3.1. *Sea $\{X_n\}$ cadena irreducible y aperiódica positiva recurrente, y sea $\pi = \{\pi_j\}_{j \in E}$ la distribución estacionaria. Entonces $p_{ij}^n \to \pi_j$ para todo i, j.*

Definición 1.3.1. *Una cadena irreducible aperiodica, positiva recurrente con medida estacionaria v, es llamada* ergódica.

Proposición 1.3.2. *Sea $\{X_n\}$ cadena irreducible y recurrente con medida estacionaria v, entocnes para todo $i, j, k, l \in E$*

$$\frac{\sum_{n=0}^{m} p_{ij}^n}{\sum_{n=0}^{m} p_{lk}^n} \to \frac{v_j}{v_k},\; m \to \infty \tag{1.4}$$

Lema 1.3.2. *La matriz $\widetilde{P}$ con elementos $\widetilde{p}_{ij} = \frac{v_{ji} p_{ji}}{v_i}$ es una matriz de transición. Ademś, el i-ésimo elementos $\widetilde{p}_{ij}^m$ de la matriz potencia $\widetilde{P}^m$ est? dada por $\widetilde{p}_{ij}^m = \frac{v_{ji} p_{ji}^m}{v_i}$.*

Lema 1.3.3. *Defínase $N_i^m = \sum_{n=0}^{m} \mathbb{1}(X_n = i)$ como el número de visitas a i antes del tiempo m. Entonces si la cadena es reducible y recurrente, $lim_{m \to \infty} \frac{\mathbb{E}_j N_i^m}{\mathbb{E}_k N_i^m} = 1$ para todo $j, k \in E$.*

1.3.1 Ejemplo de Cadena de Markov para dos Estados

Supongamos que se tiene la siguiente cadena:

$$\begin{pmatrix} 1-q & q \\ p & 1-p \end{pmatrix}. \tag{1.5}$$

Si $P[X_0 = 0] = \pi_0(0) = a$ y $P[X_0 = 1] = \pi_0(1) = b = 1 - \pi_0(0)$, con $a + b = 1$, entonces después de un procedimiento más o menos corto se tiene que:

$$P[X_n = 0] = \frac{p}{p+q} + (1-p-q)^n \left(a - \frac{p}{p+q}\right).$$
$$P[X_n = 1] = \frac{q}{p+q} + (1-p-q)^n \left(b - \frac{q}{p+q}\right).$$

donde, como $0 < p, q < 1$, se tiene que $|1-p-q| < 1$, entonces $(1-p-q)^n \rightarrow 0$ cuando $n \rightarrow \infty$. Por lo tanto

$$lim_{n\rightarrow\infty} P\left[X_n = 0\right] = \frac{p}{p+q}.$$
$$lim_{n\rightarrow\infty} P\left[X_n = 1\right] = \frac{q}{p+q}.$$

Si hacemos $v = \left(\frac{p}{p+q}, \frac{q}{p+q}\right)$, entonces

$$\left(\frac{p}{p+q}, \frac{q}{p+q}\right)\left(\begin{array}{cc} 1-q & q \\ p & 1-p \end{array}\right).$$

Proposición 1.3.3. *Suponga que la cadena es irreducible y sea E_0 un subconjunto finito de E tal que se cumple la ecuación 5.2 para alguna función h acotada que satisface $h(i) < h(j)$ para algún estado $i \notin E_0$ y todo $j \in E_0$. Entonces la cadena es transitoria.*

1.4 Procesos de Markov de Saltos

Consideremos un estado que comienza en el estado x_0 al tiempo 0, supongamos que el sistema permanece en x_0 hasta algún tiempo positivo τ_1, tiempo en el que el sistema salta a un nuevo estado $x_1 \neq x_0$. Puede ocurrir que el sistema permanezca en x_0 de manera indefinida, en este caso hacemos $\tau_1 = \infty$. Si τ_1 es finito, el sistema permanecerá en x_1 hasta τ_2, y así sucesivamente. Sea

$$X(t) = \left\{\begin{array}{cc} x_0 & 0 \leq t < \tau_1 \\ x_1 & \tau_1 \leq t < \tau_2 \\ x_2 & \tau_2 \leq t < \tau_3 \\ \vdots & \end{array}\right. \tag{1.6}$$

A este proceso se le llama *proceso de salto.* Si

$$lim_{n\rightarrow\infty}\tau_n = \left\{\begin{array}{cc} < \infty & X_t \text{ explota} \\ = \infty & X_t \text{ no explota} \end{array}\right. \tag{1.7}$$

Un proceso puro de saltos es un proceso de saltos que satisface la propiedad de Markov.

Proposición 1.4.1. *Un proceso de saltos es Markoviano si y sólo si todos los estados no absorbentes x son tales que*

$$P_x\left(\tau_1 > t+s | \tau_1 > s\right) = P_x\left(\tau_1 > t\right)$$

para $s, t \geq 0$, equivalentemente

$$\frac{1 - F_x(t+s)}{1 - F_x(s)} = 1 - F_x(t). \tag{1.8}$$

Nota 1.4.1. *Una distribución F_x satisface la ecuación (1.8) si y sólo si es una función de distribución exponencial para todos los estados no absorbentes x.*

Por un proceso de nacimiento y muerte se entiende un proceso de Markov de Saltos, $\{X_t\}_{t\geq 0}$ en $E = \mathbb{N}$ tal que del estado n sólo se puede mover a $n-1$ o $n+1$, es decir, la matriz intensidad es de la forma:

$$\Lambda = \left(\begin{array}{ccccc} -\beta_0 & \beta_0 & 0 & 0 & \ldots \\ \delta_1 & -\beta_1 - \delta_1 & \beta_1 & 0 & \ldots \\ 0 & \delta_2 & -\beta_2 - \delta_2 & \beta_2 & \ldots \\ \vdots & & & \ddots & \end{array}\right) \tag{1.9}$$

donde β_n son las probabilidades de nacimiento y δ_n las probabilidades de muerte.

La matriz de transición es

$$Q = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots \\ q_1 & 0 & p_1 & 0 & \cdots \\ 0 & q_2 & 0 & p_2 & \cdots \\ \vdots & & & \ddots & \end{pmatrix} \tag{1.10}$$

con $p_n = \frac{\beta_n}{\beta_n+\delta_n}$ y $q_n = \frac{\delta_n}{\beta_n+\delta_n}$

Proposición 1.4.2. *La recurrencia de un Proceso Markoviano de Saltos $\{X_t\}_{t\geq 0}$ con espacio de estados numerable, o equivalentemente de la cadena encajada $\{Y_n\}$ es equivalente a*

$$\sum_{n=1}^{\infty} \frac{\delta_1 \cdots \delta_n}{\beta_1 \cdots \beta_n} = \sum_{n=1}^{\infty} \frac{q_1 \cdots q_n}{p_1 \cdots p_n} = \infty \tag{1.11}$$

Lema 1.4.1. *Independientemente de la recurrencia o transitoriedad de la cadena, hay una y sólo una, salvo múltiplos, solución ν a $\nu\Lambda = 0$, dada por*

$$\nu_n = \frac{\beta_0 \cdots \beta_{n-1}}{\delta_1 \cdots \delta_n} \nu_0 \tag{1.12}$$

Corolario 1.4.1. *En el caso recurrente, la medida estacionaria μ para $\{Y_n\}$ está dada por*

$$\mu_n = \frac{p_1 \cdots p_{n-1}}{q_1 \cdots q_n} \mu_0 \tag{1.13}$$

para $n = 1, 2, \ldots$

Definición 1.4.1. *Una medida ν es estacionaria si $0 \leq \nu_j < \infty$ y para toda t se cumple que $\nu P^t = nu$.*

Definición 1.4.2. *Un proceso irreducible recurrente con medida estacionaria con masa finita es llamado ergódico.*

Teorema 1.4.1. *Un Proceso de Saltos de Markov irreducible no explosivo es ergódico si y sólo si uno puede encontrar una solución π de probabilidad, $|\pi| = 1$, $0 \leq \pi_j \leq 1$ para $\nu\Lambda = 0$. En este caso π es la distribución estacionaria.*

Corolario 1.4.2. *$\{X_t\}_{t\geq 0}$ es ergódica si y sólo si (1.21) se cumple y $S < \infty$, en cuyo caso la distribución estacionaria π está dada por*

$$\pi_0 = \frac{1}{S},\ \pi_n = \frac{1}{S} \frac{\beta_0 \cdots \beta_{n-1}}{\delta_1 \cdots \delta_n},\ n = 1, 2, \ldots \tag{1.14}$$

Sea E espacio discreto de estados, finito o numerable, y sea $\{X_t\}$ un proceso de Markov con espacio de estados E. Una medida μ en E definida por sus probabilidades puntuales μ_i, escribimos $p_{ij}^t = P^t(i, \{j\}) = P_i(X_t = j)$.

El monto del tiempo gastado en cada estado es positivo, de modo tal que las trayectorias muestrales son constantes por partes. Para un proceso de saltos denotamos por los tiempos de saltos a $S_0 = 0 < S_1 < S_2 \cdots$, los tiempos entre saltos consecutivos $T_n = S_{n+1} - S_n$ y la secuencia de estados visitados por $Y_0, Y_1, \ldots$, así las trayectorias muestrales son constantes entre S_n consecutivos, continua por la derecha, es decir, $X_{S_n} = Y_n$.

La descripción de un modelo práctico está dado usualmente en términos de las intensidades $\lambda(i)$ y las probabilidades de salto q_{ij} más que en términos de la matriz de transición P^t. Supóngase de ahora en adelante que $q_{ii} = 0$ cuando $\lambda(i) > 0$

Teorema 1.4.2. *Cualquier Proceso de Markov de Saltos satisface la Propiedad Fuerte de Markov*

Definición 1.4.3. *Una medida $v \neq 0$ es estacionaria si $0 \leq v_j < \infty$, $vP^t = v$ para toda t.*

Teorema 1.4.3. *Supongamos que $\{X_t\}$ es irreducible recurrente en E. Entonces existe una y sólo una, salvo múltiplos, medida estacionaria v. Esta v tiene la propiedad de que $0 < v_j < \infty$ para todo j y puede encontrarse en cualquiera de las siguientes formas* *i) Para algún estado i, fijo pero arbitrario, v_j es el tiempo esperado utilizado en j entre dos llegadas consecutivas al estado i;*

$$v_j = \mathbb{E}_i \int_0^{w(i)} \mathbb{1}\left(X_t = j\right) dt \tag{1.15}$$

con $w\left(i\right) = \inf\left\{t > 0 : X_t = i, X_{t^-} = \lim_{s \uparrow t} X_s \neq i\right\}$.

ii) $v_j = \frac{\mu_j}{\lambda(j)}$, donde μ es estacionaria para $\left\{Y_n\right\}$.

iii) como solución de $v\Lambda = 0$.

Definición 1.4.4. *Un proceso irreducible recurrente con medida estacionaria de masa finita es llamado ergódico.*

Teorema 1.4.4. *Un proceso de Markov de saltos irreducible no explosivo es ergódico si y sólo si se puede encontrar una solución, de probabilidad, π, con $|\pi| = 1$ y $0 \leq \pi_j \leq 1$, a $\pi\Lambda = 0$. En este caso π es la distribución estacionaria.*

Corolario 1.4.3. *Una condición suficiente para la ergodicidad de un proceso irreducible es la existencia de una probabilidad π que resuelva el sistema $\pi\Lambda = 0$ y que además tenga la propiedad de que $\sum \pi_j \lambda\left(j\right)$.*

1.5 Matriz Intensidad

Definición 1.5.1. *La matriz intensidad $\Lambda = \left(\lambda\left(i,j\right)\right)_{i,j \in E}$ del proceso de saltos $\left\{X_t\right\}_{t \geq 0}$ está dada por*

$$\begin{aligned} \lambda\left(i,j\right) &= \lambda\left(i\right) q_{i,j},\ j \neq i \\ \lambda\left(i,i\right) &= -\lambda\left(i\right) \end{aligned}$$

Proposición 1.5.1. *Una matriz $E \times E$, Λ es la matriz de intensidad de un proceso markoviano de saltos $\left\{X_t\right\}_{t \geq 0}$ si y sólo si*

$$\lambda\left(i,i\right) \leq 0,\ \lambda\left(i,j\right),\ i \neq j,\ \sum_{j \in E} \lambda\left(i,j\right) = 0.$$

Además, Λ está en correspondencia uno a uno con la distribución del proceso minimal dado por el teorema 3.1.

Para el caso particular de la Cola $M/M/1$, la matríz de itensidad está dada por

$$\Lambda = \begin{bmatrix} -\beta & \beta & 0 & 0 & 0 & \cdots \\ \delta & -\beta-\delta & \beta & 0 & 0 & \cdots \\ 0 & \delta & -\beta-\delta & \beta & 0 & \cdots \\ \vdots & & & & & \ddots \end{bmatrix}$$

1.6 Medidas Estacionarias

Definición 1.6.1. *Una medida $v \neq 0$ es estacionaria si $0 \leq v_j < \infty$, $vP^t = v$ para toda t.*

Teorema 1.6.1. *Supongamos que $\left\{X_t\right\}$ es irreducible recurrente en E. Entonces existe una y sólo una, salvo múltiplos, medida estacionaria v. Esta v tiene la propiedad de que $0 < v_j < \infty$ para todo j y puede encontrarse en cualquiera de las siguientes formas*

i) Para algún estado i, fijo pero arbitrario, v_j es el tiempo esperado utilizado en j entre dos llegadas consecutivas al estado i;

$$v_j = \mathbb{E}_i \int_0^{w(i)} \mathbb{1}\left(X_t = j\right) dt \tag{1.16}$$

con $w\left(i\right) = \inf\left\{t > 0 : X_t = i, X_{t^-} = \lim_{s \uparrow t} X_s \neq i\right\}$.

ii) $v_j = \frac{\mu_j}{\lambda(j)}$, donde μ es estacionaria para $\left\{Y_n\right\}$.

iii) como solución de $v\Lambda = 0$.

1.7 Criterios de Ergodicidad

Definición 1.7.1. *Un proceso irreducible recurrente con medida estacionaria de masa finita es llamado ergódico.*

Teorema 1.7.1. *Un proceso de Markov de saltos irreducible no explosivo es ergódico si y sólo si se puede encontrar una solución, de probabilidad, π, con $|\pi| = 1$ y $0 \leq \pi_j \leq 1$, a $\pi\Lambda = 0$. En este caso π es la distribución estacionaria.*

Corolario 1.7.1. *Una condición suficiente para la ergodicidad de un proceso irreducible es la existencia de una probabilidad π que resuelva el sistema $\pi\Lambda = 0$ y que además tenga la propiedad de que $\sum \pi_j \lambda(j) < \infty$.*

Proposición 1.7.1. *Si el proceso es ergódico, entonces existe una versión estrictamente estacionaria $\{X_t\}_{-\infty<t<\infty}$con doble tiempo infinito.*

Teorema 1.7.2. *Si $\{X_t\}$ es ergódico y π es la distribución estacionaria, entonces para todo i, j, $p_{ij}^t \rightarrow \pi_j$ cuando $t \rightarrow \infty$.*

Corolario 1.7.2. *Si $\{X_t\}$ es irreducible recurente pero no ergódica, es decir $|v| = \infty$, entonces $p_{ij}^t \rightarrow 0$ para todo $i, j \in E$.*

Corolario 1.7.3. *Para cualquier proceso Markoviano de Saltos minimal, irreducible o no, los límites $li_{t\rightarrow\infty} p_{ij}^t$ existe.*

1.8 Procesos de Nacimiento y Muerte

Proposición 1.8.1. *La recurrencia de $\{X_t\}$, o equivalentemente de $\{Y_n\}$ es equivalente a*

$$\sum_{n=1}^{\infty} \frac{\delta_1 \cdots \delta_n}{\beta_1 \cdots \beta_n} = \sum_{n=1}^{\infty} \frac{q_1 \cdots q_n}{p_1 \cdots p_n} = \infty \tag{1.17}$$

Lema 1.8.1. *Independientemente de la recurrencia o transitorieadad, existe una y sólo una, salvo múltiplos, solución a $v\Lambda = 0$, dada por*

$$v_n = \frac{\beta_0 \cdots \beta_{n-1}}{\delta_1 \cdots \delta_n} v_0 \tag{1.18}$$

para $n = 1, 2, \ldots$.

Corolario 1.8.1. *En el caso recurrente, la medida estacionaria μ para $\{Y_n\}$ está dada por*

$$\mu_n = \frac{p_1 \cdots p_{n-1}}{q_1 \cdots q_n} \mu_0 \tag{1.19}$$

para $n = 1, 2, \ldots$.

Se define a $S = 1 + \sum_{n=1}^{\infty} \frac{\beta_0\beta_1\cdots\beta_{n-1}}{\delta_1\delta_2\cdots\delta_n}$

Corolario 1.8.2. *$\{X_t\}$ es ergódica si y sólo si la ecuación (1.21) se cumple y además $S < \infty$, en cuyo caso la distribución ergódica, π, está dada por*

$$\pi_0 = \frac{1}{S},\ \pi_n = \frac{1}{S} \frac{\beta_0 \cdots \beta_{n-1}}{\delta_1 \cdots \delta_n} \tag{1.20}$$

para $n = 1, 2, \ldots$.

1.9 Procesos de Nacimiento y Muerte Generales

Por un proceso de nacimiento y muerte se entiende un proceso de saltos de markov $\{X_t\}_{t\geq 0}$ con espacio de estados a lo más numerable, con la propiedad de que sólo puede ir al estado $n+1$ o al estado $n-1$, es decir, su matriz de intensidad es de la forma

$$\Lambda = \begin{bmatrix} -\beta_0 & \beta_0 & 0 & 0 & 0 & \cdots \\ \delta_1 & -\beta_1 - \delta_1 & \beta_1 & 0 & 0 & \cdots \\ 0 & \delta_2 & -\beta_2 - \delta_2 & \beta_2 & 0 & \cdots \\ \vdots & & & & & \ddots \end{bmatrix}$$

donde β_n son las intensidades de nacimiento y δ_n las intensidades de muerte, o también se puede ver como a X_t el número de usuarios en una cola al tiempo t, un salto hacia arriba corresponde a la llegada de un nuevo usuario y un salto hacia abajo como al abandono de un usuario después de haber recibido su servicio.

La cadena de saltos $\{Y_n\}$ tiene matriz de transición dada por

$$Q = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & \cdots \\ q_1 & 0 & p_1 & 0 & 0 & \cdots \\ 0 & q_2 & 0 & p_2 & 0 & \cdots \\ \vdots & & & & & \ddots \end{bmatrix}$$

donde $p_n = \frac{\beta_n}{\beta_n + \delta_n}$ y $q_n = 1 - p_n = \frac{\delta_n}{\beta_n + \delta_n}$, donde además se asumne por el momento que p_n no puede tomar el valor 0 ó 1 para cualquier valor de n.

Proposición 1.9.1. *La recurrencia de $\{X_t\}$, o equivalentemente de $\{Y_n\}$ es equivalente a*

$$\sum_{n=1}^{\infty} \frac{\delta_1 \cdots \delta_n}{\beta_1 \cdots \beta_n} = \sum_{n=1}^{\infty} \frac{q_1 \cdots q_n}{p_1 \cdots p_n} = \infty \tag{1.21}$$

Lema 1.9.1. *Independientemente de la recurrencia o transitorieadad, existe una y sólo una, salvo múltiplos, solución a $v\Lambda = 0$, dada por*

$$v_n = \frac{\beta_0 \cdots \beta_{n-1}}{\delta_1 \cdots \delta_n} v_0 \tag{1.22}$$

para $n = 1, 2, \ldots$.

Corolario 1.9.1. *En el caso recurrente, la medida estacionaria μ para $\{Y_n\}$ está dada por*

$$\mu_n = \frac{p_1 \cdots p_{n-1}}{q_1 \cdots q_n} \mu_0 \tag{1.23}$$

para $n = 1, 2, \ldots$.

Se define a $S = 1 + \sum_{n=1}^{\infty} \frac{\beta_0 \beta_1 \cdots \beta_{n-1}}{\delta_1 \delta_2 \cdots \delta_n}$.

Corolario 1.9.2. *$\{X_t\}$ es ergódica si y sólo si la ecuación (1.21) se cumple y además $S < \infty$, en cuyo caso la distribución ergódica, π, está dada por*

$$\pi_0 = \frac{1}{S}, \ \pi_n = \frac{1}{S} \frac{\beta_0 \cdots \beta_{n-1}}{\delta_1 \cdots \delta_n} \tag{1.24}$$

para $n = 1, 2, \ldots$.

1.10 Notación Kendall-Lee

A partir de este momento se harán las siguientes consideraciones:

a) Si t_n es el tiempo aleatorio en el que llega al sistema el n-ésimo cliente, para $n = 1, 2, \ldots$, $t_0 = 0$ y $t_0 < t_1 < \cdots$ se definen los tiempos entre arribos $\tau_n = t_n - t_{n-1}$ para $n = 1, 2, \ldots$, variables aleatorias independientes e idénticamente distribuidas.

b) Los tiempos entre arribos tienen un valor medio $E(\tau)$ finito y positivo $\frac{1}{\beta}$, es decir, β se puede ver como la tasa o intensidad promedio de arribos al sistema por unidad de tiempo.

c) Además se supondrá que los servidores son identicos y si s denota la variable aleatoria que describe el tiempo de servicio, entonces $E(s) = \frac{1}{\delta}$, δ es la tasa promedio de servicio por servidor.

La notación de Kendall-Lee es una forma abreviada de describir un sistema de espera con las siguientes componentes:

a) **Fuente**: Población de clientes potenciales del sistema, esta puede ser finita o infinita.

b) **Proceso de Arribos**: Proceso determinado por la función de distribución $A(t) = P\{\tau \leq t\}$ de los tiempos entre arribos.

Además tenemos las siguientes igualdades

$$N(t) = N_q(t) + N_s(s) \tag{1.25}$$

donde

- $N(t)$ es el número de clientes en el sistema al tiempo t.
- $N_q(t)$ es el número de cliente en la cola al tiempo t.
- $N_s(t)$ es el número de clientes recibiendo servicio en el tiempo t.

Bajo la hipótesis de estacionareidad, es decir, las características de funcionamiento del sistema se han estabilizado en valores independientes del tiempo, entonces

$$N = N_q + N_s. \tag{1.26}$$

Los valores medios de las cantidades anteriores se escriben como $L = E(N)$, $L_q = E(N_q)$ y $L_s = E(N_s)$, entonces de la ecuación 1.25 se obtiene

$$L = L_q + L_s \tag{1.27}$$

Si q es el tiempo que pasa un cliente en la cola antes de recibir servicio, y W es el tiempo total que un cliente pasa en el sistema, entonces

$$w = q + s$$

por lo tanto

$$W = W_q + W_s,$$

donde $W = E(w)$, $W_q = E(q)$ y $W_s = E(s) = \frac{1}{\delta}$.

La intensidad de tráfico se define como

$$\rho = \frac{E(s)}{E(\tau)} = \frac{\beta}{\delta}. \tag{1.28}$$

La utilización por servidor es

$$u = \frac{\rho}{c} = \frac{\beta}{c\delta}. \tag{1.29}$$

donde c es el número de servidores.

Esta notación es una forma abreviada de describir un sistema de espera con componentes dados a continuación, la notación es

$$A/S/c/K/F/d \tag{1.30}$$

Cada una de las letras describe:

- A es la distribución de los tiempos entre arribos.
- S es la distribución del tiempo de servicio.
- c es el número de servidores.
- K es la capacidad del sistema.
- F es el número de individuos en la fuente.
- d es la disciplina del servicio

Usualmente se acostumbra suponer que $K = \infty$, $F = \infty$ y $d = FIFO$, es decir, First In First Out. Las distribuciones usuales para A y B son:

- GI para la distribución general de los tiempos entre arribos.
- G distribución general del tiempo de servicio.
- M Distribución exponencial para A o S.
- E_K Distribución Erlang-K, para A o S.
- D tiempos entre arribos o de servicio constantes, es decir, deterministicos.

1.10.1 Cola $M/M/1$

Este modelo corresponde a un proceso de nacimiento y muerte con $\beta_n = \beta$ y $\delta_n = \delta$ independiente del valor de n. La intensidad de tráfico $\rho = \frac{\beta}{\delta}$, implica que el criterio de recurrencia (ecuación 1.21) quede de la forma:

$$1 + \sum_{n=1}^{\infty} \rho^{-n} = \infty.$$

Equivalentemente el proceso es recurrente si y sólo si

$$\sum_{n\geq 1} \left(\frac{\beta}{\delta}\right)^n < \infty \Leftrightarrow \frac{\beta}{\delta} < 1.$$

Entonces $S = \frac{\delta}{\delta-\beta}$, luego por la ecuación 1.24 se tiene que

$$\begin{aligned}
\pi_0 &= \frac{\delta-\beta}{\delta} = 1 - \frac{\beta}{\delta}, \\
\pi_n &= \pi_0 \left(\frac{\beta}{\delta}\right)^n = \left(1 - \frac{\beta}{\delta}\right)\left(\frac{\beta}{\delta}\right)^n = (1-\rho)\,\rho^n.
\end{aligned}$$

Lo cual nos lleva a la siguiente proposición:

Proposición 1.10.1. *La cola $M/M/1$ con intensidad de tráfico ρ, es recurrente si y sólo si $\rho \leq 1$.*

Entonces por el corolario 1.9.1

Proposición 1.10.2. *La cola $M/M/1$ con intensidad de tráfico ρ es ergódica si y sólo si $\rho < 1$. En cuyo caso, la distribución de equilibrio π de la longitud de la cola es geométrica, $\pi_n = (1-\rho)\,\rho^n$, para $n = 1, 2, \ldots$.*

De la proposición anterior se desprenden varios hechos importantes.

a) $\mathbb{P}[X_t = 0] = \pi_0 = 1 - \rho$, es decir, la probabilidad de que el sistema se encuentre ocupado.

b) De las propiedades de la distribución Geométrica se desprende que

   i) $\mathbb{E}[X_t] = \frac{\rho}{1-\rho}$,

ii) $Var\left[X_t\right] = \frac{\rho}{(1-\rho)^2}$.

Si L es el número esperado de clientes en el sistema, incluyendo los que están siendo atendidos, entonces

$$L = \frac{\rho}{1-\rho}. \tag{1.31}$$

Si además W es el tiempo total del cliente en la cola: $W = W_q + W_s$, $\rho = \frac{\mathbb{E}[s]}{\mathbb{E}[\tau]} = \beta W_s$, puesto que $W_s = \mathbb{E}\left[s\right]$ y $\mathbb{E}\left[\tau\right] = \frac{1}{\delta}$. Por la fórmula de Little $L = \lambda W$

$$\begin{aligned} W &= \frac{L}{\beta} = \frac{\frac{\rho}{1-\rho}}{\beta} = \frac{\rho}{\delta}\frac{1}{1-\rho} = \frac{W_s}{1-\rho} \\ &= \frac{1}{\delta\left(1-\rho\right)} = \frac{1}{\delta-\beta}, \end{aligned}$$

luego entonces

$$\begin{aligned} W_q &= W - W_s = \frac{1}{\delta-\beta} - \frac{1}{\delta} = \frac{\beta}{\delta(\delta-\beta)} \\ &= \frac{\rho}{1-\rho}\frac{1}{\delta} = \mathbb{E}\left[s\right]\frac{\rho}{1-\rho}. \end{aligned}$$

Entonces

$$L_q = \beta W_q = \frac{\rho^2}{1-\rho}.$$

Finalmente, tenemos las siguientes proposiciones:

Proposición 1.10.3. *1.* $W\left(t\right) = 1 - e^{-\frac{t}{W}}$.

2. $W_q\left(t\right) = 1 - \rho \exp^{-\frac{t}{W}}$.
donde $W = \mathbb{E}(w)$.

Proposición 1.10.4. *La cola M/M/1 con intensidad de tráfico ρ es recurrente si y sólo si $\rho \leq 1$*

Proposición 1.10.5. *La cola M/M/1 con intensidad de tráfica ρ es ergodica si y sólo si $\rho < 1$. En este caso, la distribución de equilibrio π de la longitud de la cola es geométrica, $\pi_n = \left(1-\rho\right)\rho^n$, para $n = 0, 1, 2, \ldots$.*

1.10.2 Cola $M/M/\infty$

Este tipo de modelos se utilizan para estimar el número de líneas en uso en una gran red comunicación o para estimar valores en los sistemas $M/M/c$ o $M/M/c/c$, en el se puede pensar que siempre hay un servidor disponible para cada cliente que llega.

Se puede considerar como un proceso de nacimiento y muerte con parámetros $\beta_n = \beta$ y $\mu_n = n\mu$ para $n = 0, 1, 2, \ldots$. Este modelo corresponde al caso en que $\beta_n = \beta$ y $\delta_n = n\delta$, en este caso el parámetro de interés $\eta = \frac{\beta}{\delta}$, luego, la ecuación 1.21 queda de la forma:

$$\sum_{n=1}^{\infty} \frac{\delta_1 \cdots \delta_n}{\beta_1 \cdots \beta_n} = \sum_{n=1}^{\infty} n! \eta^{-n} = \infty$$

con $S = 1 + \sum_{n=1}^{\infty} \frac{\eta^n}{n!} = e$, entonces por la ecuación 1.24 se tiene que

$$\pi_0 = e^{\rho}, \tag{1.32}$$

$$\pi_n = e^{-\rho}\frac{\rho^n}{n!}. \tag{1.33}$$

Entonces, el número promedio de servidores ocupados es equivalente a considerar el número de clientes en el sistema, es decir,

$$L = \mathbb{E}\left[N\right] = \rho. \tag{1.34}$$
$$Var\left[N\right] = \rho. \tag{1.35}$$

Además se tiene que $W_q = 0$ y $L_q = 0$. El tiempo promedio en el sistema es el tiempo promedio de servicio, es decir, $W = \mathbb{E}\left[s\right] = \frac{1}{\delta}$.Resumiendo, tenemos la sisuguiente proposición:

Proposición 1.10.6. *La cola $M/M/\infty$ es ergódica para todos los valores de η. La distribución de equilibrio π es Poisson con media η,*

$$\pi_n = \frac{e^{-n}\eta^n}{n!}. \tag{1.36}$$

1.10.3 Cola $M/M/m$

Este sistema considera m servidores idénticos, con tiempos entre arribos y de servicio exponenciales con medias $\mathbb{E}\left[\tau\right] = \frac{1}{\beta}$ y $\mathbb{E}\left[s\right] = \frac{1}{\delta}$. definimos ahora la utilización por servidor como $u = \frac{\rho}{m}$ que también se puede interpretar como la fracción de tiempo promedio que cada servidor está ocupado.

La cola $M/M/m$ se puede considerar como un proceso de nacimiento y muerte con parámetros: $\beta_n = \beta$ para $n = 0, 1, 2, \ldots$ y

$$\delta_n = \begin{cases} n\delta & n = 0, 1, \ldots, m-1 \\ c\delta & n = m, \ldots \end{cases} \tag{1.37}$$

entonces la condición de recurrencia se va a cumplir sí y sólo si $\sum_{n\geq 1} \frac{\beta_0\cdots\beta_{n-1}}{\delta_1\cdots\delta_n} < \infty$, equivalentemente se debe de cumplir que

$$\begin{aligned} S &= 1 + \sum_{n\geq 1} \frac{\beta_0\cdots\beta_{n-1}}{\delta_1\cdots\delta_n} = \sum_{n=0}^{m-1} \frac{\beta_0\cdots\beta_{n-1}}{\delta_1\cdots\delta_n} + \sum_{n=0}^{\infty} \frac{\beta_0\cdots\beta_{n-1}}{\delta_1\cdots\delta_n} \\ &= \sum_{n=0}^{m-1} \frac{\beta^n}{n!\delta^n} + \sum_{n=0}^{\infty} \frac{\rho^m}{m!} u^n \end{aligned}$$

converja, lo cual ocurre si $u < 1$, en este caso

$$S = \sum_{n=0}^{m-1} \frac{\rho^n}{n!} + \frac{\rho^m}{m!}\left(1-u\right) \tag{1.38}$$

luego, para este caso se tiene que

$$\pi_0 = \frac{1}{S} \tag{1.39}$$

$$\pi_n = \begin{cases} \pi_0 \frac{\rho^n}{n!} & n = 0, 1, \ldots, m-1 \\ \pi_0 \frac{\rho^n}{m! m^{n-m}} & n = m, \ldots \end{cases} \tag{1.40}$$

Al igual que se hizo antes, determinaremos los valores de L_q, W_q, W y L:

$$\begin{aligned} L_q &= \mathbb{E}\left[N_q\right] = \sum_{n=0}^{\infty} \left(n-m\right)\pi_n = \sum_{n=0}^{\infty} n\pi_{n+m} \\ &= \sum_{n=0}^{\infty} n\pi_0 \frac{\rho^{n+m}}{m! m^{n+m}} = \pi_0 \frac{\rho^m}{m!} \sum_{n=0}^{\infty} n u^n = \pi_0 \frac{u\rho^m}{m!} \sum_{n=0}^{\infty} \frac{d}{du} u^n \\ &= \pi_0 \frac{u\rho^m}{m!} \frac{d}{du} \sum_{n=0}^{\infty} u^n = \pi_0 \frac{u\rho^m}{m!} \frac{d}{du}\left(\frac{1}{1-u}\right) = \pi_0 \frac{u\rho^m}{m!} \frac{1}{\left(1-u\right)^2}, \end{aligned}$$

es decir
$$L_q = \frac{u\pi_0\rho^m}{m!\left(1-u\right)^2}, \tag{1.41}$$

luego
$$W_q = \frac{L_q}{\beta}. \tag{1.42}$$

Además
$$W = W_q + \frac{1}{\delta} \tag{1.43}$$

Si definimos
$$C\left(m,\rho\right) = \frac{\pi_0\rho^m}{m!\left(1-u\right)} = \frac{\pi_m}{1-u}, \tag{1.44}$$

que es la probabilidad de que un cliente que llegue al sistema tenga que esperar en la cola. Entonces podemos reescribir las ecuaciones recién enunciadas:

$$\begin{aligned} L_q &= \frac{C\left(m,\rho\right)u}{1-u}, && (1.45)\\ W_q &= \frac{C\left(m,\rho\right)\mathbb{E}\left[s\right]}{m\left(1-u\right)} && (1.46)\\ & && (1.47)\end{aligned}$$

Por tanto tenemos las siguientes proposiciones:

Proposición 1.10.7. *La cola $M/M/m$ con intensidad de tráfico ρ es ergódica si y sólo si $\rho < 1$. En este caso la distribución ergódica π está dada por*

$$\pi_n = \begin{cases} \frac{1}{S}\frac{\eta^n}{n!} & 0 \leq n \leq m\\ \frac{1}{S}\frac{\eta^m}{m!}\rho^{n-m} & m \leq n < \infty \end{cases} \tag{1.48}$$

Proposición 1.10.8. *Para $t \geq 0$*

a)

$$W_q\left(t\right) = 1 - C\left(m,\rho\right)e^{-c\delta t(1-u)}. \tag{1.49}$$

b)

$$W\left(t\right) = \begin{cases} 1 + e^{-\delta t}\frac{\rho-m+W_q(0)}{m-1-\rho} + e^{-m\delta t(1-u)}\frac{C(m,\rho)}{m-1-\rho} & \rho \neq m-1\\ 1 - \left(1 + C\left(m,\rho\right)\delta t\right)e^{-\delta t} & \rho = m-1 \end{cases} \tag{1.50}$$

Resumiendo, para este caso $\beta_n = \beta$ y $\delta_n = m\left(n\right)\delta$, donde $m\left(n\right)$ es el número de servidores ocupados en el estado n, es decir, $m\left(n\right) = m$, para $n \geq m$ y $m\left(n\right) = m$ para $1 \leq n \leq m$. La intensidad de tráfico es $\rho = \frac{\beta}{m\delta}$ y $\frac{\beta_n}{\delta_n} = \rho$ para $n \geq m$. Así, al igual que en el caso $m = 1$, la ecuación 1.21 y la recurrencia se cumplen si y sólo si $\sum_{n=1}^{\infty}\rho^{-n} = \infty$, es decir, cuando $\rho \leq 1$.

1.10.4 Cola $M/M/m/m$

Consideremos un sistema con m servidores idénticos, pero ahora cada uno es de capacidad finita m. Si todos los servidores se encuentran ocupados, el siguiente usuario en llegar se pierde pues no se le deja esperar a que reciba servicio. Este tipo de sistemas pueden verse como un proceso de nacimiento y muerte con

$$\beta_n = \begin{cases} \beta & n = 0,1,2,\ldots,m-1\\ 0 & n \geq m \end{cases} \tag{1.51}$$

$$\delta_n = \begin{cases} n\delta & n = 0, 1, 2, \dots, m-1 \\ 0 & n \geq m \end{cases} \tag{1.52}$$

El proceso tiene epacio de estados finitos, $S = \{0, 1, \dots, m\}$, entonces de las ecuaciones que determinan la distribución estacionaria se tiene que

$$\pi_n = \begin{cases} \pi_0 \frac{\rho^n}{n!} & n = 0, 1, 2, \dots, m \\ 0 & n \geq m \end{cases} \tag{1.53}$$

y además

$$\pi_0 = \left(\sum_{n=0}^{m} \frac{\rho^n}{n!} \right)^{-1}. \tag{1.54}$$

A la ecuación 1.53 se le llama *distribución truncada*. Si definimos $\pi_m = B(m, \rho) = \pi_0 \frac{\rho^m}{m!}$, π_m representa la probabilidad de que todos los servidores se encuentren ocupados, y también se le conoce como *fórmula de pérdida de Erlang*. Necesariamente en este caso el tiempo de espera en la cola W_q y el número promedio de clientes en la cola L_q deben de ser cero puesto que no se permite esperar para recibir servicio, más aún, el tiempo de espera en el sistema y el tiempo de serivcio tienen la misma distribución, es decir,

$$W(t) = \mathbb{P}\{w \leq t\} = 1 - e^{-\mu t},$$

en particular

$$W = \mathbb{E}[w] = \mathbb{E}[s] = \frac{1}{\delta}.$$

Por otra parte, el número esperado de clientes en el sistema es

$$\begin{aligned} L &= \mathbb{E}[N] = \sum_{n=0}^{m} n\pi_n = \pi_0 \rho \sum_{n=0}^{m} \frac{\rho^{n-1}}{(n-1)!} \\ &= \pi_0 \rho \sum_{n=0}^{m-1} \frac{\rho^n}{n!} \end{aligned}$$

entonces, se tiene que

$$L = \rho(1 - B(m, \rho)) = \mathbb{E}[s](1 - B(m, \rho)). \tag{1.55}$$

Además

$$\delta_q = \delta(1 - B(m, \rho)) \tag{1.56}$$

representa la tasa promedio efectiva de arribos al sistema.

1.10.5 Cola M/G/1

Consideremos un sistema de espera con un servidor, en el que los tiempos entre arribos son exponenciales, y los tiempos de servicio tienen una distribución general G. Sea $N(t)_{t \geq 0}$ el número de clientes en el sistema al tiempo t, y sean $t_1 < t_2 < \dots$ los tiempos sucesivos en los que los clientes completan su servicio y salen del sistema.

La sucesión $\{X_n\}$ definida por $X_n = N(t_n)$ es una cadena de Markov, en específico es la Cadena encajada del proceso de llegadas de usuarios. Sea U_n el número de clientes que llegan al sistema durante el tiempo de servicio del n-ésimo cliente, entonces se tiene que

$$X_{n+1} = \begin{cases} X_n - 1 + U_{n+1} & \text{si } X_n \geq 1, \\ U_{n+1} & \text{si } X_n = 0 \end{cases}$$

Dado que los procesos de arribos de los usuarios es Poisson con parámetro λ, la probabilidad condicional de que lleguen j clientes al sistema dado que el tiempo de servicio es $s = t$, resulta:

$$\mathbb{P}\{U = j | s = t\} = e^{-\lambda t} \frac{(\lambda t)^j}{j!}, \; j = 0, 1, \dots$$

por el teorema de la probabilidad total se tiene que

$$a_j = \mathbb{P}\left\{U = j\right\} = \int_0^\infty \mathbb{P}\left\{U = j|s = t\right\} dG\left(t\right) = \int_0^\infty e^{-\lambda t}\frac{\left(\lambda t\right)^j}{j!} dG\left(t\right) \tag{1.57}$$

donde G es la distribución de los tiempos de servicio. Las probabilidades de transición de la cadena están dadas por

$$p_{0j} = \mathbb{P}\left\{U_{n+1} = j\right\} = a_j, \text{ para } j = 0, 1, \ldots \tag{1.58}$$

y para $i \geq 1$

$$p_{ij} = \begin{cases} \mathbb{P}\left\{U_{n+1} = j - i + 1\right\} = a_{j-i+1} & \text{, para } j \geq i - 1 \\ 0 & j < i - 1 \end{cases} \tag{1.59}$$

Entonces la matriz de transición es:

$$P = \begin{bmatrix} a_0 & a_1 & a_2 & a_3 & \cdots \\ a_0 & a_1 & a_2 & a_3 & \cdots \\ 0 & a_0 & a_1 & a_2 & \cdots \\ 0 & 0 & a_0 & a_1 & \cdots \\ \vdots & \vdots & \ldots & \ddots & \vdots \end{bmatrix}.$$

Sea $\rho = \sum_{n=0} n a_n$, entonces se tiene el siguiente teorema:

Teorema 1.10.1. *La cadena encajada $\{X_n\}$ es*

a) Recurrente positiva si $\rho < 1$,

b) Transitoria si $\rho > 1$,

c) Recurrente nula si $\rho = 1$.

Recordemos que si la cadena de Markov $\{X_n\}$ tiene una distribución estacionaria entonces existe una distribución de probabilidad $\pi = (\pi_0, \pi_1, \ldots,)$, con $\pi_i \geq 0$ y $\sum_{i\geq 1}\pi_i = 1$ tal que satisface la ecuación $\pi = \pi P$, equivalentemente

$$\pi_j = \sum_{i=0}^{\infty} \pi_k p_{ij}, \text{ para } j = 0, 1, 2, \ldots \tag{1.60}$$

que se puede ver como

$$\pi_j = \pi_0 a_j + \sum_{i=1}^{j+1} \pi_i a_{j-i+1}, \text{ para } j = 0, 1, \ldots \tag{1.61}$$

si definimos

$$\pi\left(z\right) = \sum_{j=0}^{\infty} \pi_j z^j \tag{1.62}$$

y

$$A\left(z\right) = \sum_{j=0}^{\infty} a_j z^j \tag{1.63}$$

con $|z_j| \leq 1$. Si la ecuación 1.61 la multiplicamos por z^j y sumando sobre j, se tiene que

$$\begin{aligned}
\sum_{j=0}^{\infty} \pi_j z^j &= \sum_{j=0}^{\infty} \pi_0 a_j z^j + \sum_{j=0}^{\infty}\sum_{i=1}^{j+1} \pi_i a_{j-i+1} z^j \\
&= \pi_0 \sum_{j=0}^{\infty} a_j z^j + \sum_{j=0}^{\infty} a_j z^j \sum_{i=1}^{\infty} \pi_i a_{i-1} \\
&= \pi_0 A\left(z\right) + A\left(z\right)\left(\frac{\pi\left(z\right) - \pi_0}{z}\right)
\end{aligned}$$

es decir,

$$\pi(z) = \pi_0 A(z) + A(z)\left(\frac{\pi(z) - \pi_0}{z}\right) \Leftrightarrow \pi(z) = \frac{\pi_0 A(z)(z-1)}{z - A(z)} \tag{1.64}$$

Si $z \to 1$, entonces $A(z) \to A(1) = 1$, y además $A^{'}(z) \to A^{'}(1) = \rho$. Si aplicamos la Regla de L'Hospital se tiene que

$$\sum_{j=0}^{\infty} \pi_j = lim_{z\to 1^-} \pi(z) = \pi_0 lim_{z\to 1^-} \frac{z-1}{z - A(z)} = \frac{\pi_0}{1-\rho}$$

Retomando,

$$a_j = \mathbb{P}\{U = j\} = \int_0^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} dG(t), \text{ para } n = 0, 1, 2, \ldots$$

entonces

$$\begin{aligned} \rho &= \sum_{n=0}^{\infty} n a_n = \sum_{n=0}^{\infty} n \int_0^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} dG(t) \\ &= \int_0^{\infty} \sum_{n=0}^{\infty} n e^{-\lambda t} \frac{(\lambda t)^n}{n!} dG(t) = \int_0^{\infty} \lambda t dG(t) = \lambda \mathbb{E}[s] \end{aligned}$$

Además, se tiene que $\rho = \beta \mathbb{E}[s] = \frac{\beta}{\delta}$ y la distribución estacionaria está dada por

$$\pi_j = \pi_0 a_j + \sum_{i=1}^{j+1} \pi_i a_{j-i+1}, \text{ para } j = 0, 1, \ldots \tag{1.65}$$

$$\pi_0 = 1 - \rho. \tag{1.66}$$

Por otra parte se tiene que

$$L = \pi^{'}(1) = \rho + \frac{A^{''}(1)}{2(1-\rho)} \tag{1.67}$$

pero $A^{''}(1) = \sum_{n=1} n(n-1) a_n = \mathbb{E}\left[U^2\right] - \mathbb{E}[U]$, $\mathbb{E}[U] = \rho$ y $\mathbb{E}\left[U^2\right] = \lambda^2 \mathbb{E}\left[s^2\right] + \rho$. Por lo tanto $L = \rho + \frac{\beta^2 \mathbb{E}\left[s^2\right]}{2(1-\rho)}$.

De las fórmulas de Little, se tiene que $W = E(w) = \frac{L}{\beta}$, también el tiempo de espera en la cola

$$W_q = \mathbb{E}(q) = \mathbb{E}(w) - \mathbb{E}(s) = \frac{L}{\beta} - \frac{1}{\delta}, \tag{1.68}$$

además el número promedio de clientes en la cola es

$$L_q = \mathbb{E}(N_q) = \beta W_q = L - \rho \tag{1.69}$$

1.10.6 Cola con Infinidad de Servidores

Este caso corresponde a $\beta_n = \beta$ y $\delta_n = n\delta$. El parámetro de interés es $\eta = \frac{\beta}{\delta}$, de donde se obtiene:

$$\sum_{n\geq 0} \frac{\delta_1 \cdots \delta_n}{\beta_1 \cdots \beta_n} = \sum_{n=1}^{\infty} n! \eta^n = \infty,$$

$$S = 1 + \sum_{n=1}^{\infty} \frac{\eta^n}{n!} = e^n.$$

Proposición 1.10.9. *La cola $M/M/\infty$ es ergodica para todos los valores de η. La distribución de equilibrio π es Poisson con media η, $\pi_n = \frac{e^{-n}\eta}{n!}$*

1.11 Redes de Colas:Sistemas Abiertos

Considerese un sistema con dos servidores, en los cuales los usuarios llegan de acuerdo a un proceso poisson con intensidad λ_1 al primer servidor, después de ser atendido se pasa a la siguiente cola en el segundo servidor. Cada servidor atiende a un usuario a la vez con tiempo exponencial con razón μ_i, para $i = 1, 2$. A este tipo de sistemas se les conoce como sistemas secuenciales.

Defínase el par (n, m) como el número de usuarios en el servidor 1 y 2 respectivamente. Las ecuaciones de balance son

$$\begin{aligned} \lambda P_{0,0} &= \mu_2 P_{0,1} && (1.70)\\ (\lambda+\mu_1) P_{n,0} &= \mu_2 P_{n,1} + \lambda P_{n-1,0} && (1.71)\\ (\lambda+\mu_2) P_{0,m} &= \mu_2 P_{0,m+1} + \mu_1 P_{1,m-1} && (1.72)\\ (\lambda+\mu_1+\mu_2) P_{n,m} &= \mu_2 P_{n,m+1} + \mu_1 P_{n+1,m-1} + \lambda P_{n-1,m} && (1.73) \end{aligned}$$

Cada servidor puede ser visto como un modelo de tipo $M/M/1$, de igual manera el proceso de salida de una cola $M/M/1$ con razón λ, nos permite asumir que el servidor 2 también es una cola $M/M/1$. Además la probabilidad de que haya n usuarios en el servidor 1 es

$$\begin{aligned} P\{n \text{ en el servidor } 1\} &= \left(\frac{\lambda}{\mu_1}\right)^n \left(1-\frac{\lambda}{\mu_1}\right) = \rho_1^n (1-\rho_1)\\ P\{m \text{ en el servidor } 2\} &= \left(\frac{\lambda}{\mu_2}\right)^n \left(1-\frac{\lambda}{\mu_2}\right) = \rho_2^m (1-\rho_2) \end{aligned}$$

Si el número de usuarios en los servidores 1 y 2 son variables aleatorias independientes, se sigue que:

$$P_{n,m} = \rho_1^n (1-\rho_1) \rho_2^m (1-\rho_2) \tag{1.74}$$

Verifiquemos que $P_{n,m}$ satisface las ecuaciones de balance (1.70) Antes de eso, enunciemos unas igualdades que nos serán de utilidad:

$$\begin{aligned} \mu_i \rho_i &= \lambda \text{ para } i = 1, 2.\\ \lambda P_{0,0} &= \lambda (1-\rho_1)(1-\rho_2)\\ \text{y } \mu_2 P_{0,1} &= \mu_2 (1-\rho_1) \rho_2 (1-\rho_2) \Rightarrow\\ \lambda P_{0,0} &= \mu_2 P_{0,1}\\ (\lambda+\mu_2) P_{0,m} &= (\lambda+\mu_2)(1-\rho_1) \rho_2^m (1-\rho_2)\\ \mu_2 P_{0,m+1} &= \lambda (1-\rho_1) \rho_2^m (1-\rho_2)\\ &= \mu_2 (1-\rho_1) \rho_2^m (1-\rho_2)\\ \mu_1 P_{1,m-1} &= \frac{\lambda}{\rho_2} (1-\rho_1) \rho_2^m (1-\rho_2) \Rightarrow\\ (\lambda+\mu_2) P_{0,m} &= \mu_2 P_{0,m+1} + \mu_1 P_{1,m-1}\\ (\lambda+\mu_1+\mu_2) P_{n,m} &= (\lambda+\mu_1+\mu_2) \rho^n (1-\rho_1) \rho_2^m (1-\rho_2)\\ \mu_2 P_{n,m+1} &= \mu_2 \rho_2 \rho_1^n (1-\rho_1) \rho_2^m (1-\rho_2)\\ \mu_1 P_{n-1,m-1} &= \mu_1 \frac{\rho_1}{\rho_2} \rho_1^n (1-\rho_1) \rho_2^m (1-\rho_2)\\ \lambda P_{n-1,m} &= \frac{\lambda}{\rho_1} \rho_1^n (1-\rho_1) \rho_2^m (1-\rho_2)\\ \Rightarrow (\lambda+\mu_1+\mu_2) P_{n,m} &= \mu_2 P_{n,m+1} + \mu_1 P_{n-1,m-1} + \lambda P_{n-1,m} \end{aligned}$$

entonces efectivamente la ecuación (1.74) satisface las ecuaciones de balance (1.70). El número promedio de usuarios en el sistema, está dado por

$$
\begin{aligned}
L &= \sum_{n,m}(n+m)\,P_{n,m} = \sum_{n,m} nP_{n,m} + \sum_{n,m} mP_{n,m}\\
&= \sum_{n}\sum_{m} nP_{n,m} + \sum_{m}\sum_{n} mP_{n,m} = \sum_{n} n\sum_{m} P_{n,m} + \sum_{m} m\sum_{n} P_{n,m}\\
&= \sum_{n} n\sum_{m} \rho_1^n\,(1-\rho_1)\,\rho_2^m\,(1-\rho_2) + \sum_{m} m\sum_{n} \rho_1^n\,(1-\rho_1)\,\rho_2^m\,(1-\rho_2)\\
&= \sum_{n} n\rho_1^n\,(1-\rho_1)\sum_{m}\rho_2^m\,(1-\rho_2) + \sum_{m} m\rho_2^m\,(1-\rho_2)\sum_{n}\rho_1^n\,(1-\rho_1)\\
&= \sum_{n} n\rho_1^n\,(1-\rho_1) + \sum_{m} m\rho_2^m\,(1-\rho_2)\\
&= \frac{\lambda}{\mu_1-\lambda} + \frac{\lambda}{\mu_2-\lambda}
\end{aligned}
$$

CAPÍTULO 2

Sistemas de Visita

2.1 Sistemas de Visitas

Los *Sistemas de Visitas* fueron introducidos a principios de los años 50, ver [85, 87, 113, 119, 131, 121], con un problema relacionado con las personas encargadas de la revisión y reparación de máquinas; más adelante fueron utilizados para estudiar problemas de control de señales de tráfico. A partir de ese momento el campo de aplicación ha crecido considerablemente, por ejemplo en: comunicación en redes de computadoras, robótica, tráfico y transporte, manufactura, producción, distribución de correo, sistema de saludpública, etc.

Un modelo de colas es un modelo matemático que describe la situación en la que uno o varios usuarios solicitan de un servicio a una instancia, computadora o persona. Aquellos usuarios que no son atendidos inmediatamente toman un lugar en una cola en espera de servicio. Un sistema de visitas consiste en modelos de colas conformadas por varias colas y un solo servidor que las visita en algún orden para atender a los usuarios que se encuentran esperando por servicio.

Uno de los principales objetivos de este tipo de sistemas es tratar de mejorar el desempeño del sistema de visitas. Una de medida de desempeño importante es el tiempo de respuesta del sistema, así como los tiempos promedios de espera en una fila y el tiempo promedio total que tarda en ser realizada una operación completa a lo largo de todo el sistema.

Algunas medidas de desempeño para los usuarios son los valores promedio de espera para ser atendidos, de servicio, de permanencia total en el sistema; mientras que para el servidor son los valores promedio de permanencia en una cola atendiendo, de traslado entre las colas, de duración del ciclo entre dos visitas consecutivas a la misma cola, entre otras medidas de desempeño estudiadas en la literatura.

En la mayoría de los modelos de colas cíclicas, la capacidad de almacenamiento es infinita, es decir la cola puede acomodar a una cantidad infinita de usuarios a la vez.

Los sistemas de visitas pueden dividirse en dos clases:

i) hay varios servidores y los usuarios que llegan al sistema eligen un servidor de entre los que están presentes.

ii) hay uno o varios servidores que son comunes a todas las colas, estos visitan a cada una de las colas y atienden a los usuarios que están presentes al momento de la visita del servidor.

La manera en que los usuarios llegan a las colas. Los usuarios llegan a las colas de manera tal que los tiempos entre arribos son independientes e idénticamente distribuidos. En la mayoría de los modelos de visitas cíclicas, la capacidad de almacenamiento es infinita, es decir la cola puede acomodar a una cantidad infinita de usuarios a la vez.

Los tiempos de servicio en una cola son usualmente considerados como muestra de una distribución de probabilidad que caracteriza a la cola, además se acostumbra considerarlos mutuamente independientes e independientes del estado actual del sistema.

La ruta de atención del servidor, es el orden en el cual el servidor visita las colas determinado por un mecanismo que puede depender del estado actual del sistema (dinámico) o puede ser independiente del estado del sistema (estático).

El mecanismo más utilizado es el cíclico. Para modelar sistemas en los cuales ciertas colas son visitadas con mayor frecuencia que otras, las colas cíclicas se han extendido a colas periódicas, en las cuales el servidor visita la cola conforme a una orden de servicio de longitud finita.

El *orden de visita* se entiende como la regla utilizada por el servidor para elegir la próxima cola. Este servicio puede ser dinámico o estático:

i) Para el caso *estático* la regla permanece invariante a lo largo del curso de la operación del sistema.

ii) Para el caso *dinámico* la cola que se elige para servicio en el momento depende de un conocimiento total o parcial del estado del sistema.

Dentro de los ordenes de tipo estático hay varios, los más comunes son:

i) *cíclico*: Si denotamos por $\{Q_i\}_{i=1}^{N}$ al conjunto de colas a las cuales el servidor visita en el orden

$$Q_1, Q_2, \ldots, Q_N, Q_1, Q_2, \ldots, Q_N.$$

ii) *periódico*: el servidor visita las colas en el orden:

$$Q_{T(1)}, Q_{T(2)}, \ldots, Q_{T(M)}, Q_{T(1)}, \ldots, Q_{T(M)}$$

caracterizada por una tabla de visitas

$$\left(T\left(1\right), T\left(2\right), \ldots, T\left(M\right)\right),$$

con $M \geq N$, $T\left(i\right) \in \{1, 2, \ldots, N\}$ e $i = \overline{1, M}$. Hay un caso especial, *colas tipo elevador* donde las colas son atendidas en el orden

$$Q_1, Q_2, \ldots, Q_N, Q_1, Q_2, \ldots, Q_{N-1}, Q_N, Q_{N-1}, \ldots, Q_1$$

.

iii) *aleatorio*: la cola Q_i es elegida para serbatendida con probabilidad p_i, $i = \overline{1, N}$, $\sum_{i=1}^{N} p_i = 1$. Una posible variación es que después de atender Q_i el servidor se desplaza a Q_j con probabilidad p_{ij}, con $i, j = \overline{1, N}$, $\sum_{j=1}^{N} p_{ij} = 1$, para $i = \overline{1, N}$.

El servidor usualmente incurrirá en tiempos de traslado para ir de una cola a otra. Un sistema de visitas puede expresarse en un par de parámetros: el número de colas, que usualmente se denotará por N, y el tráfico característico de las colas, que consiste de los procesos de arribo y los procesos de servicio, la figura (**??**) caracteriza a estos sistemas.

La disciplina de servicio especifica el número de usuarios que son atendidos durante la visita del servidor a la cola; estas pueden ser clasificadas en límite de usuarios atendidos y en usuarios atendidos en un tiempo límite, poniendo restricciones en la cantidad de tiempo utilizado por el servidor en una visita a la cola. Alternativamente pueden ser clasificadas en políticas exhaustivas y políticas cerradas, dependiendo en si los usuarios que llegaron a la cola mientras el servidor estaba dando servicio son candidatos para ser atendidos por el servidor que se encuentra en la cola dando servicio. En la política exhaustiva estos usuarios son candidatos para ser atendidos mientras que en la cerrada no lo son. De estas dos políticas se han creado híbridos los cuales pueden revisarse en [87].

La disciplina de la cola especifica el orden en el cual los usuarios presentes en la cola son atendidos. La más común es la *First-In-First-Served*.

Las políticas más comunes son las de tipo exhaustivo que consiste en que el servidor continuará trabajando hasta que la cola quede vacía; y la política cerrada, bajo la cual serán atendidos exactamente aquellos que estaban presentes al momento en que llegó el servidor a la cola.

Las políticas de servicio deben de satisfacer las siguientes propiedades:

i) No dependen de los procesos de servicio anteriores.

ii) La selección de los usuarios para ser atendidos es independiente del tiempo de servicio requerido y de los posibles arribos futuros.

iii) las políticas de servicio que son aplicadas, es decir, el número de usuarios en la cola que serán atendidos durante la visita del servidor a la misma; éstas pueden ser clasificadas por la cantidad de usuarios atendidos y por el número de usuarios atendidos en un intervalo de tiempo determinado. Las principales políticas de servicio para las cuales se han desarrollado aplicaciones son: la exhaustiva, la cerrada y la k-límite, ver [113, 131, 121]. De estas políticas se han creado híbridos los cuales pueden revisarse en Boon and Van der Mei [87].

iv) Una política de servicio es asignada a cada etapa independiente de la cola que se está atendiendo, no necesariamente es la misma para todas las etapas.

v) El servidor da servicio de manera constante.

vi) La política de servicio se asume monótona (ver [104]).

Las principales políticas deterministas de servicio son:

i) *Cerrada* donde solamente los usuarios presentes al comienzo de la etapa son considerados para ser atendidos.

ii) *Exhaustiva* en la que tanto los usuarios presentes al comienzo de la etapa como los que arriban mientras se está dando servicio son considerados para ser atendidos.

iii) k_i-limited: el número de usuarios por atender en la cola i estácotado por k_i.

iv) *tiempo limitado* la cola es atendida solo por un periodo de tiempo fijo.

- Una etapa es el periodo de tiempo durante el cual el servidor atiende de manera continua en una sola cola.
- Un ciclo es el periodo necesario para terminar l etapas.

Boxma y Groenendijk [86] enuncian la Ley de Pseudo-Conservación para la política exhaustiva como

$$\sum_{i=1}^{N}\rho_i\mathbb{E}W_i=\rho\frac{\sum_{i=1}^{N}\lambda_i\mathbb{E}\left[\delta_i^{(2)}\left(1\right)\right]}{2\left(1-\rho\right)}+\rho\frac{\delta^{(2)}}{2\delta}+\frac{\delta}{2\left(1-\rho\right)}\left[\rho^2-\sum_{i=1}^{N}\rho_i^2\right], \tag{2.1}$$

donde $\delta=\sum_{i=1}^{N}\delta_i\left(1\right)$ y $\delta_i^{(2)}$ denota el segundo momento de los tiempos de traslado entre colas del servidor, $\delta^{(2)}$ es el segundo momento de los tiempos de traslado entre las colas de todo el sistema, finalmente sea $\rho=\sum_{i=1}^{N}\rho_i$. Por otro lado, se tiene que

$$\mathbb{E}W_i=\frac{\mathbb{E}I_i^2}{2\mathbb{E}I_i}+\frac{\lambda_i\mathbb{E}\left[\eta_i^{(2)}\left(1\right)\right]}{2\left(1-\rho_i\right)}, \tag{2.2}$$

con I_i definido como el periódo de intervisita, es decir el tiempo entre una salida y el próximo arribo del servidor a la cola Q_i, dado por $I_i=C_i-V_i$, donde C_i es la longitud del ciclo, definido como el tiempo entre dos instantes de visita consecutivos a la cola Q_i y V_i es el periodo de visita, definido como el tiempo que el servidor utiliza en atender a los usuarios de la cola Q_i.

$$\mathbb{E}I_i=\frac{\left(1-\rho_i\right)}{1-\rho}\sum_{i=1}^{N}\mathbb{E}\left[\delta_i\left(1\right)\right], \tag{2.3}$$

con

$$\mathbb{E}I_i^2=\mathbb{E}\left[\delta_{i-1}^{(2)}\left(1\right)\right]-\left(\mathbb{E}\left[\delta_{i-1}\left(1\right)\right]\right)^2+\frac{1-\rho_i}{\rho_i}\sum_{j=1,j\neq i}^{N}r_{ij}+\left(\mathbb{E}I_i\right)^2, \tag{2.4}$$

donde el conjunto de valores $\{r_{ij}:i,j=1,2,\ldots,N\}$ representan la covarianza del tiempo para las colas i y j; para sistemas con servicio exhaustivo, el tiempo de estación para la cola i se define como el intervalo de tiempo entre instantes sucesivos cuando el servidor abandona la cola $i-1$ y la cola i. Hideaki

Takagi [131] proporciona expresiones cerradas para calcular r_{ij}, éstas implican resolver un sistema de N^2 Ecuaciones lineales;

$$r_{ij} = \frac{\rho_i}{1-\rho_i}\left(\sum_{m=i+1}^{N} r_{jm} + \sum_{m=1}^{j-1} r_{jm} + \sum_{m=j}^{i-1} r_{jm}\right),\ j<i, \tag{2.5}$$

$$r_{ij} = \frac{\rho_i}{1-\rho_i}\left(\sum_{m=i+1}^{j-1} r_{jm} + \sum_{m=j}^{N} r_{jm} + \sum_{m=1}^{i-1} r_{jm}\right),\ j>i, \tag{2.6}$$

$$r_{ij} = \frac{\mathbb{E}\left[\delta_{i-1}^{(2)}(1)\right] - \left(\mathbb{E}\left[\delta_{i-1}(1)\right]\right)^2}{(1-\rho_i)^2} + \frac{\lambda_i \mathbb{E}\left[\eta_i(1)^{(2)}\right]}{(1-\rho_i)^3} + \frac{\rho_i}{1-\rho_i}\sum_{j=i,j=1}^{N} r_{ij}. \tag{2.7}$$

Para el caso de la Política Cerrada la Ley de Pseudo-Conservación se expresa en los siguientes términos.

$$\sum_{i=1}^{N}\rho_i \mathbb{E}W_i = \rho\frac{\sum_{i=1}^{N}\lambda_i\mathbb{E}\left[\delta_i(1)^{(2)}\right]}{2(1-\rho)} + \rho\frac{\delta^{(2)}}{2\delta} + \frac{\delta}{2(1-\rho)}\left[\rho^2 + \sum_{i=1}^{N}\rho_i^2\right], \tag{2.8}$$

el tiempo de espera promedio para los usuarios en la cola Q_1 se puede determinar por medio de

$$\mathbb{E}W_i = \frac{(1+\rho_i)\,\mathbb{E}C_i^2}{2\mathbb{E}C_i}, \tag{2.9}$$

donde C_i denota la longitud del ciclo para la cola Q_i, definida como el tiempo entre dos instantes consecutivos de visita en Q_i, cuyo segundo momento está dado por

$$\mathbb{E}C_i^2 = \frac{1}{\rho_i}\sum_{j=1,j\neq i}^{N} r_{ij} + \sum_{j=1}^{N} r_{ij} + (\mathbb{E}C)^2, \tag{2.10}$$

con

$$\mathbb{E}C = \frac{\delta}{1-\rho},$$

donde r_{ij} representa la covarianza del tiempo de estación para las colas i y j, pero el tiempo de estación para la cola i para la política cerrada se define como el intervalo de tiempo entre instantes sucesivos cuando el servidor visita la cola i y la cola $i+1$. El conjunto $\{r_{ij} : i,j = 1,2,\ldots,N\}$ se calcula resolviendo un sistema de N^2 ecuaciones lineales

$$r_{ij} = \rho_i\left(\sum_{m=i}^{N} r_{jm} + \sum_{m=1}^{j-1} r_{jm} + \sum_{m=j}^{i-1} r_{mj}\right),\ j<i, \tag{2.11}$$

$$r_{ij} = \rho_i\left(\sum_{m=i}^{j-1} r_{jm} + \sum_{m=j}^{N} r_{jm} + \sum_{m=1}^{i-1} r_{mj}\right),\ j>i, \tag{2.12}$$

$$r_{ij} = r_{i-1}^{(2)} - \left(r_{i-1}^{(1)}\right)^2 + \lambda_i b_i^{(2)}\mathbb{E}C_i + \rho_i\sum_{j=1,j\neq i}^{N} r_{ij} + \rho_i^2\sum_{i=j,j=1}^{N} r_{ij}. \tag{2.13}$$

Finalmente, Takagi [131] proponen una aproximación para los tiempos de espera de los usuarios en cada una de las colas:

$$\begin{aligned}\sum_{i=1}^{N}\frac{\rho_i}{\rho}\left(1-\frac{\lambda_i\delta}{1-\rho}\right)\mathbb{E}\left[W_i\right] &= \sum_{i=1}^{N}\frac{\lambda_i\mathbb{E}\left[\eta_i(1)^{(2)}\right]}{2(1-\rho)}\\ +\frac{\sum_{i=1}^{N}\mathbb{E}\left[\delta_i^2\right] - \left(\mathbb{E}\left[\delta_i(1)\right]\right)^2}{2\delta} &+ \frac{\delta\left(\rho-\sum_{i=1}^{N}\rho_i^2\right)}{2\rho(1-\rho)} + \frac{\delta\sum_{i=1}^{N}\rho_i^2}{\rho(1-\rho)},\end{aligned}$$

entonces

$$\begin{aligned}\mathbb{E}W_i &\cong \frac{1-\rho+\rho_i}{1-\rho-\lambda_i\delta}\times\frac{1-\rho}{\rho\left(1-\rho\right)+\sum_{i=1}^{N}\rho_i^2}\\ &\times \left[\frac{\rho}{2\left(1-\rho\right)}\sum_{i=1}^{N}\lambda_i\mathbb{E}\left[\eta_i\left(1\right)^{(2)}\right]+\frac{\rho\Delta^2}{2\delta}+\frac{\delta}{2\left(1-\rho\right)}\sum_{i=1}^{N}\rho_i\left(1+\rho_i\right).\right]\end{aligned}$$

donde $\Delta^2=\sum_{i=1}^{N}\delta_i^2$.

El modelo está compuesto por c colas de capacidad infinita, etiquetadas de 1 a c las cuales son atendidas por s servidores. Los servidores atienden de acuerdo a una cadena de Markov independiente $\left(X_n^i\right)_n$ con $1\leq i\leq s$ y $n\in\{1,2,\ldots,c\}$ con la misma matriz de transición $r_{k,l}$ y única medida invariante (p_k). Cada servidor permanece atendiendo en la cola un periodo llamado de visita y determinada por la política de servicio asignada a la cola.

Los usuarios llegan a la cola k con una tasa λ_k y son atendidos a una razón μ_k. Las sucesiones de tiempos de interarribo $\left(\tau_k\left(n\right)\right)_n$, la de tiempos de servicio $\left(\sigma_k^i\left(n\right)\right)_n$ y la de tiempos de cambio $\left(\sigma_{k,l}^{0,i}\left(n\right)\right)_n$ requeridas en la cola k para el servidor i son sucesiones independientes e idénticamente distribuidas con distribución general independiente de i, con media $\sigma_k=\frac{1}{\mu_k}$, respectivamente $\sigma_{k,l}^0=\frac{1}{\mu_{k,l}^0}$, e independiente de las cadenas de Markov $\left(X_n^i\right)_n$. Además se supone que los tiempos de interarribo se asume son acotados, para cada $\rho_k=\lambda_k\sigma_k<s$ para asegurar la estabilidad de la cola k cuando opera como una cola $M/GM/1$.

Una política de servicio determina el número de usuarios que serán atendidos sin interrupción en periodo de servicio por los servidores que atienden a la cola. Para un solo servidor esta se define a través de una función f donde $f\left(x,a\right)$ es el número de usuarios que son atendidos sin interrupción cuando el servidor llega a la cola y encuentra x usuarios esperando dado el tiempo transcurrido de interarribo a. Sea $v\left(x,a\right)$ la du ración del periodo de servicio para una sola condición inicial $\left(x,a\right)$.

Las políticas de servicio consideradas satisfacen las siguientes propiedades:

i) Hay conservación del trabajo, es decir

$$v\left(x,a\right)=\sum_{l=1}^{f(x,a)}\sigma\left(l\right)$$

con $f\left(0,a\right)=v\left(0,a\right)=0$, donde $\left(\sigma\left(l\right)\right)_l$ es una sucesión independiente e idénticamente distribuida de los tiempos de servicio solicitados.

ii) La selección de usuarios para se atendidos es independiente de sus correspondientes tiempos de servicio y del pasado hasta el inicio del periodo de servicio. Así las distribución $\left(f,v\right)$ no depende del orden en el cuál son atendidos los usuarios.

iii) La política de servicio es monótona en el sentido de que para cada $a\geq 0$ los números $f\left(x,a\right)$ son monótonos en distribución en x y su límite en distribución cuando $x\rightarrow\infty$ es una variable aleatoria F^{*0} que no depende de a.

iv) El número de usuarios atendidos por cada servidor es acotado por $f^{min}\left(x\right)$ de la longitud de la cola x que además converge monótonamente en distribución a F^* cuando $x\rightarrow\infty$

Un sistema de visitas o sistema de colas consiste en un cierto número de filas o colas atendidas por un solo servidor en un orden determinado, estos se puede aplicar en situaciones en las cuales varios tipos de usuarios intentan tener acceso a una fuente en común que está disponible para un solo tipo de usuario a la vez.

Un sistema de visitas consiste en varias colas a las cuales los usuarios llegan conforme a un proceso poisson con tasa λ_i; la capacidad de las mismas, es decir, el número de lugares disponibles; el número de servidores que llegan a la cola correspondiente para dar servicio a los usuarios; la manera en que los servidores dan servicio; el tiempo que tarda el servidor en ir de una a otra cola, así como el orden y la disciplina de servicio de la cola; la política de servicio determina cuales y cuantos usuarios serán atendidos durante la visita del servidor a la cola.

Las políticas más comunes son las de tipo exhaustivo que consiste en que el servidor continuará trabajando hasta que la cola quede vacía; y la política cerrada, bajo la cual serán atendidos exactamente aquellos que estaban presentes al momento en que llegó el servidor a la cola. El esquema de ruta determina en que orden el servidor visitara a las colas. La decisión del servidor sobre la próxima cola que visitará puede depender de la información disponible para el servidor, por ejemplo las longitudes de las colas.

El servidor usualmente incurrirá en tiempos de traslado para ir de una cola a otra. Un sistema de visitas puede expresarse en un par de parámetros: el número de colas, que usualmente se denotará por N, y el tráfico característico de las colas, que consiste de los procesos de arribo y los procesos de servicio, la figura (**??**) caracteriza a estos sistemas.

Algunas medidas de desempeño para los usuarios son los valores promedio de espera para ser atendidos, de servicio, de permanencia total en el sistema; mientras que para el servidor son los valores promedio de permanencia en una cola atendiendo, de traslado entre las colas, de duración del ciclo entre dos visitas consecutivas a la misma cola, entre otras medidas de desempeño estudiaddas en la literatura.

2.2 Función Generadora de Probabilidades

Teorema 2.2.1 (Teorema de Continuidad)**.** *Supóngase que $\{X_n, n = 1, 2, 3, \ldots\}$ son variables aleatorias finitas, no negativas con valores enteros tales que $P(X_n = k) = p_k^{(n)}$, para $n = 1, 2, 3, \ldots$, $k = 0, 1, 2, \ldots$, con $\sum_{k=0}^{\infty} p_k^{(n)} = 1$, para $n = 1, 2, 3, \ldots$. Sea g_n la PGF para la variable aleatoria X_n. Entonces existe una sucesión $\{p_k\}$ tal que*

$$lim_{n\rightarrow\infty} p_k^{(n)} = p_k \text{ para } 0 < s < 1.$$

En este caso, $g(s) = \sum_{k=0}^{\infty} s^k p_k$. Además

$$\sum_{k=0}^{\infty} p_k = 1 \text{ si y sólo si } lim_{s\uparrow 1} g(s) = 1$$

Teorema 2.2.2. *Sea N una variable aleatoria con valores enteros no negativos finita tal que $P(N = k) = p_k$, para $k = 0, 1, 2, \ldots$, y $\sum_{k=0}^{\infty} p_k = P(N < \infty) = 1$. Sea Φ la PGF de N tal que $g(s) = \mathbb{E}\left[s^N\right] = \sum_{k=0}^{\infty} s^k p_k$ con $g(1) = 1$. Si $0 \leq p_1 \leq 1$ y $\mathbb{E}[N] = g^{'}(1) \leq 1$, entonces no existe solución de la ecuación $g(s) = s$ en el intervalo $[0, 1)$. Si $\mathbb{E}[N] = g^{'}(1) > 1$, lo cual implica que $0 \leq p_1 < 1$, entonces existe una única solución de la ecuación $g(s) = s$ en el intervalo $[0, 1)$.*

Teorema 2.2.3. *Si X y Y tienen PGF G_X y G_Y respectivamente, entonces,*

$$G_X(s) = G_Y(s)$$

para tod s, si y sólo si

$$P(X = k)) = P(Y = k)$$

para toda $k = 0, 1, \ldots,$, es decir, si y sólo si X y Y tienen la misma distribución de probabilidad.

Teorema 2.2.4. *Para cada n fijo, sea la sucesióin de probabilidades $\{a_{0,n}, a_{1,n}, \ldots, \}$, tales que $a_{k,n} \geq 0$ para toda $k = 0, 1, 2, \ldots$, y $\sum_{k\geq 0} a_{k,n} = 1$, y sea $G_n(s)$ la correspondiente función generadora, $G_n(s) = \sum_{k\geq 0} a_{k,n} s^k$. De modo que para cada valor fijo de k*

$$lim_{n\rightarrow\infty} a_{k,n} = a_k,$$

es decir converge en distribución, es necesario y suficiente que para cada valor fijo $s \in [0,)$,

$$lim_{n\rightarrow\infty} G_n(s) = G(s),$$

donde $G(s) = \sum_{k\geq 0} p_k s^k$, para cualquier la función generadora del límite de la sucesión.

Teorema 2.2.5 (Teorema de Abel)**.** *Sea* $G(s) = \sum_{k\geq 0} a_k s^k$ *para cualquier* $\{p_0, p_1, \ldots, \}$*, tales que* $p_k \geq 0$ *para toda* $k = 0, 1, 2, \ldots,$*. Entonces* $G(s)$ *es continua por la derecha en* $s = 1$*, es decir*

$$lim_{s\uparrow 1} G(s) = \sum_{k\geq 0} p_k = G(),$$

sin importar si la suma es finita o no.

Nota 2.2.1. *El radio de Convergencia para cualquier PGF es* $R \geq 1$*, entonces, el Teorema de Abel nos dice que aún en el peor escenario, cuando* $R = 1$*, aún se puede confiar en que la PGF será continua en* $s = 1$*, en contraste, no se puede asegurar que la PGF será continua en el límite inferior* $-R$*, puesto que la PGF es simétrica alrededor del cero: la PGF converge para todo* $s \in (-R, R)$*, y no lo hace para* $s < -R$ *o* $s > R$*. Además nos dice que podemos escribir* $G_X(1)$ *como una abreviación de* $lim_{s\uparrow 1} G_X(s)$*.*

Entonces si suponemos que la diferenciación término a término está permitida, entonces

$$G_X^{'}(s) = \sum_{x=1}^{\infty} x s^{x-1} p_x$$

el Teorema de Abel nos dice que

$$\begin{aligned}
\mathbb{E}(X) &= \lim_{s\uparrow 1} G_X^{'}(s) : \\
\mathbb{E}[X] &= = \sum_{x=1}^{\infty} x p_x = G_X^{'}(1) \\
&= \lim_{s\uparrow 1} G_X^{'}(s),
\end{aligned}$$

dado que el Teorema de Abel se aplica a

$$G_X^{'}(s) = \sum_{x=1}^{\infty} x s^{x-1} p_x,$$

estableciendo así que $G_X^{'}(s)$ es continua en $s = 1$. Sin el Teorema de Abel no se podría asegurar que el límite de $G_X^{'}(s)$ conforme $s \uparrow 1$ sea la respuesta correcta para $\mathbb{E}[X]$.

Nota 2.2.2. *La PGF converge para todo* $|s| < R$*, para algún* R*. De hecho la PGF converge absolutamente si* $|s| < R$*. La PGF además converge uniformemente en conjuntos de la forma* $\left\{s : |s| < R^{'}\right\}$*, donde* $R^{'} < R$*, es decir,* $\forall \epsilon > 0, \exists n_0 \in \mathbb{Z}$ *tal que* $\forall s$*, con* $|s| < R^{'}$*, y* $\forall n \geq n_0$*,*

$$|\sum_{x=0}^{n} s^x \mathbb{P}(X = x) - G_X(s)| < \epsilon.$$

De hecho, la convergencia uniforme es la que nos permite diferenciar término a término:

$$G_X(s) = \mathbb{E}\left[s^X\right] = \sum_{x=0}^{\infty} s^x \mathbb{P}(X = x),$$

y sea $s < R$*.*

1.

$$\begin{aligned}
G_X^{'}(s) &= \frac{d}{ds}\left(\sum_{x=0}^{\infty} s^x \mathbb{P}(X = x)\right) = \sum_{x=0}^{\infty} \frac{d}{ds}\left(s^x \mathbb{P}(X = x)\right) \\
&= \sum_{x=0}^{n} x s^{x-1} \mathbb{P}(X = x).
\end{aligned}$$

2.

$$\begin{aligned}\int_a^b G_X(s)\,ds &= \int_a^b \left(\sum_{x=0}^{\infty} s^x \mathbb{P}(X=x)\right) ds = \sum_{x=0}^{\infty}\left(\int_a^b s^x \mathbb{P}(X=x)\,ds\right)\\ &= \sum_{x=0}^{\infty} \frac{s^{x+1}}{x+1}\mathbb{P}(X=x),\end{aligned}$$

para $-R < a < b < R$.

Teorema 2.2.6 (Teorema de Convergencia Monótona para PGF)**.** *Sean X y X_n variables aleatorias no negativas, con valores en los enteros, finitas, tales que*

$$lim_{n\rightarrow\infty} G_{X_n}(s) = G_X(s)$$

para $0 \leq s \leq 1$*, entonces*

$$lim_{n\rightarrow\infty} P(X_n = k) = P(X = k),$$

para $k = 0, 1, 2, \ldots$.

El teorema anterior requiere del siguiente lema

Lema 2.2.1. *Sean $a_{n,k} \in \mathbb{Z}^+$, $n \in \mathbb{N}$ constantes no negativas con $\sum_{k\geq 0} a_{k,n} \leq 1$. Supóngase que para $0 \leq s \leq 1$, se tiene*

$$a_n(s) = \sum_{k=0}^{\infty} a_{k,n} s^k \rightarrow a(s) = \sum_{k=0}^{\infty} a_k s^k.$$

Entonces

$$a_{0,n} \rightarrow a_0.$$

Consideremos un sistema que consta de únicamente un servidor y una sola cola, a la cual los usuarios arriban conforme a un proceso poisson cuya tasa promedio de llegada es $1/\lambda$; la tasa promedio con la cual el servidor da servicio es $1/\mu$, además los tiempos entre arribos y los tiempos de servicio son independientes entre sí.

Se define la carga de tráfico $\rho := \frac{\lambda}{\mu}$, para este modelo existe un teorema que nos dice la relación que hay entre el valor de ρ y la estabilidad de la cola:

Proposición 2.2.1. *La cola $M/M/1$ con carga de tráfico ρ, es estable si y sólo si $\rho < 1$.*

Este teorema nos permite determinar las principales medidas de desempeño: Tiempo de espera en el sistema, W, el número esperado de clientes en el sistema, L, además de los tiempos promedio e espera tanto en la cola como de servicio, s representa el tiempo de servicio para un cliente:

$$L = \frac{\rho}{1-\rho}, \tag{2.14}$$

$$W = \frac{1}{\mu-\lambda}, \tag{2.15}$$

$$W_q = \mathbb{E}[s]\frac{\rho}{1-\rho}, \text{ y} \tag{2.16}$$

$$L_q = \frac{\rho^2}{1-\rho}. \tag{2.17}$$

Esta es la idea general, poder determinar la principales medidas de desempeño para un sistema de colas o sistema de visitas, para este fin es necesario realizar los siguientes supuestos. En teoría de colas hay casos particulares, para los cuales es posible determinar específicamente medidas de desempeño del sistema bajo condiciones de estabilidad, tales como los tiempos promedio de espera y de servicio, tanto en el sistema como en cada una de las colas.

En teoría de colas hay casos particulares, para los cuales es posible determinar específicamente medidas de desempeño del sistema bajo condiciones de estabilidad, tales como los tiempos promedio de espera y de servicio, tanto en el sistema como en cada una de las colas. Se considerarán intervalos de tiempo de la forma $[t, t+1]$. Los usuarios arriban por paquetes de manera independiente del resto de las colas. Se define el grupo de usuarios que llegan a cada una de las colas del sistema 1, caracterizadas por Q_1 y Q_2 respectivamente, en el intervalo de tiempo $[t, t+1]$ por $X_1(t), X_2(t)$.

Para cada uno de los procesos anteriores se define su Función Generadora de Probabilidades (PGF):

$$P_1(z_1) = \mathbb{E}\left[z_1^{X_1(t)}\right], \quad P_2(z_2) = \mathbb{E}\left[z_2^{X_2(t)}\right].$$

Con primer momento definidos por

$$\begin{aligned}
\mu_1 &= \mathbb{E}[X_1(t)] = P_1^{(1)}(1), \\
\mu_2 &= \mathbb{E}[X_2(t)] = P_2^{(1)}(1).
\end{aligned}$$

En lo que respecta al servidor, en términos de los tiempos de visita a cada una de las colas, se denotarán por τ_1, τ_2 para Q_1, Q_2 respectivamente; y a los tiempos en que el servidor termina de atender en las colas Q_1, Q_2, se les denotará por $\overline{\tau}_1, \overline{\tau}_2$ respectivamente. Entonces, los tiempos de servicio están dados por las diferencias $\overline{\tau}_1 - \tau_1, \overline{\tau}_2 - \tau_2$ para Q_1, Q_2. Análogamente los tiempos de traslado del servidor desde el momento en que termina de atender a una cola y llega a la siguiente para comenzar a dar servicio están dados por $\tau_2 - \overline{\tau}_1, \tau_1 - \overline{\tau}_2$.

La FGP para estos tiempos de traslado están dados por

$$R_1(z_1) = \mathbb{E}\left[z_1^{\tau_2 - \overline{\tau}_1}\right], \quad R_2(z_2) = \mathbb{E}\left[z_2^{\tau_1 - \overline{\tau}_2}\right],$$

y al igual que como se hizo con anterioridad

$$r_1 = R_1^{(1)}(1) = \mathbb{E}[\tau_2 - \overline{\tau}_1], \quad r_2 = R_2^{(1)}(1) = \mathbb{E}[\tau_1 - \overline{\tau}_2],$$

Sean α_1, α_2 el número de usuarios que arriban en grupo a la cola Q_1 y Q_2 respectivamente. Sus PGF's están definidas como

$$A_1(z) = \mathbb{E}\left[z^{\alpha_1(t)}\right], \quad A_2(z) = \mathbb{E}\left[z^{\alpha_2(t)}\right].$$

Su primer momento está dado por

$$\lambda_1 = \mathbb{E}[\alpha_1(t)] = A_1^{(1)}(1), \quad \lambda_2 = \mathbb{E}[\alpha_2(t)] = A_2^{(1)}(1).$$

Sean β_1, β_2 el número de usuarios que arriban en el grupo α_1, α_2 a la cola Q_1 y Q_2, respectivamente, de igual manera se definen sus PGF's

$$B_1(z) = \mathbb{E}\left[z^{\beta_1(t)}\right], \quad B_2(z) = \mathbb{E}\left[z^{\beta_2(t)}\right],$$

con

$$b_1 = \mathbb{E}[\beta_1(t)] = B_1^{(1)}(1), \quad b_2 = \mathbb{E}[\beta_2(t)] = B_2^{(1)}(1).$$

La distribución para el número de grupos que arriban al sistema en cada una de las colas se definen por:

$$P_1(z_1) = A_1[B_1(z_1)] = \mathbb{E}\left[B_1(z_1)^{\alpha_1(t)}\right], \quad P_2(z_1) = A_1[B_1(z_1)] = \mathbb{E}\left[B_1(z_1)^{\alpha_1(t)}\right],$$

entonces

$$\begin{aligned} P_1^{(1)}(1) &= \mathbb{E}\left[\alpha_1(t)\, B_1^{(1)}(1)\right] = B_1^{(1)}(1)\, \mathbb{E}\left[\alpha_1(t)\right] = \lambda_1 b_1 \\ P_2^{(1)}(1) &= \mathbb{E}\left[\alpha_2(t)\, B_2^{(1)}(1)\right] = B_2^{(1)}(1)\, \mathbb{E}\left[\alpha_2(t)\right] = \lambda_2 b_2. \end{aligned}$$

De lo desarrollado hasta ahora se tiene lo siguiente

$$\begin{aligned} & \mathbb{E}\left[z_1^{L_1(\overline{\tau}_1)} z_2^{L_2(\overline{\tau}_1)}\right] = \mathbb{E}\left[z_2^{L_2(\overline{\tau}_1)}\right] = \mathbb{E}\left[z_2^{L_2(\tau_1)+X_2(\overline{\tau}_1-\tau_1)}\right] \\ = & \mathbb{E}\left[\left\{z_2^{L_2(\tau_1)}\right\}\left\{z_2^{X_2(\overline{\tau}_1-\tau_1)}\right\}\right] = \mathbb{E}\left[\left\{z_2^{L_2(\tau_1)}\right\}\left\{P_2(z_2)\right\}^{\overline{\tau}_1-\tau_1}\right] \\ = & \mathbb{E}\left[\left\{z_2^{L_2(\tau_1)}\right\}\left\{\theta_1(P_2(z_2))\right\}^{L_1(\tau_1)}\right] = F_1(\theta_1(P_2(z_2)), z_2) \end{aligned}$$

es decir

$$\mathbb{E}\left[z_1^{L_1(\overline{\tau}_1)} z_2^{L_2(\overline{\tau}_1)}\right] = F_1(\theta_1(P_2(z_2)), z_2). \tag{2.18}$$

Procediendo de manera análoga para $\overline{\tau}_2$:

$$\begin{aligned} \mathbb{E}\left[z_1^{L_1(\overline{\tau}_2)} z_2^{L_2(\overline{\tau}_2)}\right] &= \mathbb{E}\left[z_1^{L_1(\overline{\tau}_2)}\right] = \mathbb{E}\left[z_1^{L_1(\tau_2)+X_1(\overline{\tau}_2-\tau_2)}\right] = \mathbb{E}\left[\left\{z_1^{L_1(\tau_2)}\right\}\left\{z_1^{X_1(\overline{\tau}_2-\tau_2)}\right\}\right] \\ &= \mathbb{E}\left[\left\{z_1^{L_1(\tau_2)}\right\}\left\{P_1(z_1)\right\}^{\overline{\tau}_2-\tau_2}\right] = \mathbb{E}\left[\left\{z_1^{L_1(\tau_2)}\right\}\left\{\theta_2(P_1(z_1))\right\}^{L_2(\tau_2)}\right] \\ &= F_2(z_1, \theta_2(P_1(z_1))) \end{aligned}$$

por tanto

$$\mathbb{E}\left[z_1^{L_1(\overline{\tau}_2)} z_2^{L_2(\overline{\tau}_2)}\right] = F_2(z_1, \theta_2(P_1(z_1))) \tag{2.19}$$

Ahora, para el intervalo de tiempo $[\overline{\tau}_1, \tau_2]$ y $[\overline{\tau}_2, \tau_1]$, los arribos de los usuarios modifican el número de usuarios que llegan a las colas, es decir, los procesos $L_1(t)$ y $L_2(t)$. La PGF para el número de arribos a todas las estaciones durante el intervalo $[\overline{\tau}_1, \tau_2]$ cuya distribución está especificada por la distribución compuesta $R_1(\mathbf{z}), R_2(\mathbf{z})$:

$$\begin{aligned} R_1(\mathbf{z}) = R_1\left(\prod_{i=1}^{2} P(z_i)\right) = \mathbb{E}\left[\left\{\prod_{i=1}^{2} P(z_i)\right\}^{\tau_2-\overline{\tau}_1}\right] \\ R_2(\mathbf{z}) = R_2\left(\prod_{i=1}^{2} P(z_i)\right) = \mathbb{E}\left[\left\{\prod_{i=1}^{2} P(z_i)\right\}^{\tau_1-\overline{\tau}_2}\right] \end{aligned}$$

Dado que los eventos en $[\tau_1, \overline{\tau}_1]$ y $[\overline{\tau}_1, \tau_2]$ son independientes, la PGF conjunta para el número de usuarios en el sistema al tiempo $t = \tau_2$ la PGF conjunta para el número de usuarios en el sistema están dadas por

$$\begin{aligned} F_1(\mathbf{z}) &= R_2\left(\prod_{i=1}^{2} P(z_i)\right) F_2(z_1, \theta_2(P_1(z_1))) \\ F_2(\mathbf{z}) &= R_1\left(\prod_{i=1}^{2} P(z_i)\right) F_1(\theta_1(P_2(z_2)), z_2) \end{aligned}$$

Entonces debemos de determinar las siguientes expresiones:

$$
\begin{aligned}
f_1(1) &= \tfrac{\partial F_1(\mathbf{z})}{\partial z_1}|_{\mathbf{z}=1}, \quad f_1(2) = \tfrac{\partial F_1(\mathbf{z})}{\partial z_2}|_{\mathbf{z}=1},\\
f_2(1) &= \tfrac{\partial F_2(\mathbf{z})}{\partial z_1}|_{\mathbf{z}=1}, \quad f_2(2) = \tfrac{\partial F_2(\mathbf{z})}{\partial z_2}|_{\mathbf{z}=1},
\end{aligned}
$$

calculando las derivadas parciales

$$
\begin{aligned}
\frac{\partial R_1(\mathbf{z})}{\partial z_1}|_{\mathbf{z}=1} &= R_1^{(1)}(1)\,P_1^{(1)}(1)\\
\frac{\partial R_1(\mathbf{z})}{\partial z_2}|_{\mathbf{z}=1} &= R_1^{(1)}(1)\,P_2^{(1)}(1)\\
\frac{\partial R_2(\mathbf{z})}{\partial z_1}|_{\mathbf{z}=1} &= R_2^{(1)}(1)\,P_1^{(1)}(1)\\
\frac{\partial R_2(\mathbf{z})}{\partial z_2}|_{\mathbf{z}=1} &= R_2^{(1)}(1)\,P_2^{(1)}(1)
\end{aligned}
$$

igualando a cero

$$
\begin{aligned}
\frac{\partial}{\partial z_1}F_1(\theta_1(P_2(z_2)),z_2) &= 0\\
\frac{\partial}{\partial z_2}F_1(\theta_1(P_2(z_2)),z_2) &= \frac{\partial F_1}{\partial z_2}+\frac{\partial F_1}{\partial z_1}\theta_1^{(1)}P_2^{(1)}(1)\\
\frac{\partial}{\partial z_1}F_2(z_1,\theta_2(P_1(z_1))) &= \frac{\partial F_2}{\partial z_1}+\frac{\partial F_2}{\partial z_2}\theta_2^{(1)}P_1^{(1)}(1)\\
\frac{\partial}{\partial z_2}F_2(z_1,\theta_2(P_1(z_1))) &= 0.
\end{aligned}
$$

Por lo tanto de las dos secciones anteriores se tiene que:

$$
\begin{aligned}
\frac{\partial F_1}{\partial z_1} &= \frac{\partial R_2}{\partial z_1}|_{\mathbf{z}=1}+\frac{\partial F_2}{\partial z_1}|_{\mathbf{z}=1} = R_2^{(1)}(1)\,P_1^{(1)}(1)+f_2(1)+f_2(2)\,\theta_2^{(1)}(1)\,P_1^{(1)}(1)\\
\frac{\partial F_1}{\partial z_2} &= \frac{\partial R_2}{\partial z_2}|_{\mathbf{z}=1}+\frac{\partial F_2}{\partial z_2}|_{\mathbf{z}=1} = R_2^{(1)}(1)\,P_2^{(1)}(1)\\
\frac{\partial F_2}{\partial z_1} &= \frac{\partial R_1}{\partial z_1}|_{\mathbf{z}=1}+\frac{\partial F_1}{\partial z_1}|_{\mathbf{z}=1} = R_1^{(1)}(1)\,P_1^{(1)}(1)\\
\frac{\partial F_2}{\partial z_2} &= \frac{\partial R_1}{\partial z_2}|_{\mathbf{z}=1}+\frac{\partial F_1}{\partial z_2}|_{\mathbf{z}=1} = R_1^{(1)}(1)\,P_2^{(1)}(1)+f_1(1)\,\theta_1^{(1)}(1)\,P_2^{(1)}(1)
\end{aligned}
$$

El cual se puede escribir en forma equivalente:

$$
\begin{aligned}
f_1(1) &= r_2\mu_1+f_2(1)+f_2(2)\,\frac{\mu_1}{1-\mu_2}\\
f_1(2) &= r_2\mu_2\\
f_2(1) &= r_1\mu_1\\
f_2(2) &= r_1\mu_2+f_1(2)+f_1(1)\,\frac{\mu_2}{1-\mu_1}
\end{aligned}
$$

De donde:

$$
\begin{aligned}
f_1(1) &= \mu_1\left[r_2+\frac{f_2(2)}{1-\mu_2}\right]+f_2(1)\\
f_2(2) &= \mu_2\left[r_1+\frac{f_1(1)}{1-\mu_1}\right]+f_1(2)
\end{aligned}
$$

Resolviendo para $f_1(1)$:

$$\begin{aligned} f_1(1) &= r_2\mu_1 + f_2(2)\frac{\mu_1}{1-\mu_2} = r_2\mu_1 + r_1\mu_1 + f_2(2)\frac{\mu_1}{1-\mu_2} \\ &= \mu_1(r_2+r_1) + f_2(2)\frac{\mu_1}{1-\mu_2} = \mu_1\left(r + \frac{f_2(2)}{1-\mu_2}\right), \end{aligned}$$

entonces

$$\begin{aligned} f_2(2) &= \mu_2\left(r_1 + \frac{f_1(1)}{1-\mu_1}\right) + f_1(2) = \mu_2\left(r_1 + \frac{f_1(1)}{1-\mu_1}\right) + r_2\mu_2 \\ &= \mu_2\left[r_1 + r_2 + \frac{f_1(1)}{1-\mu_1}\right] = \mu_2\left[r + \frac{f_1(1)}{1-\mu_1}\right] \\ &= \mu_2 r + \mu_1\left(r + \frac{f_2(2)}{1-\mu_2}\right)\frac{\mu_2}{1-\mu_1} \\ &= \mu_2 r + \mu_2\frac{r\mu_1}{1-\mu_1} + f_2(2)\frac{\mu_1\mu_2}{(1-\mu_1)(1-\mu_2)} \\ &= \mu_2\left(r + \frac{r\mu_1}{1-\mu_1}\right) + f_2(2)\frac{\mu_1\mu_2}{(1-\mu_1)(1-\mu_2)} \\ &= \mu_2\left(\frac{r}{1-\mu_1}\right) + f_2(2)\frac{\mu_1\mu_2}{(1-\mu_1)(1-\mu_2)} \end{aligned}$$

entonces

$$\begin{aligned} f_2(2) - f_2(2)\frac{\mu_1\mu_2}{(1-\mu_1)(1-\mu_2)} &= \mu_2\left(\frac{r}{1-\mu_1}\right) \\ f_2(2)\left(1 - \frac{\mu_1\mu_2}{(1-\mu_1)(1-\mu_2)}\right) &= \mu_2\left(\frac{r}{1-\mu_1}\right) \\ f_2(2)\left(\frac{1-\mu_1-\mu_2+\mu_1\mu_2-\mu_1\mu_2}{(1-\mu_1)(1-\mu_2)}\right) &= \mu_2\left(\frac{r}{1-\mu_1}\right) \\ f_2(2)\left(\frac{1-\mu}{(1-\mu_1)(1-\mu_2)}\right) &= \mu_2\left(\frac{r}{1-\mu_1}\right) \end{aligned}$$

por tanto

$$\begin{aligned} f_2(2) &= \frac{r\frac{\mu_2}{1-\mu_1}}{\frac{1-\mu}{(1-\mu_1)(1-\mu_2)}} = \frac{r\mu_2(1-\mu_1)(1-\mu_2)}{(1-\mu_1)(1-\mu)} \\ &= \frac{\mu_2(1-\mu_2)}{1-\mu}r = r\mu_2\frac{1-\mu_2}{1-\mu}. \end{aligned}$$

es decir

$$f_2(2) = r\mu_2\frac{1-\mu_2}{1-\mu}. \tag{2.20}$$

Entonces

$$\begin{aligned} f_1(1) &= \mu_1 r + f_2(2)\frac{\mu_1}{1-\mu_2} = \mu_1 r + \left(\frac{\mu_2(1-\mu_2)}{1-\mu}r\right)\frac{\mu_1}{1-\mu_2} \\ &= \mu_1 r + \mu_1 r\left(\frac{\mu_2}{1-\mu}\right) = \mu_1 r\left[1 + \frac{\mu_2}{1-\mu}\right] \\ &= r\mu_1\frac{1-\mu_1}{1-\mu} \end{aligned}$$

2.3 El problema de la ruina del jugador

Supongamos que se tiene un jugador que cuenta con un capital inicial de $\tilde{L}_0 \geq 0$ unidades, esta persona realiza una serie de dos juegos simultáneos e independientes de manera sucesiva, dichos eventos son independientes e idénticos entre sí para cada realización. Para $n \geq 0$ fijo, la ganancia en el n-ésimo juego es $\tilde{X}_n = X_n + Y_n$ unidades de las cuales se resta una cuota de 1 unidad por cada juego simultáneo, es decir, se restan dos unidades por cada juego realizado. En términos de la teoría de colas puede pensarse como el número de usuarios que llegan a una cola vía dos procesos de arribo distintos e independientes entre sí. Su Función Generadora de Probabilidades (FGP) está dada por $F(z) = \mathbb{E}\left[z^{\tilde{L}_0}\right]$ para $z \in \mathbb{C}$, además

$$\tilde{P}(z) = \mathbb{E}\left[z^{\tilde{X}_n}\right] = \mathbb{E}\left[z^{X_n+Y_n}\right] = \mathbb{E}\left[z^{X_n} z^{Y_n}\right] = \mathbb{E}\left[z^{X_n}\right]\mathbb{E}\left[z^{Y_n}\right] = P(z)\check{P}(z),$$

con $\tilde{\mu} = \mathbb{E}\left[\tilde{X}_n\right] = \tilde{P}[z] < 1$.

Sea $\tilde{L}_n$ el capital remanente después del n-ésimo juego. Entonces

$$\tilde{L}_n = \tilde{L}_0 + \tilde{X}_1 + \tilde{X}_2 + \cdots + \tilde{X}_n - 2n.$$

La ruina del jugador ocurre después del n-ésimo juego, es decir, la cola se vacía después del n-ésimo juego, entonces sea T definida como $T = min\left\{\tilde{L}_n = 0\right\}$. Si $\tilde{L}_0 = 0$, entonces claramente $T = 0$. En este sentido T puede interpretarse como la longitud del periodo de tiempo que el servidor ocupa para dar servicio en la cola, comenzando con $\tilde{L}_0$ grupos de usuarios presentes en la cola, quienes arribaron conforme a un proceso dado por $\tilde{P}(z)$.

Sea $g_{n,k}$ la probabilidad del evento de que el jugador no caiga en ruina antes del n-ésimo juego, y que además tenga un capital de k unidades antes del n-ésimo juego, es decir, dada $n \in \{1, 2, \ldots\}$ y $k \in \{0, 1, 2, \ldots\}$

$$g_{n,k} := P\left\{\tilde{L}_j > 0, j = 1, \ldots, n, \tilde{L}_n = k\right\}, \tag{2.21}$$

la cual además se puede escribir como:

$$\begin{aligned}
g_{n,k} &= P\left\{\tilde{L}_j > 0, j = 1, \ldots, n, \tilde{L}_n = k\right\} = \sum_{j=1}^{k+1} g_{n-1,j} P\left\{\tilde{X}_n = k - j + 1\right\}\\
&= \sum_{j=1}^{k+1} g_{n-1,j} P\left\{X_n + Y_n = k - j + 1\right\} = \sum_{j=1}^{k+1}\sum_{l=1}^{j} g_{n-1,j} P\left\{X_n + Y_n = k - j + 1, Y_n = l\right\}\\
&= \sum_{j=1}^{k+1}\sum_{l=1}^{j} g_{n-1,j} P\left\{X_n + Y_n = k - j + 1 | Y_n = l\right\} P\left\{Y_n = l\right\}\\
&= \sum_{j=1}^{k+1}\sum_{l=1}^{j} g_{n-1,j} P\left\{X_n = k - j - l + 1\right\} P\left\{Y_n = l\right\},
\end{aligned}$$

es decir

$$g_{n,k} = \sum_{j=1}^{k+1}\sum_{l=1}^{j} g_{n-1,j} P\left\{X_n = k - j - l + 1\right\} P\left\{Y_n = l\right\}. \tag{2.22}$$

Además

$$g_{0,k} = P\left\{\tilde{L}_0 = k\right\}. \tag{2.23}$$

Se definen las siguientes FGP:

$$G_n(z) = \sum_{k=0}^{\infty} g_{n,k} z^k, \text{ para } n = 0, 1, \ldots, \tag{2.24}$$

y

$$G(z,w) = \sum_{n=0}^{\infty} G_n(z)\, w^n, z, w \in \mathbb{C}. \tag{2.25}$$

En particular para $k = 0$,

$$g_{n,0} = G_n(0) = P\left\{\tilde{L}_j > 0, \text{ para } j < n, \text{ y } \tilde{L}_n = 0\right\} = P\{T = n\},$$

además

$$G(0,w) = \sum_{n=0}^{\infty} G_n(0)\, w^n = \sum_{n=0}^{\infty} P\{T = n\}\, w^n = \mathbb{E}\left[w^T\right]$$

la cuál resulta ser la FGP del tiempo de ruina T.

Proposición 2.3.1. *Sean $z, w \in \mathbb{C}$ y sea $n \geq 0$ fijo. Para $G_n(z)$ y $G(z,w)$ definidas como en (2.24) y (2.25) respectivamente, se tiene que*

$$G_n(z) = \frac{1}{z}\left[G_{n-1}(z) - G_{n-1}(0)\right] \tilde{P}(z). \tag{2.26}$$

Además

$$G(z,w) = \frac{zF(z) - wP(z)\, G(0,w)}{z - wR(z)}, \tag{2.27}$$

con un único polo en el círculo unitario, además, el polo es de la forma $z = \theta(w)$ y satisface que

i) $\tilde{\theta}(1) = 1$,

ii) $\tilde{\theta}^{(1)}(1) = \frac{1}{1-\tilde{\mu}}$,

iii) $\tilde{\theta}^{(2)}(1) = \frac{\tilde{\mu}}{(1-\tilde{\mu})^2} + \frac{\tilde{\sigma}}{(1-\tilde{\mu})^3}$.

Finalmente, además se cumple que

$$\mathbb{E}\left[w^T\right] = G(0,w) = F\left[\tilde{\theta}(w)\right]. \tag{2.28}$$

Proof. Multiplicando las ecuaciones (2.22) y (2.23) por el término z^k:

$$\begin{aligned}
g_{n,k} z^k &= \sum_{j=1}^{k+1} \sum_{l=1}^{j} g_{n-1,j} P\{X_n = k - j - l + 1\} P\{Y_n = l\} z^k, \\
g_{0,k} z^k &= P\left\{\tilde{L}_0 = k\right\} z^k,
\end{aligned}$$

ahora sumamos sobre k

$$\begin{aligned}
\sum_{k=0}^{\infty} g_{n,k} z^{k} &= \sum_{k=0}^{\infty}\sum_{j=1}^{k+1}\sum_{l=1}^{j} g_{n-1,j} P\left\{X_n = k-j-l+1\right\} P\left\{Y_n = l\right\} z^{k}\\
&= \sum_{k=0}^{\infty} z^{k}\sum_{j=1}^{k+1}\sum_{l=1}^{j} g_{n-1,j} P\left\{X_n = k-(j+l-1)\right\} P\left\{Y_n = l\right\}\\
&= \sum_{k=0}^{\infty} z^{k+(j+l-1)-(j+l-1)}\sum_{j=1}^{k+1}\sum_{l=1}^{j} g_{n-1,j} P\left\{X_n = k-(j+l-1)\right\} P\left\{Y_n = l\right\}\\
&= \sum_{k=0}^{\infty}\sum_{j=1}^{k+1}\sum_{l=1}^{j} g_{n-1,j} z^{j-1} P\left\{X_n = k-(j+l-1)\right\} z^{k-(j+l-1)} P\left\{Y_n = l\right\} z^{l}\\
&= \sum_{j=1}^{\infty}\sum_{l=1}^{j} g_{n-1,j} z^{j-1}\sum_{k=j+l-1}^{\infty} P\left\{X_n = k-(j+l-1)\right\} z^{k-(j+l-1)} P\left\{Y_n = l\right\} z^{l}\\
&= \sum_{j=1}^{\infty} g_{n-1,j} z^{j-1}\sum_{l=1}^{j}\sum_{k=j+l-1}^{\infty} P\left\{X_n = k-(j+l-1)\right\} z^{k-(j+l-1)} P\left\{Y_n = l\right\} z^{l}\\
&= \sum_{j=1}^{\infty} g_{n-1,j} z^{j-1}\sum_{k=j+l-1}^{\infty}\sum_{l=1}^{j} P\left\{X_n = k-(j+l-1)\right\} z^{k-(j+l-1)} P\left\{Y_n = l\right\} z^{l}\\
&= \sum_{j=1}^{\infty} g_{n-1,j} z^{j-1}\sum_{k=j+l-1}^{\infty}\sum_{l=1}^{j} P\left\{X_n = k-(j+l-1)\right\} z^{k-(j+l-1)}\sum_{l=1}^{j} P\left\{Y_n = l\right\} z^{l}\\
&= \sum_{j=1}^{\infty} g_{n-1,j} z^{j-1}\sum_{l=1}^{\infty} P\left\{Y_n = l\right\} z^{l}\sum_{k=j+l-1}^{\infty}\sum_{l=1}^{j} P\left\{X_n = k-(j+l-1)\right\} z^{k-(j+l-1)}\\
&= \frac{1}{z}\left[G_{n-1}\left(z\right) - G_{n-1}\left(0\right)\right]\check{P}\left(z\right)\sum_{k=j+l-1}^{\infty}\sum_{l=1}^{j} P\left\{X_n = k-(j+l-1)\right\} z^{k-(j+l-1)}\\
&= \frac{1}{z}\left[G_{n-1}\left(z\right) - G_{n-1}\left(0\right)\right]\check{P}\left(z\right) P\left(z\right) = \frac{1}{z}\left[G_{n-1}\left(z\right) - G_{n-1}\left(0\right)\right]\tilde{P}\left(z\right),
\end{aligned}$$

es decir la ecuación (2.24) se puede reescribir como

$$G_n\left(z\right) = \frac{1}{z}\left[G_{n-1}\left(z\right) - G_{n-1}\left(0\right)\right]\tilde{P}\left(z\right). \tag{2.29}$$

Por otra parte recordemos la ecuación (2.24)

$$G_n\left(z\right) = \sum_{k=0}^{\infty} g_{n,k} z^{k},\ \text{entonces}\ \frac{G_n\left(z\right)}{z} = \sum_{k=1}^{\infty} g_{n,k} z^{k-1},$$

por lo tanto utilizando la ecuación (2.29):

$$\begin{aligned}
G\left(z,w\right) &= \sum_{n=0}^{\infty} G_n\left(z\right) w^{n} = G_0\left(z\right) + \sum_{n=1}^{\infty} G_n\left(z\right) w^{n} = F\left(z\right) + \sum_{n=0}^{\infty}\left[G_n\left(z\right) - G_n\left(0\right)\right] w^{n}\frac{\tilde{P}\left(z\right)}{z}\\
&= F\left(z\right) + \frac{w}{z}\sum_{n=0}^{\infty}\left[G_n\left(z\right) - G_n\left(0\right)\right] w^{n-1}\tilde{P}\left(z\right)
\end{aligned}$$

es decir

$$G\left(z,w\right) = F\left(z\right) + \frac{w}{z}\left[G\left(z,w\right) - G\left(0,w\right)\right]\tilde{P}\left(z\right),$$

entonces

$$\begin{aligned} G(z,w) = F(z) + \frac{w}{z}\left[G(z,w) - G(0,w)\right]\tilde{P}(z) &= F(z) + \frac{w}{z}\tilde{P}(z)\,G(z,w) - \frac{w}{z}\tilde{P}(z)\,G(0,w) \\ &\Leftrightarrow \\ G(z,w)\left\{1 - \frac{w}{z}\tilde{P}(z)\right\} &= F(z) - \frac{w}{z}\tilde{P}(z)\,G(0,w), \end{aligned}$$

por lo tanto,

$$G(z,w) = \frac{zF(z) - w\tilde{P}(z)\,G(0,w)}{1 - w\tilde{P}(z)}. \tag{2.30}$$

Ahora $G(z,w)$ es analítica en $|z| = 1$. Sean z, w tales que $|z| = 1$ y $|w| \leq 1$, como $\tilde{P}(z)$ es FGP

$$|z - \left(z - w\tilde{P}(z)\right)| < |z| \Leftrightarrow |w\tilde{P}(z)| < |z|$$

es decir, se cumplen las condiciones del Teorema de Rouché y por tanto, z y $z - w\tilde{P}(z)$ tienen el mismo número de ceros en $|z| = 1$. Sea $z = \tilde{\theta}(w)$ la solución única de $z - w\tilde{P}(z)$, es decir

$$\tilde{\theta}(w) - w\tilde{P}\left(\tilde{\theta}(w)\right) = 0, \tag{2.31}$$

con $|\tilde{\theta}(w)| < 1$. Cabe hacer mención que $\tilde{\theta}(w)$ es la FGP para el tiempo de ruina cuando $\tilde{L}_0 = 1$. Considerando la ecuación (2.31)

$$\begin{aligned} 0 &= \frac{\partial}{\partial w}\tilde{\theta}(w)|_{w=1} - \frac{\partial}{\partial w}\left\{w\tilde{P}\left(\tilde{\theta}(w)\right)\right\}|_{w=1} = \tilde{\theta}^{(1)}(w)|_{w=1} - \frac{\partial}{\partial w}w\left\{\tilde{P}\left(\tilde{\theta}(w)\right)\right\}|_{w=1} \\ &- w\frac{\partial}{\partial w}\tilde{P}\left(\tilde{\theta}(w)\right)|_{w=1} = \tilde{\theta}^{(1)}(1) - \tilde{P}\left(\tilde{\theta}(1)\right) - w\left\{\frac{\partial\tilde{P}\left(\tilde{\theta}(w)\right)}{\partial\tilde{\theta}(w)} \cdot \frac{\partial\tilde{\theta}(w)}{\partial w}|_{w=1}\right\} \\ &= \tilde{\theta}^{(1)}(1) - \tilde{P}\left(\tilde{\theta}(1)\right) - \tilde{P}^{(1)}\left(\tilde{\theta}(1)\right) \cdot \tilde{\theta}^{(1)}(1), \end{aligned}$$

luego

$$\tilde{P}\left(\tilde{\theta}(1)\right) = \tilde{\theta}^{(1)}(1) - \tilde{P}^{(1)}\left(\tilde{\theta}(1)\right) \cdot \tilde{\theta}^{(1)}(1) = \tilde{\theta}^{(1)}(1)\left(1 - \tilde{P}^{(1)}\left(\tilde{\theta}(1)\right)\right),$$

por tanto

$$\tilde{\theta}^{(1)}(1) = \frac{\tilde{P}\left(\tilde{\theta}(1)\right)}{\left(1 - \tilde{P}^{(1)}\left(\tilde{\theta}(1)\right)\right)} = \frac{1}{1 - \tilde{\mu}}.$$

Ahora determinemos el segundo momento de $\tilde{\theta}(w)$, nuevamente consideremos la ecuación (2.31):

$$0 = \tilde{\theta}(w) - w\tilde{P}\left(\tilde{\theta}(w)\right) \Rightarrow 0 = \frac{\partial}{\partial w}\left\{\tilde{\theta}(w) - w\tilde{P}\left(\tilde{\theta}(w)\right)\right\} \Rightarrow 0 = \frac{\partial}{\partial w}\left\{\frac{\partial}{\partial w}\left\{\tilde{\theta}(w) - w\tilde{P}\left(\tilde{\theta}(w)\right)\right\}\right\}$$

luego se tiene

$$\begin{aligned}
&\frac{\partial}{\partial w}\left\{\frac{\partial}{\partial w}\tilde{\theta}(w)-\frac{\partial}{\partial w}\left[w\tilde{P}\left(\tilde{\theta}(w)\right)\right]\right\}=\frac{\partial}{\partial w}\left\{\frac{\partial}{\partial w}\tilde{\theta}(w)-\frac{\partial}{\partial w}\left[w\tilde{P}\left(\tilde{\theta}(w)\right)\right]\right\}\\
=&\ \frac{\partial}{\partial w}\left\{\frac{\partial\tilde{\theta}(w)}{\partial w}-\left[\tilde{P}\left(\tilde{\theta}(w)\right)+w\frac{\partial}{\partial w}P\left(\tilde{\theta}(w)\right)\right]\right\}\\
=&\ \frac{\partial}{\partial w}\left\{\frac{\partial\tilde{\theta}(w)}{\partial w}-\left(\tilde{P}\left(\tilde{\theta}(w)\right)+w\frac{\partial\tilde{P}\left(\tilde{\theta}(w)\right)}{\partial w}\frac{\partial\tilde{\theta}(w)}{\partial w}\right]\right\}\\
=&\ \frac{\partial}{\partial w}\left\{\tilde{\theta}^{(1)}(w)-\tilde{P}\left(\tilde{\theta}(w)\right)-w\tilde{P}^{(1)}\left(\tilde{\theta}(w)\right)\tilde{\theta}^{(1)}(w)\right\}\\
=&\ \frac{\partial}{\partial w}\tilde{\theta}^{(1)}(w)-\frac{\partial}{\partial w}\tilde{P}\left(\tilde{\theta}(w)\right)-\frac{\partial}{\partial w}\left[w\tilde{P}^{(1)}\left(\tilde{\theta}(w)\right)\tilde{\theta}^{(1)}(w)\right]\\
=&\ \frac{\partial}{\partial w}\tilde{\theta}^{(1)}(w)-\frac{\partial\tilde{P}\left(\tilde{\theta}(w)\right)}{\partial\tilde{\theta}(w)}\frac{\partial\tilde{\theta}(w)}{\partial w}-\tilde{P}^{(1)}\left(\tilde{\theta}(w)\right)\tilde{\theta}^{(1)}(w)-w\frac{\partial\tilde{P}^{(1)}\left(\tilde{\theta}(w)\right)}{\partial w}\tilde{\theta}^{(1)}(w)\\
-&\ w\tilde{P}^{(1)}\left(\tilde{\theta}(w)\right)\frac{\partial\tilde{\theta}^{(1)}(w)}{\partial w}\\
=&\ \tilde{\theta}^{(2)}(w)-\tilde{P}^{(1)}\left(\tilde{\theta}(w)\right)\tilde{\theta}^{(1)}(w)-\tilde{P}^{(1)}\left(\tilde{\theta}(w)\right)\tilde{\theta}^{(1)}(w)-w\tilde{P}^{(2)}\left(\tilde{\theta}(w)\right)\left(\tilde{\theta}^{(1)}(w)\right)^2\\
-&\ w\tilde{P}^{(1)}\left(\tilde{\theta}(w)\right)\tilde{\theta}^{(2)}(w)\\
=&\ \tilde{\theta}^{(2)}(w)-2\tilde{P}^{(1)}\left(\tilde{\theta}(w)\right)\tilde{\theta}^{(1)}(w)-w\tilde{P}^{(2)}\left(\tilde{\theta}(w)\right)\left(\tilde{\theta}^{(1)}(w)\right)^2-w\tilde{P}^{(1)}\left(\tilde{\theta}(w)\right)\tilde{\theta}^{(2)}(w)\\
=&\ \tilde{\theta}^{(2)}(w)\left[1-w\tilde{P}^{(1)}\left(\tilde{\theta}(w)\right)\right]-\tilde{\theta}^{(1)}(w)\left[w\tilde{\theta}^{(1)}(w)\tilde{P}^{(2)}\left(\tilde{\theta}(w)\right)+2\tilde{P}^{(1)}\left(\tilde{\theta}(w)\right)\right]
\end{aligned}$$

luego

$$\begin{aligned}
\tilde{\theta}^{(2)}(w) &\quad \left[1-w\tilde{P}^{(1)}\left(\tilde{\theta}(w)\right)\right]-\tilde{\theta}^{(1)}(w)\left[w\tilde{\theta}^{(1)}(w)\tilde{P}^{(2)}\left(\tilde{\theta}(w)\right)+2\tilde{P}^{(1)}\left(\tilde{\theta}(w)\right)\right]=0\\
\tilde{\theta}^{(2)}(w) &= \frac{\tilde{\theta}^{(1)}(w)\left[w\tilde{\theta}^{(1)}(w)\tilde{P}^{(2)}\left(\tilde{\theta}(w)\right)+2P^{(1)}\left(\tilde{\theta}(w)\right)\right]}{1-w\tilde{P}^{(1)}\left(\tilde{\theta}(w)\right)}\\
&= \frac{\tilde{\theta}^{(1)}(w)\,w\tilde{\theta}^{(1)}(w)\tilde{P}^{(2)}\left(\tilde{\theta}(w)\right)}{1-w\tilde{P}^{(1)}\left(\tilde{\theta}(w)\right)}+\frac{2\tilde{\theta}^{(1)}(w)\tilde{P}^{(1)}\left(\tilde{\theta}(w)\right)}{1-w\tilde{P}^{(1)}\left(\tilde{\theta}(w)\right)}
\end{aligned}$$

si evaluamos la expresión anterior en $w=1$:

$$\begin{aligned}
\tilde{\theta}^{(2)}(1) &= \frac{\left(\tilde{\theta}^{(1)}(1)\right)^2\tilde{P}^{(2)}\left(\tilde{\theta}(1)\right)}{1-\tilde{P}^{(1)}\left(\tilde{\theta}(1)\right)}+\frac{2\tilde{\theta}^{(1)}(1)\tilde{P}^{(1)}\left(\tilde{\theta}(1)\right)}{1-\tilde{P}^{(1)}\left(\tilde{\theta}(1)\right)}=\frac{\left(\tilde{\theta}^{(1)}(1)\right)^2\tilde{P}^{(2)}(1)}{1-\tilde{P}^{(1)}(1)}+\frac{2\tilde{\theta}^{(1)}(1)\tilde{P}^{(1)}(1)}{1-\tilde{P}^{(1)}(1)}\\
&= \frac{\left(\frac{1}{1-\tilde{\mu}}\right)^2\tilde{P}^{(2)}(1)}{1-\tilde{\mu}}+\frac{2\left(\frac{1}{1-\tilde{\mu}}\right)\tilde{\mu}}{1-\tilde{\mu}}=\frac{\tilde{P}^{(2)}(1)}{(1-\tilde{\mu})^3}+\frac{2\tilde{\mu}}{(1-\tilde{\mu})^2}=\frac{\sigma^2-\tilde{\mu}+\tilde{\mu}^2}{(1-\tilde{\mu})^3}+\frac{2\tilde{\mu}}{(1-\tilde{\mu})^2}\\
&= \frac{\sigma^2-\tilde{\mu}+\tilde{\mu}^2+2\tilde{\mu}(1-\tilde{\mu})}{(1-\tilde{\mu})^3}
\end{aligned}$$

es decir

$$\tilde{\theta}^{(2)}(1) = \frac{\sigma^2}{(1-\tilde{\mu})^3}+\frac{\tilde{\mu}}{(1-\tilde{\mu})^2}.$$

□

Corolario 2.3.1. *El tiempo de ruina del jugador tiene primer y segundo momento dados por*

$$\mathbb{E}\left[T\right] = \frac{\mathbb{E}\left[\tilde{L}_0\right]}{1-\tilde{\mu}} \tag{2.32}$$

$$Var\left[T\right] = \frac{Var\left[\tilde{L}_0\right]}{\left(1-\tilde{\mu}\right)^2}+\frac{\sigma^2\mathbb{E}\left[\tilde{L}_0\right]}{\left(1-\tilde{\mu}\right)^3}. \tag{2.33}$$

Se considerarán intervalos de tiempo de la forma $[t,t+1]$. Los usuarios arriban por paquetes de manera independiente del resto de las colas. Se define el grupo de usuarios que llegan a cada una de las colas del sistema 1, caracterizadas por Q_1 y Q_2 respectivamente, en el intervalo de tiempo $[t,t+1]$ por $X_1\left(t\right), X_2\left(t\right)$.

2.4 Ecuaciones Centrales

Proposición 2.4.1. *Supongamos*

$$f_i\left(i\right)-f_j\left(i\right)=\mu_i\left[\sum_{k=j}^{i-1}r_k+\sum_{k=j}^{i-1}\frac{f_k\left(k\right)}{1-\mu_k}\right] \tag{2.34}$$

$$f_{i+1}\left(i\right)=r_i\mu_i, \tag{2.35}$$

Demostrar que

$$f_i\left(i\right) = \mu_i\left[\sum_{k=1}^{N}r_k+\sum_{k=1,k\neq i}^{N}\frac{f_k\left(k\right)}{1-\mu_k}\right].$$

En la Ecuación (2.35) hagamos $j=i+1$, entonces se tiene $f_j=r_i\mu_i$, lo mismo para (2.34)

$$\begin{aligned} f_i\left(i\right) &= r_i\mu_i+\mu_i\left[\sum_{k=j}^{i-1}r_k+\sum_{k=j}^{i-1}\frac{f_k\left(k\right)}{1-\mu_k}\right] \\ &= \mu_i\left[\sum_{k=j}^{i}r_k+\sum_{k=j}^{i-1}\frac{f_k\left(k\right)}{1-\mu_k}\right] \end{aligned}$$

entonces, tomando sobre todo valor de $1,\ldots,N$, tanto para antes de i como para después de i, entonces

$$f_i\left(i\right) = \mu_i\left[\sum_{k=1}^{N}r_k+\sum_{k=1,k\neq i}^{N}\frac{f_k\left(k\right)}{1-\mu_k}\right].$$

Ahora, supongamos nuevamente la ecuación (2.34)

$$\begin{aligned}
f_i(i) - f_j(i) &= \mu_i\left[\sum_{k=j}^{i-1} r_k + \sum_{k=j}^{i-1}\frac{f_k(k)}{1-\mu_k}\right]\\
&\Leftrightarrow\\
f_j(j) - f_i(j) &= \mu_j\left[\sum_{k=i}^{j-1} r_k + \sum_{k=i}^{j-1}\frac{f_k(k)}{1-\mu_k}\right]\\
f_i(j) &= f_j(j) - \mu_j\left[\sum_{k=i}^{j-1} r_k + \sum_{k=i}^{j-1}\frac{f_k(k)}{1-\mu_k}\right]\\
&= \mu_j(1-\mu_j)\frac{r}{1-\mu} - \mu_j\left[\sum_{k=i}^{j-1} r_k + \sum_{k=i}^{j-1}\frac{f_k(k)}{1-\mu_k}\right]\\
&= \mu_j\left[(1-\mu_j)\frac{r}{1-\mu} - \sum_{k=i}^{j-1} r_k - \sum_{k=i}^{j-1}\frac{f_k(k)}{1-\mu_k}\right]\\
&= \mu_j\left[(1-\mu_j)\frac{r}{1-\mu} - \sum_{k=i}^{j-1} r_k - \frac{r}{1-\mu}\sum_{k=i}^{j-1}\mu_k\right]\\
&= \mu_j\left[\frac{r}{1-\mu}\left(1-\mu_j - \sum_{k=i}^{j-1}\mu_k\right) - \sum_{k=i}^{j-1} r_k\right]\\
&= \mu_j\left[\frac{r}{1-\mu}\left(1 - \sum_{k=i}^{j}\mu_k\right) - \sum_{k=i}^{j-1} r_k\right].
\end{aligned}$$

Ahora,

$$\begin{aligned}
1 - \sum_{k=i}^{j}\mu_k &= 1 - \sum_{k=1}^{N}\mu_k + \sum_{k=j+1}^{i-1}\mu_k\\
&\Leftrightarrow\\
\sum_{k=i}^{j}\mu_k &= \sum_{k=1}^{N}\mu_k - \sum_{k=j+1}^{i-1}\mu_k\\
&\Leftrightarrow\\
\sum_{k=1}^{N}\mu_k &= \sum_{k=i}^{j}\mu_k + \sum_{k=j+1}^{i-1}\mu_k
\end{aligned}$$

Por tanto

$$f_i(j) = \mu_j\left[\frac{r}{1-\mu}\sum_{k=j+1}^{i-1}\mu_k + \sum_{k=j}^{i-1} r_k\right].$$

Teorema 2.4.1 (Teorema de Continuidad). *Supóngase que* $\{X_n, n = 1,2,3,\ldots\}$ *son variables aleatorias finitas, no negativas con valores enteros tales que* $P(X_n = k) = p_k^{(n)}$*, para* $n = 1,2,3,\ldots$, $k = 0,1,2,\ldots$, *con* $\sum_{k=0}^{\infty} p_k^{(n)} = 1$*, para* $n = 1,2,3,\ldots$. *Sea* g_n *la PGF para la variable aleatoria* X_n. *Entonces existe una sucesión* $\{p_k\}$ *tal que*

$$lim_{n\to\infty} p_k^{(n)} = p_k \text{ para } 0 < s < 1.$$

En este caso, $g(s) = \sum_{k=0}^{\infty} s^k p_k$. *Además*

$$\sum_{k=0}^{\infty} p_k = 1 \text{ si y sólo si } lim_{s\uparrow 1} g(s) = 1$$

Teorema 2.4.2. *Sea N una variable aleatoria con valores enteros no negativos finita tal que $P(N=k)=p_k$, para $k=0,1,2,\ldots,$ y $\sum_{k=0}^{\infty} p_k = P(N<\infty)=1$. Sea Φ la PGF de N tal que $g(s)=\mathbb{E}\left[s^N\right]=\sum_{k=0}^{\infty} s^k p_k$ con $g(1)=1$. Si $0\leq p_1\leq 1$ y $\mathbb{E}[N]=g^{'}(1)\leq 1$, entonces no existe solución de la ecuación $g(s)=s$ en el intervalo $[0,1)$. Si $\mathbb{E}[N]=g^{'}(1)>1$, lo cual implica que $0\leq p_1<1$, entonces existe una única solución de la ecuación $g(s)=s$ en el intervalo $[0,1)$.*

Teorema 2.4.3. *Si X y Y tienen PGF G_X y G_Y respectivamente, entonces,*

$$G_X(s)=G_Y(s)$$

para toda s, si y sólo si

$$P(X=k))=P(Y=k)$$

para toda $k=0,1,\ldots,.$, es decir, si y sólo si X y Y tienen la misma distribución de probabilidad.

Teorema 2.4.4. *Para cada n fijo, sea la sucesióin de probabilidades $\{a_{0,n},a_{1,n},\ldots,\}$, tales que $a_{k,n}\geq 0$ para toda $k=0,1,2,\ldots,$ y $\sum_{k\geq 0} a_{k,n}=1$, y sea $G_n(s)$ la correspondiente función generadora, $G_n(s)=\sum_{k\geq 0} a_{k,n}s^k$. De modo que para cada valor fijo de k*

$$lim_{n\rightarrow\infty}a_{k,n}=a_k,$$

es decir converge en distribución, es necesario y suficiente que para cada valor fijo $s\in[0,)$,

$$lim_{n\rightarrow\infty}G_n(s)=G(s),$$

donde $G(s)=\sum_{k\geq 0} p_k s^k$, para cualquier la función generadora del límite de la sucesión.

Teorema 2.4.5 (Teorema de Abel)**.** *Sea $G(s)=\sum_{k\geq 0} a_k s^k$ para cualquier $\{p_0,p_1,\ldots,\}$, tales que $p_k\geq 0$ para toda $k=0,1,2,\ldots,$. Entonces $G(s)$ es continua por la derecha en $s=1$, es decir*

$$lim_{s\uparrow 1}G(s)=\sum_{k\geq 0} p_k = G(),$$

sin importar si la suma es finita o no.

Nota 2.4.1. *El radio de Convergencia para cualquier PGF es $R\geq 1$, entonces, el Teorema de Abel nos dice que aún en el peor escenario, cuando $R=1$, aún se puede confiar en que la PGF será continua en $s=1$, en contraste, no se puede asegurar que la PGF será continua en el límite inferior $-R$, puesto que la PGF es simétrica alrededor del cero: la PGF converge para todo $s\in(-R,R)$, y no lo hace para $s<-R$ o $s>R$. Además nos dice que podemos escribir $G_X(1)$ como una abreviación de $lim_{s\uparrow 1}G_X(s)$.*

Entonces si suponemos que la diferenciación término a término está permitida, entonces

$$G_X^{'}(s) = \sum_{x=1}^{\infty} x s^{x-1} p_x$$

el Teorema de Abel nos dice que

$$\begin{aligned}
\mathbb{E}(X) &= \lim_{s\uparrow 1} G_X^{'}(s): \\
\mathbb{E}[X] &= =\sum_{x=1}^{\infty} x p_x = G_X^{'}(1) \\
&= \lim_{s\uparrow 1} G_X^{'}(s),
\end{aligned}$$

dado que el Teorema de Abel se aplica a

$$G_X^{'}(s) = \sum_{x=1}^{\infty} x s^{x-1} p_x,$$

estableciendo así que $G_X^{'}(s)$ es continua en $s=1$. Sin el Teorema de Abel no se podría asegurar que el límite de $G_X^{'}(s)$ conforme $s\uparrow 1$ sea la respuesta correcta para $\mathbb{E}[X]$.

Nota 2.4.2. *La PGF converge para todo $|s| < R$, para algún R. De hecho la PGF converge absolutamente si $|s| < R$. La PGF además converge uniformemente en conjuntos de la forma $\left\{s : |s| < R^{'}\right\}$, donde $R^{'} < R$, es decir, $\forall \epsilon > 0, \exists n_0 \in \mathbb{Z}$ tal que $\forall s$, con $|s| < R^{'}$, y $\forall n \geq n_0$,*

$$|\sum_{x=0}^{n} s^x \mathbb{P}\left(X = x\right) - G_X\left(s\right)| < \epsilon.$$

De hecho, la convergencia uniforme es la que nos permite diferenciar término a término:

$$G_X\left(s\right) = \mathbb{E}\left[s^X\right] = \sum_{x=0}^{\infty} s^x \mathbb{P}\left(X = x\right),$$

y sea $s < R$.

1.

$$\begin{aligned} G_X^{'}\left(s\right) &= \frac{d}{ds}\left(\sum_{x=0}^{\infty} s^x \mathbb{P}\left(X = x\right)\right) = \sum_{x=0}^{\infty} \frac{d}{ds}\left(s^x \mathbb{P}\left(X = x\right)\right) \\ &= \sum_{x=0}^{n} x s^{x-1} \mathbb{P}\left(X = x\right). \end{aligned}$$

2.

$$\begin{aligned} \int_a^b G_X\left(s\right) ds &= \int_a^b \left(\sum_{x=0}^{\infty} s^x \mathbb{P}\left(X = x\right)\right) ds = \sum_{x=0}^{\infty} \left(\int_a^b s^x \mathbb{P}\left(X = x\right) ds\right) \\ &= \sum_{x=0}^{\infty} \frac{s^{x+1}}{x+1} \mathbb{P}\left(X = x\right), \end{aligned}$$

para $-R < a < b < R$.

Teorema 2.4.6 (Teorema de Convergencia Monótona para PGF)**.** *Sean X y X_n variables aleatorias no negativas, con valores en los enteros, finitas, tales que*

$$lim_{n\rightarrow\infty} G_{X_n}\left(s\right) = G_X\left(s\right)$$

para $0 \leq s \leq 1$, entonces

$$lim_{n\rightarrow\infty} P\left(X_n = k\right) = P\left(X = k\right),$$

para $k = 0, 1, 2, \ldots$.

El teorema anterior requiere del siguiente lema

Lema 2.4.1. *Sean $a_{n,k} \in \mathbb{Z}^+$, $n \in \mathbb{N}$ constantes no negativas con $\sum_{k\geq 0} a_{k,n} \leq 1$. Supóngase que para $0 \leq s \leq 1$, se tiene*

$$a_n\left(s\right) = \sum_{k=0}^{\infty} a_{k,n} s^k \rightarrow a\left(s\right) = \sum_{k=0}^{\infty} a_k s^k.$$

Entonces

$$a_{0,n} \rightarrow a_0.$$

2.5 Redes de Jackson

Cuando se considera la cantidad de usuarios que llegan a cada uno de los nodos desde fuera del sistema más los que provienen del resto de los nodos, se dice que la red es abierta y recibe el nombre de *Red de Jackson Abierta.*

Si denotamos por $Q_1(t), Q_2(t), \ldots, Q_K(t)$ el número de usuarios presentes en la cola $1, 2, \ldots, K$ respectivamente al tiempo t, entonces se tiene la colección de colas $\{Q_1, Q_2, \ldots, Q_K\}$, donde después de que el usuario es atendido en la cola i, se traslada a la cola j con probabilidad p_{ij}. En caso de que un usuario decida volver a ser atendido en i, este permanecerá en la misma cola con probabilidad p_{ii}. Para considerar a los usuarios que entran al sistema por primera vez por i, más aquellos que provienen de otra cola, es necesario considerar un estado adicional 0, con probabilidad de transición $p_{00} = 0$, $p_{0j} \geq 0$ y $p_{j0} \geq 0$, para $j = 1, 2, \ldots, K$, entonces en general la probabilidad de transición de una cola a otra puede representarse por $P = (p_{ij})_{i,j=0}^{K}$.

Para el caso específico en el que en cada una de las colas los tiempos entre arribos y los tiempos de servicio sean exponenciales con parámetro de intensidad λ y media μ, respectivamente, con m servidores y sin restricciones en la capacidad de almacenamiento en cada una de las colas, en Chee-Hook y Boon-Hee [107], cap. 6, se muestra que el número de usuarios en las K colas, en el caso estacionario, puede determinarse por la ecuación (2.36) que a continuación se presenta, además de que la distribución límite de la misma es (2.37).

El número de usuarios en las K colas en su estado estacionario, ver [84], se define como

$$p_{q_1 q_2 \cdots q_K} = P\left[Q_1 = q_1, Q_2 = q_2, \ldots, Q_K = q_K\right]. \tag{2.36}$$

Jackson (1957), demostró que la distribución límite $p_{q_1 q_2 \cdots q_K}$ de (2.36) es

$$p_{q_1 q_2 \cdots q_K} = P_1(q_1) P_2(q_2) \cdots P_K(q_K), \tag{2.37}$$

donde

$$p_i(r) = \begin{cases} p_i(0) \frac{(\gamma_i/\mu_i)^r}{r!}, & r = 0, 1, 2, \ldots, m, \\ p_i(0) \frac{(\gamma_i/\mu_i)^r}{m! m^{r-m}}, & r = m, m+1, \ldots. \end{cases} \tag{2.38}$$

y

$$\gamma_i = \lambda_i + \sum p_{ji}\gamma_j, \; i = 1, 2, \ldots, K. \tag{2.39}$$

La relación (2.39) es importante puesto que considera no solamente los arribos externos si no que además permite considerar intercambio de clientes entre las distintas colas que conforman el sistema.

Dados λ_i y p_{ij}, la cantidad γ_i puede determinarse a partir de la ecuación (2.39) de manera recursiva. Además $p_i(0)$ puede determinarse utilizando la condición de normalidad

$$\sum_{q_1}\sum_{q_2} \cdots \sum_{q_K} p_{q_1 q_2 \cdots q_K} = 1.$$

Sin embargo las Redes de Jackson tienen el inconveniente de que no consideran el caso en que existan tiempos de traslado entre las colas.

CAPÍTULO 3

Funciones Generadoras de Probabilidades

CAPÍTULO 4

Modelos de Flujo

CAPÍTULO 5

Modelos de Flujo: Preliminario

CAPÍTULO 6

Teorema de Down

CAPÍTULO 7

Revión de Procesos Regenerativos

CAPÍTULO 8

Procesos Estocasticos

CAPÍTULO 9

Procesos Regenerativos

CAPÍTULO 10

Introducción a los Procesos de Renovacion

CAPÍTULO 11

Thorisson

CAPÍTULO 12

Transformaciones

CAPÍTULO 13

Simulacion de Variables Aleatorias

CAPÍTULO 14

Procesos de Renovacion

CAPÍTULO 15

Revision Procesos Regenerativos

Part II

BIBLIOGRAFIA

CAPÍTULO 16

Bibliografía

Bibliography

[1] Asmussen, S. (1987). *Applied Probability and Queues*. John Wiley and Sons.

[2] Bhat, N. (2008). *An Introduction to Queueing Theory: Modelling and Analysis in Applications*. Birkhauser.

[3] Boxma, J. O. (1987). Pseudoconservation Laws in Cyclic Service Systems. *Journal of Applied Probability*, 24, 949-964.

[4] Borovkov, A. A., and Schassberger, R. (1995). Ergodicity of a Polling Network. *Stochastic Processes and their Applications*, 50(2), 253-262.

[5] van den Bos, L., and Boon, M. (2013). *Networks of Polling Systems* (report). Eindhoven University of Technology.

[6] Boon, M. A. A., van der Mei, R. D., and Winands, E. M. M. (2011). Applications of Polling Systems. February 24, 2011.

[7] Ng, C. H., and Soong, B. H. (2008). *Queueing Modelling Fundamentals with Applications in Communications Networks*. John Wiley and Sons.

[8] Chen, H. (1995). Fluid Approximations and Stability of Multiclass Queueing Networks I: Work-conserving Disciplines. *Annals of Applied Probability*.

[9] Cooper, R. B. (1970). Queues Served in Cyclic Order: Waiting Times. *The Bell System Technical Journal*, 49, 399-413.

[10] Dai, J. G. (1995). On Positive Harris Recurrence of Multiclass Queueing Networks: A Unified Approach Via Fluid Limit Models. *The Annals of Applied Probability*, 5(1), 49-77.

[11] Dai, J. G., and Meyn, S. P. (1995). Stability and Convergence of Moments for Multiclass Queueing Networks Via Fluid Limit Models. *IEEE Transactions on Automatic Control*, 40(11), 1889-1904.

[12] Dai, J. G., and Weiss, G. (1996). Stability and Instability of Fluid Models for Reentrant Lines. *Mathematics of Operations Research*, 21(1), 115-134.

[13] Daley, D. J. (1968). The Correlation Structure of the Output Process of Some Single Server Queueing Systems. *The Annals of Mathematical Statistics*, 39(3), 1007-1019.

[14] Davis, M. H. A. (1984). Piecewise Deterministic Markov Processes: A General Class of Nondiffusion Stochastic Models. *Journal of the Royal Statistical Society, Series B*, 46(3), 353-388.

[15] Down, D. (1998). On the Stability of Polling Models with Multiple Servers. *Journal of Applied Probability*, 35(3), 925-935.

[16] Disney, R. L., Farrell, R. L., and Morais, P. R. (1973). A Characterization of $M/G/1$ Queues with Renewal Departure Processes. *Management Science*, 19(11), Theory Series, 1222-1228.

[17] Eisenberg, M. (1972). Queues with Periodic Service and Changeover Time. *Operations Research*, 20(2), 440-451.

[18] Fricker, C., and Jaïbi, M. R. (1994). Monotonicity and Stability of Periodic Polling Models. *Queueing Systems*, 15(1-4), 211-238.

[19] Getoor, R. K. (1979). Transience and Recurrence of Markov Processes. In J. Azema and M. Yor (Eds.), *Seminaire de Probabilités XIV* (pp. 397-409). New York: Springer-Verlag.

[20] Gut, A. (1995). *Stopped Random Walks: Limit Theorems and Applications.* Applied Probability.

[21] Kaspi, H., and Mandelbaum, A. (1992). Regenerative Closed Queueing Networks. *Stochastics: An International Journal of Probability and Stochastic Processes*, 39(4), 239-258.

[22] Kleinrock, L. (1975). *Queueing Systems: Theory, Volume 1.* Wiley-Interscience.

[23] Konheim, A. G., Levy, H., and Srinivasan, M. M. (1994). Descendant Set: An Efficient Approach for the Analysis of Polling Systems. *IEEE Transactions on Communications*, 42(2-4), 1245-1253.

[24] Lang, S. (1973). *Calculus of Several Variables.* Addison-Wesley Publishing Company.

[25] Levy, H., and Sidi, M. (1990). Polling Systems: Applications, Modeling, and Optimization. *IEEE Transactions on Communications*, 30(10), 1750-1760.

[26] Fricker, C., and Jaïbi, M. R. (1998). Stability of Multi-server Polling Models. Institute National de Recherche en Informatique et en Automatique.

[27] Meyn, S. P., and Tweedie, R. L. (1993). *Markov Chains and Stochastic Stability.* Springer-Verlag.

[28] Meyn, S. P., and Down, D. (1994). Stability of Generalized Jackson Networks. *The Annals of Applied Probability.*

[29] van der Mei, R. D., and Borst, S. C. (1997). Analysis of Multiple-server Polling Systems by Means of the Power-Series Algorithm. *Stochastic Models*, 13(2), 339-369.

[30] Meyn, S. P. (1995). Transience of Multiclass Queueing Networks via Fluid Limit Models. *The Annals of Applied Probability*, 5(4), 946-957.

[31] Adan, I. J. B. F., van Leeuwaarden, J. S. H., and Winands, E. M. M. (2006). On the Application of Rouche's Theorem in Queueing Theory. *Operations Research Letters*, 34(3), 355-360.

[32] Roubos, A. (2007). *Polling Systems and their Applications.* Vrije Universiteit Amsterdam.

[33] Saavedra, B. P. (2011). Informe Técnico del Microsimulador. Departamento de Matemáticas.

[34] Vishnevskii, V. M., and Semenova, O. V. (2006). Mathematical Methods to Study the Polling Systems. *Automation and Remote Control*, 67(2), 173-220.

[35] Serfozo, R. (2009). *Basics of Applied Stochastic Processes.* Springer-Verlag.

[36] Sharpe, M. (1998). *General Theory of Markov Processes.* Academic.

[37] Sigman, K. (1990). The Stability of Open Queueing Networks. *Stochastic Processes and their Applications*, 35(1), 11-15.

[38] Sidi, M., and Levy, H. (1990). Customer Routing in Polling Systems. In P. King, I. Mitrani, and R. Pooley (Eds.), *Proceedings Performance '90.* North-Holland, Amsterdam.

[39] Sidi, M., and Levy, H. (1991). Polling Systems with Simultaneous Arrivals. *IEEE Transactions on Communications*, 39(6), 823-827.

[40] Sigman, K., Thorisson, H., and Wolff, R. W. (1994). A Note on the Existence of Regeneration Times. *Journal of Applied Probability*, 31, 1116-1122.

[41] Stidham, S. Jr. (1972). Regenerative Processes in the Theory of Queues, with Applications to the Alternating Priority Queue. *Advances in Applied Probability*, 4(3), 542-577.

[42] Takagi, H. (1986). *Analysis of Polling Systems.* Cambridge: MIT Press.

[43] Takagi, H., and Kleinrock. (1986). Analysis of Polling Systems. Cambridge: MIT Press.

[44] Takagi, H. (1988). Queueing Analysis of Polling Models. *ACM Computing Surveys*, 20(1), 5-28.

[45] Thorisson, H. (2000). *Coupling, Stationarity, and Regeneration.* Probability and its Applications. Springer-Verlag.

[46] Winands, E. M. M., Adan, I. J. B. F., and van Houtum, G. J. (2006). Mean Value Analysis for Polling Systems. *Queueing Systems*, 54, 35-44.

[47] James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An Introduction to Statistical Learning: with Applications in R.* Springer.

[48] Hosmer, D. W., Lemeshow, S., and Sturdivant, R. X. (2013). *Applied Logistic Regression* (3rd ed.). Wiley.

[49] Bishop, C. M. (2006). *Pattern Recognition and Machine Learning.* Springer.

[50] Harrell, F. E. (2015). *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis.* Springer.

[51] R Documentation and Tutorials: `https://cran.r-project.org/manuals.html`

[52] Tutorials on R-bloggers: `https://www.r-bloggers.com/`

[53] Coursera: *Machine Learning* by Andrew Ng.

[54] edX: *Data Science and Machine Learning Essentials* by Microsoft.

[55] Ross, S. M. (2014). *Introduction to Probability and Statistics for Engineers and Scientists.* Academic Press.

[56] DeGroot, M. H., and Schervish, M. J. (2012). *Probability and Statistics* (4th ed.). Pearson.

[57] Hogg, R. V., McKean, J., and Craig, A. T. (2019). *Introduction to Mathematical Statistics* (8th ed.). Pearson.

[58] Kleinbaum, D. G., and Klein, M. (2010). *Logistic Regression: A Self-Learning Text* (3rd ed.). Springer.

[59] Wasserman, L. (2004). *All of Statistics: A Concise Course in Statistical Inference.* Springer.

[60] Probability and Statistics Tutorials on Khan Academy: `https://www.khanacademy.org/math/statistics-probability`

[61] Online Statistics Education: `http://onlinestatbook.com/`

[62] Peng, C. Y. J., Lee, K. L., and Ingersoll, G. M. (2002). *An Introduction to Logistic Regression Analysis and Reporting.* The Journal of Educational Research.

[63] Agresti, A. (2007). *An Introduction to Categorical Data Analysis* (2nd ed.). Wiley.

[64] Han, J., Pei, J., and Kamber, M. (2011). *Data Mining: Concepts and Techniques.* Morgan Kaufmann.

[65] Data Cleaning and Preprocessing on Towards Data Science: `https://towardsdatascience.com/data-cleaning-and-preprocessing`

[66] Molinaro, A. M., Simon, R., and Pfeiffer, R. M. (2005). *Prediction Error Estimation: A Comparison of Resampling Methods.* Bioinformatics.

[67] Evaluating Machine Learning Models on Towards Data Science: `https://towardsdatascience.com/evaluating-machine-learning-models`

[68] Practical Guide to Logistic Regression in R on Towards Data Science: `https://towardsdatascience.com/practical-guide-to-logistic-regression-in-r`

[69] Coursera: *Statistics with R* by Duke University.

[70] edX: *Data Science: Probability* by Harvard University.

[71] Coursera: *Logistic Regression* by Stanford University.

[72] edX: *Data Science: Inference and Modeling* by Harvard University.

[73] Coursera: *Data Science: Wrangling and Cleaning* by Johns Hopkins University.

[74] edX: *Data Science: R Basics* by Harvard University.

[75] Coursera: *Regression Models* by Johns Hopkins University.

[76] edX: *Data Science: Statistical Inference* by Harvard University.

[77] An Introduction to Survival Analysis on Towards Data Science: `https://towardsdatascience.com/an-introduction-to-survival-analysis`

[78] Multinomial Logistic Regression on DataCamp: `https://www.datacamp.com/community/tutorials/multinomial-logistic-regression-R`

[79] Coursera: *Survival Analysis* by Johns Hopkins University.

[80] edX: *Data Science: Statistical Inference and Modeling for High-throughput Experiments* by Harvard University.

[81] Sigman, K., and Wolff, R. W. (1993). A Review of Regenerative Processes. *SIAM Review*, 38(2), 269-288.

[82] I.J.B.F. Adan, J.S.H. van Leeuwaarden, and E.M.M. Winands. On the application of Rouches theorem in queueing theory. Operations Research Letters, 34(3):355-360, 2006.

[83] Asmussen Soren, Applied Probability and Queues, John Wiley and Sons, 1987.

[84] Bhat Narayan, An Introduction to Queueing Theory Modelling and Analysis in Applications, Birkhauser, 2008.

[85] Boxma J. O., Analysis and Optimization of Polling Systems, Queueing, Performance and Control in ATM, pp. 173-183, 1991.

[86] Boxma J. O., Pseudoconservation Laws in Cyclic Service Systems, Journal of Applied Probability, vol. 24, 1987, pp. 949-964.

[87] Boon M.A.A., Van der Mei R.D. and Winands E.M.M., Applications of Polling Systems, 2011.

[88] Borovkov. A. A. and Schassberger R., Ergodicity of a Polling Network, Stochastic Processes and their Applications, vol. 50, no. 2, 1995, pp. 253-262.

[89] Laurent van den Bos and Marko Boon, Networks of Polling Systems (report), Eindhoven University of Technology, 2013.

[90] Chee-Hook Ng. y Boon-Hee Soong, Queueing Modelling Fundemantals with Applications in Communications Networks, John-Wiley and Sons, Ltd, 2008.

[91] Chen H., Fluid Approximations and Stability of Multiclass Queueing Networks I: Work-Conserving Disciplines, Annals Applied Probabability, vol. 5, no. 3, 1995, pp. 637-665.

[92] R.B. Cooper, G. Murray, Queues served in cyclic order (The Bell System Technical Journal, 48 (1969) 675-689).

[93] R.B. Cooper, Queues served in cyclic order: waiting times (The Bell System Technical Journal, 49 (1970) 399-413).

[94] Dai Jean G., On positive Harris Recurrence of Multiclass Queueing Networks: A Unified Approach Via Fluid Limit Models, The Annals of Applied Probability, vol. 5, No. 1, 1995, pp. 49-77.

[95] Dai Jim G. and Meyn Sean P., Stability and Convergence of Moments for Multiclass Queueing Networks via Fluid Limit Models, IEEE transactions on Automatic Control, vol. 40, No. 11, 1995, pp. 1889-1904.

[96] Dai Jim G. and Weiss G., Stability and Inestability of Fluid Models for Reentrant Lines, Mathematics of Operation Research, vol. 21, no. 1, 1996, pp. 115-134.

[97] D.J. Daley, The correlation structure of the output process of some single server queueing systems, The Annals of Mathematical Statistics, Vol. 39. No. 3, pp. 1007-1019, 1968.

[98] Davis, M. H. A., Piecewise Deterministic Markov Processes: A General Class of Nondifussion Stochastic Models. Journal of Royal Statistics Society Serie B, vol. 46, no. 3, 1984, pp. 353-388.

[99] Down D., On the Stability of Polling Models with Multiple Servers, Journal of Applied Probability, vol. 35, no. 3, 1998, pp. 925-935.

[100] D.J. Daley, The correlation structure of the output process of some single server queueing systems, The Annals of Mathematical Statistics, Vol. 39. No. 3, pp. 1007-1019, 1968.

[101] Ralpjh L. Disney, Robert L. Farrell and Paulo Renato Morais, A Characterization of $M/G/1$ Queues with Renewal Departure Processes, Manage of Science, Vol. 19, No. 11, Theory Series, pp. 1222-1228, 1973.

[102] M. Eisenberg, Queues with periodic service and changeover time (Operations Research, 20 (2)(1972) 440-451).

[103] Fricker Christine and Jaïbi M.R., Monotonicity and Stability of Periodic Polling Models, Queueing Systems, vol. 15, no. 1-4, 1994, pp. 211-238.

[104] Fricker Christine and Jaïbi M.R., Stability of Multi-Server Polling Models, Institute National de Recherche en Informatique et en Automatique, Issue 3347, 1998.

[105] Getoor R. K., Transience and recurrence of Markov processes, Siminaire de Probabilitis XIV, J. Azema and M Yor, Eds. New York Springer-Verlag, 1979.

[106] Gut A., Stopped Random Walks: Limit Theorems and Applications, Applied Probability, 1995.

[107] Chee-Hook Ng. y Boon-Hee Soong, Queueing Modelling Fundemantals with Applications in Communications Networks, John-Wiley and Sons, Ltd, 2008.

[108] Kaspi H. and Mandelbaum A., Regenerative Closed Queueing Networks, Stochastics: An International Journal of Probability and Stochastic Processes, vol. 39, no. 4, 1992, pp. 239-258.

[109] A.G. Konheim, H. Levy, M.M. Srinivasan, Descendant set: an efficient approach for the analysis of polling systems (IEEE Transactions on Communications, 42 (2/3/4)(1994) 1245-1253).

[110] Leonard Kleinrock, Theory, Volume 1, Queueing Systems Wiley-Interscience, 1975,

[111] A. G. Konheim, h. Levy and M. M. Srinivasan. Descendant set: an efficient approach for the analysis of polling systems. IEEE Trabsanctions on Communications, 42(2-4): pp. 1245-1253, 1994.

[112] Serge Lang, Calculus of Several Variables, Addison-Wesley Publishing Company, 1973.

[113] Levy Hanoch y Sidi Moshe , Polling Systems: Applications, Modeling, and Optimization, IEEE Transactions on Communications, vol. 30, no. 10, 1990, pp. 1750-1760.

[114] van der Mei, R.D. and Borst, S. C., Analysis of Multiple-Server Polling Systems by Means of the Power-Series Algorithm, Stochastic Models, vol. 13, no. 2, 1997, pp. 339-369.

[115] Meyn Sean P., Transience of Multiclass Queueing Networks via Fluid Limit Models, The Annals of Applied Probability, vol. 5, no. 4, 1995, pp.946-957.

[116] Meyn S. P. and Tweedie R. L., Stability of Markovian Processes II:Continuous Time Processes and Sample Chains, Advanced Applied Pobability, vol. 25, 1993, pp. 487-517.

[117] Meyn S. P. and Tweedie R. L., Markov Chains and Stochastic Stability, 1993.

[118] Meyn, S.P. and Down, D., Stability of Generalized Jackson Networks, The Annals of Applied Probability, 1994.

[119] Roubos Alex, Polling Systems and their Applications, Vrije Universiteit Amsterdam, 2007.

[120] Saavedra B. P., Informe Técnico del Microsimulador, Departamento de Matemátias, 2011.

[121] Vishnevskii V.M. and Semenova O.V., Mathematical Methods to Study the Polling Systems, Automation and Remote Control, vol. 67, no. 2, 2006, pp. 173-220.

[122] Richard Serfozo, Basics of Applied Stochastic Processes, Springer-Verlag, 2009.

[123] Sharpe Michael , General Theory of Markov Processes. Boston, M.A. Academic, 1998.

[124] Sigman Karl, The Stability of Open Queueing Networks, Stochastic Processes and their Applications, vol. 35, no. 1, 1990, pp. 11-15.

[125] Sidi M. and Levy H., Customer Routing in Polling Systems, In: P. King, I. Mitrani, and R. Pooley (Eds.), Proceedings Performance '90, North-Holland, Amsterdam, 1990.

[126] Sidi M. and Levy H., Polling Systems with Simultaneus Arrivals. IEEE Transactions on Communications, vol. 39, no. 6, 1991, pp. 823-827.

[127] Karl Sigman, Hermann Thorisson and Ronald W. Wolff, A Note on the Existence of Regeneration Times, Journal of Applied Probability, vol. 31, pp. 1116-1122, 1994.

[128] Shaler Stidham, Jr., Regenerative Processes in the theory ow queues, with applications to the alternating priority queue, Advances in Applied Probability, Vol. 4, no. 3, 1972,pp. 542-577.

[129] Takagi H., Analysis of Polling Systems, Cambdrige: MIT Press, 1986

[130] Takagi H. and Kleinrock, Analysis of Polling Systems, Cambdrige: MIT Press, 1986

[131] Takagi Hideaki, Queueing Analysis of Polling Models, ACM computing Surveys, vol. 20, no. 1, 1988, pp. 5-28.

[132] Hermann Thorisson, Coupling, Stationarity, and Regeneration, Probability and its Applications, Springer-Verlag, 2000.

[133] E.M.M. Winands, I.J.B.F. Adan adn G.J. van Houtum, Mean value analysis for polling systems, Queueing Systems (2006), 54:35-44,

[134] Konheim, A.G. and Meister, B., Waiting Lines and Times in a System with Polling, J. ACM, 1974, vol. 21, n0. 3, pp. 470-490.

[135] Levy, H. and Kleinrock, L., Polling Systems with Zero Switch-over Periods: A General Method for Analysis the Expected DElay, Performance Evaluat., 1991, vol. 13, no. 2, pp. 97-107.

[136] Yechiali, U., Analysis and Control of Polling systems, in Performance Evaluation of Computer and Communications Systems, Donatiello, L. and Nelson R., Eds. Berlin: Springer Verlag, 1993, pp. 630-650.