Aplicación de Cadenas de Markov

Carlos Ernesto Martínez Rodríguez

Universidad Autónoma de la Ciudad de México, Casa Libertad

Índice

Notación de Kendall

A partir de este momento se harán las siguientes consideraciones: 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. 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. 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.

Notación de Kendall

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

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

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

Notación de Kendall

Además tenemos las siguientes igualdades

$$N(t) = N_q(t) + N_s(s) \tag{1}$$

donde

- $N(t)$ es el número de clientes en el sistema al tiempo t.
- $N_q(t)$ es el número de clientes 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

$$N = N_q + N_s. \tag{2}$$

Los valores medios de las cantidades anteriores se escriben como $L = E(N)$, $L_q = E\left(N_q\right)$ y $L_s = E(N_s)$, entonces de la ecuación 1 se obtiene

Notación de Kendall

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

$$\rho = \frac{E(s)}{E(\tau)} = \frac{\beta}{\delta}. \tag{4}$$

La utilización por servidor es

$$u = \frac{\rho}{c} = \frac{\beta}{c\delta}. \tag{5}$$

donde c es el número de servidores.

Más sobre la notación *Kendall-Lee*

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

$$A/S/c/K/F/d \tag{6}$$

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.

Más sobre la notación *Kendall-Lee*

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.

Procesos de Nacimiento y Muerte

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.

Procesos de Nacimiento y Muerte

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.

Procesos de Nacimiento y Muerte

Proposición

La recurrencia de $\{X_t\}$, o equivalentemente de $\{Y_n\}$ es equivalente a

$$\sum_{n=1}^{\infty} \frac{\delta_1 \cdots \delta_n}{\beta_1 \cdots \beta_n} = \sum_{n=1}^{\infty} \frac{q_1 \cdots q_n}{p_1 \cdots p_n} = \infty \tag{7}$$

Lema

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

$$v_n = \frac{\beta_0 \cdots \beta_{n-1}}{\delta_1 \cdots \delta_n} v_0 \tag{8}$$

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

Procesos de Nacimiento y Muerte

Corolario

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

$$\mu_n = \frac{p_1 \cdots p_{n-1}}{q_1 \cdots q_n} \mu_0 \tag{9}$$

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

Procesos de Nacimiento y Muerte

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

$\{X_t\}$ es ergódica si y sólo si la ecuación (36) se cumple y además $S < \infty$, en cuyo caso la distribución ergódica, π, está dada por

$$\pi_0 = \frac{1}{S}, \pi_n = \frac{1}{S}\frac{\beta_0\cdots\beta_{n-1}}{\delta_1\cdots\delta_n} \tag{10}$$

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

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 36) 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 39 se tiene que

Cola $M/M/1$

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

Lo cual nos lleva a la siguiente

Proposición

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

Cola $M/M/1$

Entonces por el corolario 3.1

Proposición

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.

1. $\mathbb{P}[X_t = 0] = \pi_0 = 1 - \rho$, es decir, la probabilidad de que el sistema se encuentre ocupado.
2. De las propiedades de la distribución Geométrica se desprende que
   1. $\mathbb{E}[X_t] = \frac{\rho}{1-\rho}$,
   2. $Var[X_t] = \frac{\rho}{(1-\rho)^2}$.

Cola $M/M/1$

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

$$L = \frac{\rho}{1-\rho}$$

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 $W_s = \mathbb{E}[s]$ y $\mathbb{E}[\tau] = \frac{1}{\delta}$. Por la fórmula de Little $L = \lambda W$

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

Cola $M/M/1$

luego entonces

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

Entonces

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

Finalmente

Cola $M/M/1$

Proposición

1. $W(t) = 1 - e^{-\frac{t}{W}}$.
2. $W_q(t) = 1 - \rho \exp^{-\frac{t}{W}}$.

donde $W = \mathbb{E}(w)$.

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 $n = 0, 1, 2, \ldots$, entonces por la ecuación 39 se tiene que

$$\pi_0 = e^{\rho}$$
$$\pi_n = e^{-\rho}\frac{\rho^n}{n!}$$

Cola $M/M/\infty$

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

$$L = \mathbb{E}[N] = \rho$$
$$Var[N] = \rho$$

Cola $M/M/\infty$

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, $W = \mathbb{E}[s] = \frac{1}{\delta}$. Resumiendo, tenemos la siguiente proposición:

Proposición

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

Cola $M/M/m$

Este sistema considera m servidores idénticos, con tiempos entre arribos y de servicio exponenciales con medias $\mathbb{E}[\tau] = \frac{1}{\beta}$ y $\mathbb{E}[s] = \frac{1}{\delta}$. 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

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

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

Cola $M/M/m$

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

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

Cola $M/M/m$

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

luego, para este caso se tiene que

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

Cola $M/M/m$

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

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

Cola $M/M/m$

es decir

$$L_q = \frac{u\pi_0\rho^m}{m!\,(1-u)^2} \tag{11}$$

luego

$$W_q = \frac{L_q}{\beta} \tag{12}$$

$$W = W_q + \frac{1}{\delta} \tag{13}$$

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

Cola $M/M/m$

$$L_q = \frac{C(m,\rho)\,u}{1-u}$$

$$W_q = \frac{C(m,\rho)\,\mathbb{E}[s]}{m(1-u)}$$

Proposición

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

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

Cola $M/M/m$

Proposición

Para $t \geq 0$

a) $W_q(t) = 1 - C(m,\rho)\, e^{-c\delta t(1-u)}$

b)

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

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 $t_1 < t_2 < \ldots$ los tiempos sucesivos en los que los clientes completan su servicio y salen del sistema.

La sucesión $\{X_n\}$ definida por $X_n = N(t)$ 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

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

Cola M/G/1

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:

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

por el teorema de la probabilidad total se tiene que

$$a_j = \mathbb{P}\left\{U = j\right\} = \int_0^\infty \mathbb{P}\left\{U = j | s = t\right\} dG\left(t\right) = \int_0^\infty e^{-\lambda t}\frac{(\lambda t)^j}{j!} dG\left(t\right) \qquad (14)$$

donde G es la distribución de los tiempos de servicio.

Cola M/G/1

Las probabilidades de transición de la cadena están dadas por

$$p_{0j} = \mathbb{P}\left\{U_{n+1} = j\right\} = a_j, \text{ para } j = 0, 1, \ldots \tag{15}$$

y para $i \geq 1$

$$p_{ij} = \begin{cases} \mathbb{P}\left\{U_{n+1} = j - i + 1\right\} = a_{j-i+1} & , \text{para } j \geq i - 1 \\ 0 & j < i - 1 \end{cases} \tag{16}$$

Cola M/G/1

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

Teorema

La cadena encajada $\{X_n\}$ es

a) *Recurrente positiva si $\rho < 1$,*
b) *Transitoria si $\rho > 1$,*
c) *Recurrente nula si $\rho = 1$.*

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

$$\pi_j = \pi_0 a_j + \sum_{i=1}^{j+1} \pi_i a_{j-i+1}, \text{ para } j = 0, 1, \ldots \tag{17}$$

$$\pi_0 = 1 - \rho \tag{18}$$

Cola M/G/1

Además se tiene que

$$L = \pi^{'}(1) = \rho + \frac{A^{''}(1)}{2(1-\rho)} \tag{19}$$

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 $L = \rho + \frac{\beta^2\mathbb{E}[s^2]}{2(1-\rho)}$.
De las fórmulas de Little, se tiene que $W = E(w) = \frac{L}{\beta}$, también el tiempo de espera en la cola

$$W_q = \mathbb{E}(q) = \mathbb{E}(w) - \mathbb{E}(s) = \frac{L}{\beta} - \frac{1}{\delta}, \tag{20}$$

además el número promedio de clientes en la cola es

$$L_q = \mathbb{E}\left(N_q\right) = \beta W_q = L - \rho \tag{21}$$

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

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

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

Cola M/M/m/m

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

$$\pi_n = \begin{cases} \pi_0 \frac{\rho^n}{n!} & n = 0, 1, 2, \ldots, m \\ 0 & n \geq m \end{cases} \tag{22}$$

y ademas

$$\pi_0 = \left(\sum_{n=0}^{m} \frac{\rho^n}{n!} \right)^{-1} \tag{23}$$

A la ecuación 22 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*.

Cola M/M/m/m

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 tiempode 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

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

Cola M/M/m/m

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

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

entonces, se tiene que

$$L = \rho\left(1 - B\left(m,\rho\right)\right) = \mathbb{E}\left[s\right]\left(1 - B\left(m,\rho\right)\right). \tag{24}$$

Además

$$\delta_q = \delta\left(1 - B\left(m,\rho\right)\right) \tag{25}$$

representa la tasa promedio efectiva de arribos al sistema.

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 siemas secuenciales.

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

$$\lambda P_{0,0} = \mu_2 P_{0,1} \tag{26}$$

$$(\lambda + \mu_1) P_{n,0} = \mu_2 P_{n,1} + \lambda P_{n-1,0} \tag{27}$$

$$(\lambda + \mu_2) P_{0,m} = \mu_2 P_{0,m+1} + \mu_1 P_{1,m-1} \tag{28}$$

$$(\lambda + \mu_1 + \mu_2) P_{n,m} = \mu_2 P_{n,m+1} + \mu_1 P_{n+1,m-1} + \lambda P_{n-1,m} \tag{29}$$

Sistemas Abiertos

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

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

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

$$P_{n,m} = \rho_1^n \left(1 - \rho_1\right) \rho_2^m \left(1 - \rho_2\right) \tag{30}$$

Sistemas Abiertos

Verifiquemos que $P_{n,m}$ satisface las ecuaciones de balance (26) Antes de eso, enunciemos unas igualdades que nos serán de utilidad:

$$\mu_i \rho_i = \lambda \text{ para } i = 1, 2.$$

Sistemas Abiertos

$$\begin{aligned}
\lambda P_{0,0} &= \lambda\left(1-\rho_1\right)\left(1-\rho_2\right) \\
\text{y } \mu_2 P_{0,1} &= \mu_2\left(1-\rho_1\right)\rho_2\left(1-\rho_2\right) \\
\Rightarrow \lambda P_{0,0} &= \mu_2 P_{0,1} \\
\left(\lambda+\mu_2\right)P_{0,m} &= \left(\lambda+\mu_2\right)\left(1-\rho_1\right)\rho_2^m\left(1-\rho_2\right) \\
\mu_2 P_{0,m+1} &= \lambda\left(1-\rho_1\right)\rho_2^m\left(1-\rho_2\right) \\
&= \mu_2\left(1-\rho_1\right)\rho_2^m\left(1-\rho_2\right) \\
\mu_1 P_{1,m-1} &= \frac{\lambda}{\rho_2}\left(1-\rho_1\right)\rho_2^m\left(1-\rho_2\right) \\
\Rightarrow \left(\lambda+\mu_2\right)P_{0,m} &= \mu_2 P_{0,m+1}+\mu_1 P_{1,m-1} \\
\left(\lambda+\mu_1+\mu_2\right)P_{n,m} &= \left(\lambda+\mu_1+\mu_2\right)\rho^n\left(1-\rho_1\right)\rho_2^m\left(1-\rho_2\right) \\
\mu_2 P_{n,m+1} &= \mu_2\rho_2\rho_1^n\left(1-\rho_1\right)\rho_2^m\left(1-\rho_2\right)
\end{aligned}$$

Procesos de Salto

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

$$X(t) = \begin{cases} x_0 & 0 \leq t < \tau_1 \\ x_1 & \tau_1 \leq t < \tau_2 \\ x_2 & \tau_2 \leq t < \tau_3 \\ \vdots & \end{cases} \tag{31}$$

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

$$lim_{n\rightarrow\infty}\tau_n = \begin{cases} < \infty & X_t \text{ explota} \\ = \infty & X_t \text{ no explota} \end{cases} \tag{32}$$

Procesos de Salto

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

Proposición

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

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

para $s, t \geq 0$, *equivalentemente*

$$\frac{1 - F_x(t+s)}{1 - F_x(s)} = 1 - F_x(t). \tag{33}$$

Procesos de Salto

Observaciones

Una distribución F_x satisface la ecuación (33) 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:

$$\Lambda = \begin{pmatrix} -\beta_0 & \beta_0 & 0 & 0 & \ldots \\ \delta_1 & -\beta_1-\delta_1 & \beta_1 & 0 & \ldots \\ 0 & \delta_2 & -\beta_2-\delta_2 & \beta_2 & \ldots \\ \vdots & & & \ddots & \end{pmatrix} \tag{34}$$

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

Procesos de Salto

La matriz de transición es

$$Q = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots \\ q_1 & 0 & p_1 & 0 & \dots \\ 0 & q_2 & 0 & p_2 & \dots \\ \vdots & & & \ddots & \end{pmatrix} \tag{35}$$

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

Proposición

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

$$\sum_{n=1}^{\infty} \frac{\delta_1 \cdots \delta_n}{\beta_1 \cdots \beta_n} = \sum_{n=1}^{\infty} \frac{q_1 \cdots q_n}{p_1 \cdots p_n} = \infty \tag{36}$$

Lema

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

$$\nu_n = \frac{\beta_0 \cdots \beta_{n-1}}{\delta_1 \cdots \delta_n} \nu_0 \tag{37}$$

Corolario

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

$$\mu_n = \frac{p_1 \cdots p_{n-1}}{q_1 \cdots q_n}\mu_0 \tag{38}$$

para $n = 1, 2, \ldots$

Definición

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

Procesos de Salto

Definición

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

Teorema

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

$\{X_t\}_{t\geq 0}$ es ergódica si y sólo si (36) se cumple y $S < \infty$, en cuyo caso la distribución estacionaria π está dada por

$$\pi_0 = \frac{1}{S}, \pi_n = \frac{1}{S}\frac{\beta_0 \cdots \beta_{n-1}}{\delta_1 \cdots \delta_n}, n = 1, 2, \ldots \tag{39}$$

Proposición

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

Proposición

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, $\pi_n = (1-\rho)\rho^n$, para $n = 0, 1, 2, \ldots$.

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

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

En este caso $\beta_n = \beta$ y $\delta_n = m(n)\delta$, donde $m(n) = n$, $1 \leq n \leq m$. La intensidad de tráfico es $\rho = \frac{\beta}{m\delta}$, se tiene entonces que $\beta_n/\delta_n = \rho$ para $n \geq m$. Así, para el caso $m = 1$,

Cadena de Markov con dos estados

Supongamos que se tiene la siguiente cadena:

$$\begin{pmatrix} 1-q & q \\ p & 1-p \end{pmatrix} \tag{40}$$

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 $(1-p-q)^n \to 0$ cuando $n \to \infty$. Por lo tanto

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

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

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

Estacionareidad

Teorema

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 $n \in Z$.*

Estacionareidad

Teorema

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,

$$v_j = \mathbb{E}_i \sum_{n=0}^{\tau(i)-1} \mathbb{1}\left(X_n = i\right) = \sum_{n=0}^{\infty} \mathbb{P}_i\left[X_n = j, \tau(i) > n\right] \tag{41}$$

Estacionareidad

Teorema

Si la cadena es irreducible y recurrente, entonces una medida estacionaria v existe, 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* $c \in (0, \infty)$.

Estacionareidad

Corolario

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

$$\pi_j = \frac{1}{\mathbb{E}_i \tau_i} \mathbb{E}_i \sum_{n=0}^{\tau(i)-1} \mathbb{1}\left(X_n = j\right) = \frac{1}{\mathbb{E}_j \tau\left(j\right)}. \tag{42}$$

Corolario

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

Estacionareidad y Teoría Ergódica

Lema

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

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

Utilizando el teorema (2.2) y el corolario refCor.3.5, se demuestra el siguiente resultado importante.

Estacionareidad y Teoría Ergódica

Teorema

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

Definición

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

Estacionareidad

Proposición

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

$$\frac{\sum_{n=0}^{m} p_{ij}^{n}}{\sum_{n=0}^{m} p_{lk}^{n}} \rightarrow \frac{v_j}{v_k}, \; m \rightarrow \infty \tag{43}$$

Lema

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

Estacionareidad y Teoría Ergódica

Lema

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,* $lim_{m\rightarrow\infty} \frac{\mathbb{E}_j N_i^m}{\mathbb{E}_k N_i^m} = 1$ *para todo* $j, k \in E$.

Criterios de Ergodicidad

Definición

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

Teorema

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

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 $\sum \pi_j \lambda(j) < \infty$.

Criterios de Ergodicidad

Proposición

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

Teorema

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

Corolario

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

Criterios de Ergodicidad

Corolario

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

Carlos Ernesto Martínez Rodríguez
Academia de Matemáticas
Plantel Casa Libertad
Universidad Autónoma de la Ciudad de México
http://carlosmartinez.expresauacm.org/
carlos.martinez@expresauacm.org

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.