PNV 3421 – Processos Estocásticos

Notas de Aula de Teoria de Filas 19/11/2020

Roteiro da Aula

1 – Formula de Little.

2 – Processos de Fila com Chegada Poisson – a Visão de quem Chega.

3 – Para uma Fila M/G/1 $p_j = \pi_j$.

4 – Fila M/M/1 – Distribuição do tempo de permanência no sistema e distribuição dos intervalos entre saídas consecutivas.

5. Estudo da Fila $M/E_k/1$ como cadeia de Markov em tempo contínuo.

6. Comentários sobre a 3ª Série de Problemas

1 – Formula de Little

De início, é importante destacar mais uma vez que o foco desse estudo de processos de fila é o regime estacionário. A ênfase tem sido dada ao estudo do número de clientes no sistema; assim, no caso das cadeias de Markov em tempo

discreto, o objetivo primeiro era determinar a distribuição estacionária de probabilidades π_j, $j = 0, 1, 2, \ldots$, e agora, no estudo de cadeias de Markov em tempo contínuo é obter as probabilidades estacionárias p_j. Na introdução à Teoria de Filas foi mencionado que outra medida importante de avaliação de um processo de fila é o tempo de espera em fila ou o tempo de permanência do cliente no sistema.
Para alguns processos de fila, é possível deduzir a distribuição do tempo de permanência no sistema (veja item 4 desta aula) ou do tempo de espera (resolva a Questão 10 da 3ª Série de Problemas). Há, porém, mesmo que não se conheçam tais distribuições, uma relação importante para calcular o tempo médio de permanência no sistema (e, por consequência, o tempo médio de espera) a partir do número médio de clientes no sistema – é a fórmula de Little.

Sejam:

$L = \sum_{j=0}^{\infty} j p_j$ o número médio de clientes no sistema, média da distribuição das probabilidades estacionárias p_j;

$W = \lim_{N \to \infty} \left(\sum_{n=1}^{N} W_n \right) / N$ o tempo médio de permanência de um cliente no sistema; e

λ a taxa de chegada de clientes ao sistema.

A Formula de Little expressa a relação entre estes 3 elementos:

$$L = \lambda W \quad (1)$$

Little foi quem primeiro demonstrou a validade da relação (1) para um modelo específico de fila. Posteriormente, foi mostrado que a relação tem uma aplicação bem abrangente.

Seja $N(t)$ o número de clientes no sistema para um dado processo de fila. Se o processo estocástico $\{N(t), t \geq 0\}$ for um processo regenerativo, vale para ele a formula de Little.

<u>def</u> Um processo estocástico $\{N(t), t \geq 0\}$ é um processo regenerativo se existem, com probabilidade igual a 1, instantes $T_1, T_2, T_3, \ldots$, a partir dos quais o processo é uma réplica probabilística do processo começando no instante zero. Por exemplo, se o instante zero é o instante da primeira chegada ao sistema, dando início ao primeiro período ocupado, T_1 marca o fim do primeiro período vazio e dá início ao segundo

período ocupado, T_2 indica o fim do segundo período vazio e início do terceiro período ocupado e, assim, sucessivamente. Para filas com índice de congestionamento menor que 1, o processo $\{N(t), t \geq 0\}$ é um processo regenerativo.

Uma demonstração informal da fórmula de Little pode ser feita a partir da Figura 1 utilizada na aula introdutória de Teoria de Filas. Nesta figura $A(t)$ conta o número de chegadas até o instante t, $D(t)$ conta o número de saídas até o instante t, $N(t)$ é o número de clientes no sistema no instante t, V_n é o tempo de atendimento do cliente, W_{q_n} é o tempo de espera em fila do cliente n e a soma destas duas parcelas fornece o tempo de permanência do cliente n no sistema. Isto é, $W_n = W_{q_n} + V_n$. A área compreendida entre as curvas $A(t)$ e $D(t)$ representa, de um lado, a integral de $N(t) = A(t) - D(t)$ e, de outro, a soma dos tempos de permanência dos clientes no sistema. Seja T o instante final do primeiro período ocupado e $N(T)$ o número de clientes atendidos no primeiro período ocupado. Então:

$$\int_0^T N(t)\,dt = \sum_{n=1}^{N(T)} W_n \qquad (2)$$

Dividindo (2) por T, obtém-se:

$$\frac{1}{T}\int_0^T N(t)\,dt = \frac{1}{T}\sum_{n=1}^{N(T)} W_n =$$

$$= \frac{N(T)}{T}\sum_{n=1}^{N(T)} \frac{W_n}{N(T)} \qquad (3)$$

Ou

$$\bar{N}_{PC1} = \frac{N(T)}{T}\,\bar{W}_{PC1} \qquad (4)$$

em que $\bar{N}_{PC1}$ é o número médio de clientes no sistema até o fim do primeiro período ocupado e $\bar{W}_{PC1}$ é o tempo médio de permanência no sistema dos clientes atendidos até o fim do primeiro período ocupado

Considerando que, quando o índice de congestionamento for menor que 1, haverá uma infinidade de períodos ocupados, entremeados por períodos vazios, e que $\lim_{t\to\infty} N(t)/t = \lambda$, de 4 resulta:

$$L = \lambda W \qquad (1)$$

$N(t) = A(t) - D(t)$

$\hookrightarrow$ número de navios no terminal no instante t

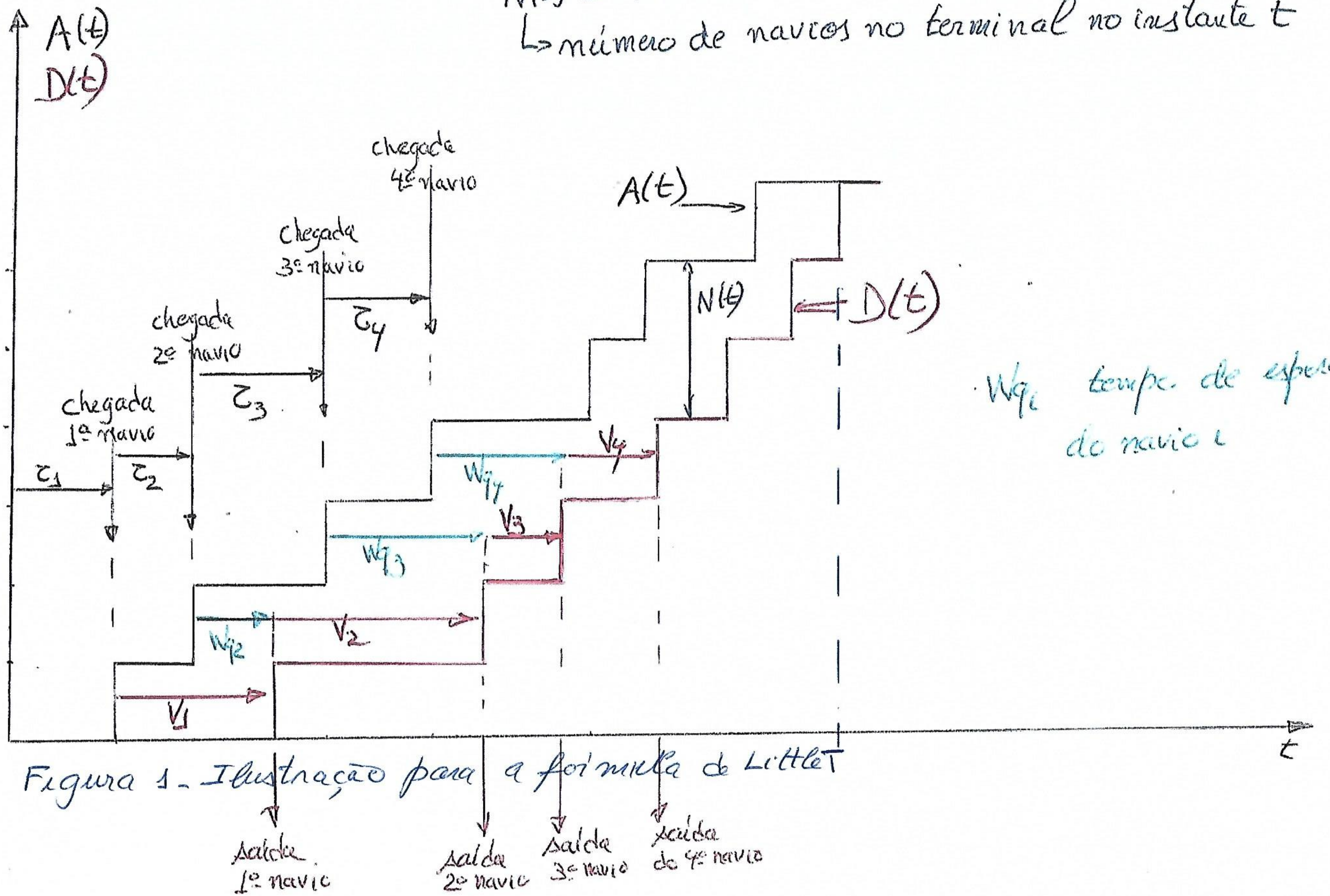


Figura 1 - Ilustração para a fórmula de Little

2 – Processos de Fila com Chegada Poisson – a Visão de quem Chega

Seja $\{N(t), t \geq 0\}$ o número de clientes no sistema para uma fila em que o processo de chegada é Poisson. Sejam $p_j(t) = P[N(t)=j]$ e $a_j(t)$ a probabilidade de que um cliente, ao chegar num instante t, encontre j clientes no sistema. O objetivo é mostrar que $a_j(t) = p_j(t)$; isto é, deseja-se mostrar que a probabilidade de que um cliente, ao chegar, encontre j clientes no sistema no instante t é igual à probabilidade de haver j clientes no sistema no instante t.

De acordo com a definição de $a_j(t)$,

$$a_j(t) = \lim_{\Delta t \to 0} P[N(t)=j \,/\, N(t+\Delta t)-N(t)=1] \quad (5)$$

Trabalhando com a definição de probabilidade condicional e, em seguida, com o fato de o processo de Poisson ter incrementos independentes, de (5) resulta:

$$a_j(t) = \lim_{\Delta t \to 0} \left[\frac{P[N(t)=j,\ N(t+\Delta t)-N(t)=1]}{P[N(t+\Delta t)-N(t)=1]} \right] =$$

$$= \lim_{\Delta t \to 0} \left[\frac{P[N(t)=j]\, P[N(t+\Delta t) - N(t) = 1]}{P[N(t+\Delta t) - N(t) = 1]} \right] = p_j(t)$$

Ou $$a_j(t) = p_j(t) \qquad (6)$$

Em regime estacionário, $t \to \infty$

$$a_j = p_j \qquad (7)$$

3 – Para uma fila M/G/1, $p_j = \pi_j$

Seja $\{N(t), t \geq 0\}$ o número de clientes no sistema no instante para uma fila M/G/1. Como se admite uma distribuição de probabilidades genérica para o tempo de atendimento, o processo $\{N(t), t \geq 0\}$ não é uma cadeia de Markov em tempo contínuo. No entanto, se $\rho = \lambda E(V) < 1$, existe a distribuição estacionária p_j, $j = 0, 1, 2, \ldots$

$$p_j = \lim_{t \to \infty} P[N(t) = j / N(0) = i]$$

O objetivo é mostrar que esta distribuição p_j é igual à distribuição π_j da cadeia de Markov associada à fila M/G/1. Para tanto, sejam:

a_j – a probabilidade de que, em regime estacionário, um cliente encontre, ao chegar, j clientes no sistema; e

d_j – a probabilidade de que, em regime estacionário, um cliente deixe, ao sair, j clientes no sistema.

Propriedade $\quad a_j = d_j$

Considerem-se os processos de contagem $\{A(t), t \geq 0\}$ e $\{D(t), t \geq 0\}$ sendo $A(t)$ o número de chegadas até o instante t e $D(t)$ o número de saídas até o instante t. $N(t) = A(t) - D(t)$ é o número de clientes no sistema no instante t (premissa $N(0) = 0$). Sejam ainda $\{A_j(t), t \geq 0\}$ o contador de chegadas em que o cliente encontra j clientes no sistema e $\{D_j(t), t \geq 0\}$ o contador de saídas em que o cliente deixa j clientes no sistema. Em outras palavras, o primeiro processo conta as passagens de $N(t) = j$ para $N(t) = j+1$ e o segundo conta as passagens de $N(t) = j+1$ para $N(t) = j$.

Para qualquer processo de fila com chegada individual e atendimento individual, admitindo $N(0) = A(0) = D(0) = 0$; tem-se

$$A_j(t) - D_j(t) \leq 1 \qquad (8)$$

e

$$\frac{A_j(t) - D_j(t)}{A(t)} \leq \frac{1}{A(t)} \qquad (9)$$

Considere-se o limite de (9) quando t tende a infinito

$$\lim_{t\to\infty} \frac{A_j(t)}{A(t)} = \lim_{t\to\infty} \frac{D_j(t)}{A(t)} =$$

$$= \lim_{t\to\infty} \frac{D_j(t)}{D(t)} \times \frac{D(t)}{A(t)} =$$

$$= \lim_{t\to\infty} \frac{D_j(t)}{D(t)} \times \lim_{t\to\infty} \frac{D(t)}{A(t)} = \lim_{t\to\infty} \frac{D_j(t)}{D(t)}$$

ou reescrevendo

$$\lim_{t\to\infty} \frac{A_j(t)}{A(t)} = \lim_{t\to\infty} \frac{D_j(t)}{D(t)} \qquad (10)$$

Como o lado esquerdo de (10) é igual a a_j e o lado direito é igual a d_j, fica demonstrado:

$$a_j = d_j \qquad (11)$$

Tendo em vista que:

i) na seção 2 destas notas de aula foi mostrado que, para filas com chegada Poisson, $a_j = p_j$; e

ii) a cadeia de Markov da fila M/G/1 é observada nos instantes de saídas dos clientes e não inclui aquele que está saindo o que implica que $d_j = \pi_j$

conclui-se que, para a fila M/G/1,

$$p_j = \pi_j \quad , \; j = 0, 1, 2, \ldots \qquad (12)$$

Resultados Complementares

1 - A partir de (12),

$$L = \sum_{j=0}^{\infty} j\, p_j = \sum_{j=0}^{\infty} j\, \pi_j = L' =$$

$$= \rho + \frac{\rho^2 (1+\alpha^2)}{2(1-\rho)} \qquad (P\text{-}K)$$

2 - Utilizando agora a fórmula de Little, obtém-se a expressão para o tempo médio

de permanência de um cliente no sistema:

$$W = \frac{L}{\lambda} = E(V) + \frac{\rho^2(1+\alpha^2)}{2\lambda(1-\rho)} \qquad (13)$$

Portanto, o tempo médio de espera para um cliente de fila M/G/1 é dado por:

$$W_q = \frac{\rho^2(1+\alpha^2)}{2\lambda(1-\rho)} \qquad (14)$$

e o número médio de clientes na fila, L_q, é dado por:

$$L_q = \frac{\rho^2(1+\alpha^2)}{2(1-\rho)} \qquad (15)$$

4_ Fila M/M/1 _ Distribuição do tempo de permanência no sistema e distribuição dos intervalos entre saídas consecutivas

4.1 Distribuição do tempo de permanência no sistema (em regime estacionário)

Admita-se que um cliente, ao chegar, encontre j clientes no sistema: um sendo atendido e (j−1) na fila. Assim, o cliente espera em fila durante:

1 - o tempo para completar o atendimento do cliente em serviço, que, pela falta de memória dos tempos de atendimento, é uma variável exponencial de média $1/\mu$;

2 - os tempos de atendimentos dos $(j-1)$ clientes em fila quando ele chegar.

O seu tempo de espera é a soma de j variáveis exponenciais de média $1/\mu$. Considerando que seu tempo de atendimento é também exponencial de média $1/\mu$, o tempo de permanência do cliente que, ao chegar, encontra j clientes no sistema é uma variável Erlang de ordem $(j+1)$. Isto é,

$$h_{TP/encontrar\,j}^{(t)} = \frac{\mu^{j+1} t^j e^{-\mu t}}{j!} \qquad (16)$$

Eliminando-se a condicionalidade,

$$h_{TP}^{(t)} = \sum_{j=0}^{\infty} h_{TP/encontrar\,j}^{(t)} \times P[encontrar\ j] =$$

$$\sum_{j=0}^{\infty} \frac{\mu^{j+1} t^j e^{-\mu t}}{j!} a_j = \sum_{j=0}^{\infty} \frac{\mu^{j+1} t^j e^{-\mu t}}{j!} \times p_j =$$

$$= \sum_{j=0}^{\infty} \frac{\mu^{j+1} t^j e^{-\mu t}}{j!} (1-\rho)\rho^j =$$

$$= \mu(1-\rho)e^{-\mu t} \sum_{j=0}^{\infty} \frac{(\rho\mu t)^j}{j!} =$$

$$= \mu(1-\rho)e^{-\mu t} e^{\rho\mu t} = \mu(1-\rho)\, e^{-\mu(1-\rho)t} \qquad (17)$$

Isto é, a distribuição do tempo de permanência de um cliente da fila M/M/1 é exponencial com média igual a $1/\mu(1-\rho)$.

4.2 Distribuição dos intervalos entre saídas consecutivas de uma fila M/M/1

Admita-se que um cliente, ao sair, deixe $j \geq 1$ clientes no sistema; em tal situação, o intervalo até a próxima saída corresponde ao tempo de atendimento do cliente que vai sair, tendo, portanto, distribuição exponencial de média $1/\mu$. Caso um cliente, ao sair, deixe zero clientes no sistema, o intervalo até a próxima saída será a soma do intervalo de tempo residual até a próxima chegada, que é uma variável exponencial de média $1/\lambda$, com o tempo de atendimento do cliente que chegar. Isto é

$$U = \theta + V$$

sendo θ exponencial de média $1/\lambda$ e V exponencial de média $1/\mu$.

$$P[U \leq t / \theta = t'] = P[V \leq t - t'] =$$

$$= 1 - e^{-\mu(t-t')}$$

Eliminando-se a condicionalidade,

$$P[U \leq t] = \int_0^t P[U \leq t/\theta = t'] f_\theta(t')dt' =$$

$$\int_0^t (1 - e^{-\mu(t-t')}) \lambda e^{-\lambda t'} dt'$$

Tarefa para os alunos

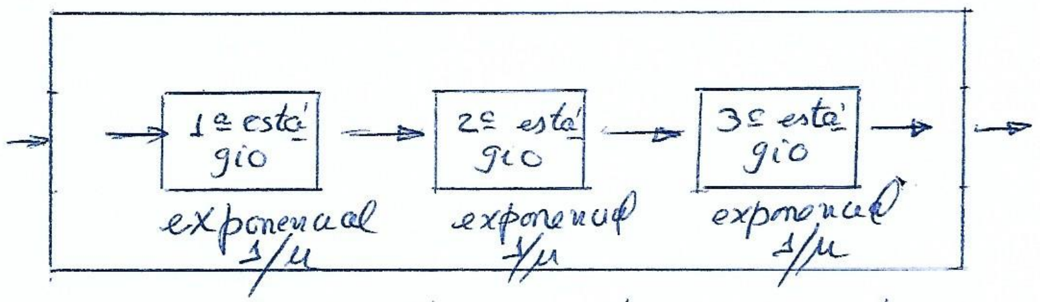
Completar a dedução, observando que a distribuição não condicional dos intervalos entre saídas consecutivas é a soma de uma densidade exponencial, ponderada por $(1-d_0)$, com a função densidade da variável U, ponderada por d_0, e que $d_0 = p_0$.

Estudo da Fila $M/E_k/1$ por meio de cadeia de Markov em tempo contínuo

Considere-se uma fila $M/G/1$ em que os tempos de atendimento têm distribuição Erlang de ordem k, $g_V(t) = \mu^k t^{k-1} e^{-\mu t}/k!$.

Como já mencionado na seção 3 dessas notas de aula o processo $\{N(t), t \geq 0\}$, sendo $N(t)$ o número de clientes no sistema no instante t, não é uma cadeia de Markov em tempo contínuo. Tendo em vista, no entanto que o tempo de atendimento pode ser decomposto em k estágios de atendimento exponencial de média $1/\mu$, é possível estudar essa fila por meio de cadeia de Markov em tempo contínuo, mudando-se o foco do estudo. Em lugar do número de clientes no sistema no instante t, $N(t)$, estuda-se o número de estágios de serviço no sistema no instante t, $N_e(t)$.

Com a finalidade de trabalhar com um exemplo específico, admita-se que o tempo de atendimento seja uma variável aleatória com distribuição Erlang de ordem 3 ; assim, como mostra a ilustração abaixo, um atendimento é composto por 3 estágios de serviço exponencial com média $1/\mu$



Tempo de atendimento Erlang de ordem 3 decomposto em 3 estágios de atendimento exponencial

O diagrama das taxas de transição para o processo estocástico $\{N_e(t), t \geq 0\}$ para a fila $M/E_3/1$ é mostrado abaixo.

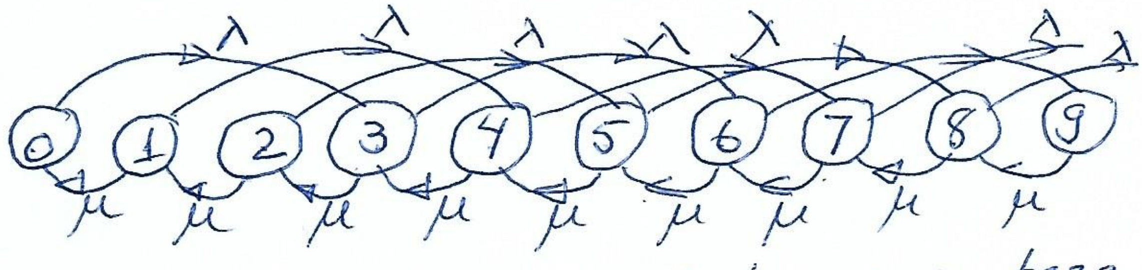


Diagrama das taxas de transição para o número de estágios de serviço de uma fila $M/E_3/1$

A partir do diagrama das taxas de transição, são escritas as equações de equilíbrio (em regime estacionário) para cada estado j do processo $\{N_e(t), t \geq 0\}$.

Para o estado zero $\lambda p_0^e = \mu p_1^e$

Para o estado 1 $(\lambda+\mu) p_1^e = \mu p_2^e$

Para o estado 2 $(\lambda+\mu) p_2^e = \mu p_3^e$

Para o estado 3 $(\lambda+\mu) p_3^e = \lambda p_0^e + \mu p_4^e$

Para o estado 4 $(\lambda+\mu) p_4^e = \lambda p_1^e + \mu p_5^e$

e, assim, sucessivamente. Obtendo-se a expressão de cada p_j^e em função de p_0^e, utiliza-se a relação $\sum_{j=0}^{\infty} p_j^e = 1$ a fim de determinar a condição de existência da distribuição estacionária e calcular o valor de p_0^e.

O passo seguinte é obter a distribuição estacionária $p_j = \lim_{t\to\infty} P[N(t)=j / N(0)=i]$, utilizando as seguintes relações:

$$p_0 = p_0^e$$
$$p_1 = p_1^e + p_2^e + p_3^e$$
$$p_2 = p_4^e + p_5^e + p_6^e$$
$$p_3 = p_7^e + p_8^e + p_9^e$$

e, assim, sucessivamente.