Roberto Ramos de Morais

3 TEORIA DAS FILAS

No cotidiano encontram-se filas em diversas situações: ao ir à padaria para comprar pão, na sala de espera do dentista, na compra de ingressos para o cinema, veículos em uma barreira de pedágio. São situações fáceis de visualizar por fazer parte da vida das pessoas. Mas, há outras situações distantes da realidade da maioria das pessoas, como peças esperando para serem processadas por uma máquina. Outras, ainda, são intangíveis, como a espera para ser atendido ao telefone pelo funcionário de um SAC.

O que interessa no contexto deste trabalho é estudar os parâmetros das filas (quantidade de elementos na fila e tempo de espera, por exemplo) como indicadores de desempenho de um processo ou de um sistema.

O fenômeno das filas foi estudado inicialmente por Erlang em 1909, na Dinamarca, quando ele trabalhava para a Companhia Telefônica daquele país. O intuito era diminuir a quantidade de clientes esperando pelo atendimento das telefonistas.

Outros autores que contribuíram para o avanço da teoria das filas foram (FOGLIATTI; MATTOS, 2007):

- Adams, em 1936, e Tanner, em 1951, sobre tempo médio de espera para pedestres atravessarem uma rua sem semáforo;
- Everett, em 1953, sobre fluxo de barcos em terminais portuários;
- Cobham, em 1954, sobre reparo de máquinas;
- Bailey, em 1954, sobre o fluxo de pacientes em pronto socorros;
- Morse, em 1962, e Prabhu, em 1965, sobre formação de estoques;
- Bitran e Morabito, em 1996, sobre sistemas de manufatura;
- Mendonça e Morabito, em 2001, sobre sistemas médicos de emergência.

Com o avanço da computação, os estudos de filas também se desenvolvem rapidamente, permitindo a criação de modelos cada vez mais próximos da realidade.

AGNER KRARUP ERLANG (1878-1929)

Matemático e engenheiro dinamarquês. foi a primeira pessoa a estudar o problema de redes de telefonia. Trabalhando para a Companhia Telefônica, estudou o fluxo de ligações em um pequeno vilarejo. A partir desse estudo, criou a fórmula conhecida como a fórmula de Erlang, para calcular a fração de ligações que tentavam chamar alguém fora do vilarejo e que tinham que esperar (fila) devido todas as linhas estarem em uso. Embora o modelo de Erlang seja simples, as complexas redes de telefonia atuais ainda se baseiam em seu trabalho. Em 1909 surge o primeiro trabalho, *The Theory of Probabilities and Telephone Conversations*, demonstrando que ligações telefônicas distribuídas aleatoriamente seguiam a lei de distribuição de Poisson. Outro trabalho publicado é *Solution of some Problems in the Theory of Probabilities of Significance in Automatic Telephone Exchanges* (1917).

Fórmulas de Erlang

A fórmula de Erlang B é utilizada no estudo de sistemas com perdas e utilizada para dimensionamento de troncos telefônicos e qualquer outro equipamento que receba tráfego.

$$B(s,a) = \frac{a^s/_{s!}}{\sum_{k=0}^{s} a^k/_{k!}}$$

- B(s,a) = Probabilidade de todos os troncos estarem ocupados ou "perda admissível"
- s = Quantidade de troncos
- a = Densidade de tráfego. É a relação entre as chamadas que chegam e as que são atendidas em um determinado intervalo de tempo

A fórmula de Erlang C é utilizada no estudo de sistemas com perdas e é utilizada para dimensionamento de recursos em qualquer sistema constituído por filas, inclusive em centrais de atendimento.

$$C(s,a) = \frac{a^s/_{s!\,(1-a/_s)}}{\sum_{k=0}^{s} \frac{a^k}{k!} + a^s/_{s!\,(1-a/_s)}}$$

- C(s,a) = Probabilidade de todos os agentes estarem ocupados ou a probabilidade do cliente ter que esperar na fila
- s = Quantidade de agentes
- a = Densidade de tráfego. É a relação entre as chamadas que chegam e as que são atendidas em um determinado intervalo de tempo

Fonte: Wikipedia, 2018; Erlang.com.br, 2007

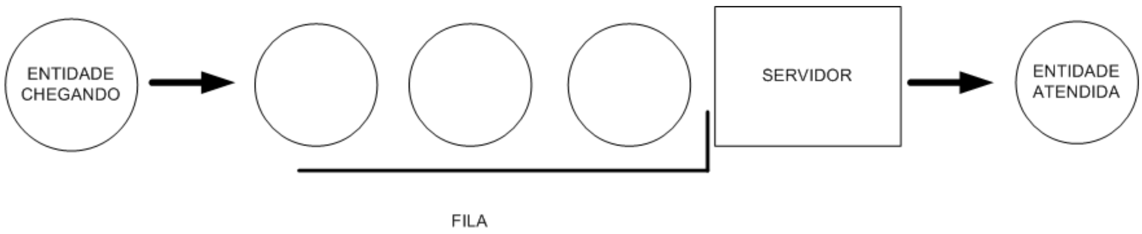
3.1 Terminologia e notação básica da teoria de filas

Os elementos que compõem uma fila são:

- **Entidades**: é o elemento que entra, é atendido ou processado, e sai do sistema, atendido ou transformado. Exemplos: clientes em uma agência bancária, peças em uma linha de produção.
- **Servidor ou atendente**: é o recurso que atende ou processa a entidade, podendo ser um ou mais. Exemplos: caixas em uma agência bancária, máquinas operatrizes em uma linha de produção.
- **Área de espera**: local onde a fila se forma.

Estes elementos são apresentados na Figura 3.1:

Figura 3.1: Elementos de um sistema com fila



Fonte: o autor

Há vários parâmetros que definem a fila, como comprimento ou número da fila, que se refere à quantidade de entidades esperando pelo atendimento; quantidade de entidades no sistemas, igual à soma do comprimento de fila e as entidades que estão sendo atendidas; mais aqueles que estão sendo atendidos; tempo de espera, referente à permanência da entidade na fila; tempo de atendimento, correspondente ao tempo consumido pelo servidor no processamento da entidade; tempo de atravessamento, soma dos tempos de espera e de atendimento, indicando o tempo total entre a chegada e saída da entidade do sistema.

- O tamanho da fila varia no tempo de forma aleatória, em função das chegadas dos clientes que são servidos e deixam o sistema

3.2 Classificação das filas

As filas podem ser classificadas conforme a quantidade de estágios ou atendentes pelos quais a entidade deve passar (MAISTER, 1995 apud CORRÊA; CORRÊA, 2019):

1. **Filas de sistemas estágio único:** neste tipo há uma fila e um único ponto de atendimento, após o qual a entidade sai do sistema. Por exemplo, o caixa de pagamento em uma lanchonete. Este tipo se subdivide em:
    a. *Servidores paralelos:* há vários atendentes, cada um com sua própria fila. Exemplo: caixa de supermercados, barreiras de pedágio.
    b. *Fila única:* há uma única fila que é atendida por vários atendentes em paralelo, como em um banco.
    c. *Filas concorrentes:* há um único servidor que atende a várias filas. Por exemplo: um semáforo em um cruzamento.
    d. *Filas discriminadas:* há a separação por tipo de entidade ou objetivo de atendimento, como caixas bancárias para idosos, caixas de supermercados para clientes com compras até dez itens.
2. **Filas de sistemas de estágios múltiplos:** as entidades devem passar por diversos atendentes, como matéria prima em uma linha de produção que passa por vários processos até ser transformada em produto acabado ou pacientes que passam por diversos exames em um pronto socorro.

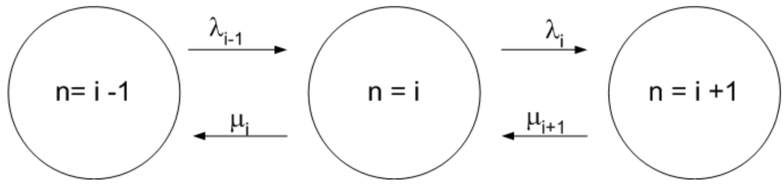
3.3 Processos estocásticos

Entende-se por processo estocástico todo aquele cujas variáveis são aleatórias, ou seja, assumem valores (estados) conforme uma distribuição de probabilidades ao longo do tempo (FOGLIATTI; MATTOS, 2007; HILLIER; LIEBERMAN, 2013; TAHA, 2008). Estes processos podem ser dos tipos:

- **Processos de estado-contínuo:** o estado do sistema é representado por variáveis contínuas (p. ex., tempo de espera);
- **Processos de estado-discreto:** representados por variáveis discretas (p. ex., clientes em uma fila).

- **Processos de Markov:** os estados futuros do processo dependem apenas do estado presente, sendo independentes do passado. Neste caso, o tempo de um estado possui uma distribuição exponencial (sem memória - *memoryless*). O processo estocástico que possui esta característica é chamado de cadeia de Markov.
- **Processos de nascimento e morte:** são processos discretos de Markov nos quais as transições de estado estão restritas aos estados vizinhos. O sistema flui de um estado para outro dependendo de λ (taxa média de chegada de clientes) e μ (taxa média de serviços realizados pelo sistema). O estado *n* do sistema pode ser representado por valores inteiros e a mudança só se dará para os estados *n+1* (próximo estado) ou *n-1* (estado anterior). Esse caso é classificado como estado absorvente, isto é, pode retornar ao mesmo estado em uma transição (saída ou chegada). Exemplo: central telefônica no estado *n*, que é alterado para *n-1* pelo encerramento de uma ligação (morte) ou para *n+1* por uma nova ligação (nascimento). A Figura 3.2 apresenta estas transições:

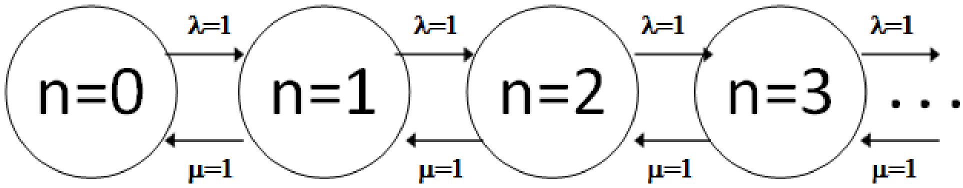
Figura 3.2: mudanças de estados de um sistema.



Fonte: o autor.

Por exemplo, considerando o comportamento a entrada de clientes em uma loja. A chegada de um cliente não influencia a chegada do próximo cliente e nem foi influenciada pela chegada do anterior. O estado (quantidade de clientes na loja) é função apenas da quantidade anterior que havia na loja. Se havia 2 clientes e mais um entra na loja, o sistema passa do estado 2 para o estado 3. Se um cliente sai, o sistema retorna ao estado 2. Ou seja, o estado aumenta conforme a taxa de chegada λ e diminui conforme a taxa de atendimento μ (Figura 3.3):

Figura 3.3: Exemplo.



Fonte: o autor.

Andrei Andreyevich Markov

Matemático russo, nascido em Riazan, 14 de junho de 1856 e falecido em São Petersburgo, 20 de julho de 1922. Formado na Universidade Estatal de São Petersburgo em 1878, onde também exerceu o cargo de professor em 1886. Iniciou seus estudos em limite de integrais e teoria da aproximação. A partir de 1900 trabalhou com métodos de frações contínuas, que iniciou com Pafnuti Tchebychev na teoria da probabilidade, além de provar o teorema do limite central. Markov é lembrado pelo seu estudo de cadeias de Markov.
Fonte: Wikipedia (2018b)

3.3 Notação de Kendal

A notação de Kendal, desenvolvida em 1953, visa indicar os tipos de distribuição de probabilidade associada à cada etapa do processo. A especificação de um modelo de fila requer que seis características (parâmetros) sejam declaradas e a notação mais conhecida é a forma A/S/m/B/K/SD (RAGSDALE, 2014):

- **A:** Processo de chegada, caracterizado pela distribuição de probabilidade dos períodos de tempo decorridos entre as chegadas de clientes no sistema de fila. Os tipos de distribuição para tempos de chegada são:
  - *M:* distribuição exponencial ou markoviana (M do inglês *memoryless*);
  - E_k: distribuição Erlang com parâmetro k;
  - H_k: distribuição hiperexponencial com parâmetro k;

    - *G:* distribuição geral (modelo não especificado e resultados válidos para todas as distribuições)
    - *D:* determinística (tempos constantes com variância zero)
- **S:** Tempo de serviço, caracterizado pela distribuição de probabilidade dos períodos de tempo de serviço para cada cliente servido. São dos mesmos tipos das distribuições dos tempos de chegada.
- **m:** Número de servidores, que indica a quantidade de servidores ou atendentes disponíveis. Conceitualmente serão todos idênticos e qualquer cliente pode ser alocado a qualquer dos servidores disponíveis.
- **B:** Capacidade da fila, indicador do número máximo de clientes que podem entrar e permanecer na fila.
- **K:** Tamanho da população, correspondente número total de clientes que, potencialmente, podem vir a utilizar os recursos do sistema de fila
- **SD:** Disciplina de atendimento, ordem como os clientes aguardam para acessar os serviços. A disciplina mais comum é a do tipo FIFO, mas existem LIFO, randômico, menor tempo de serviço primeiro (SPTF)

  Quando não for especificado:
    - a capacidade dos sistemas (B) é infinita (ou indicação de ∞);
    - o tamanho da população (K) é infinito (ou indicação de ∞);
    - o tipo de disciplina de serviço é definido com FIFO;
    - as chegadas são individuais (os clientes não chegam agrupados);

Por exemplo um sistema de fila que foi retratado com o emprego da notação de Kendall como M/G/4/50/2000/LIFO significa

- M: os tempos entre chegadas são de acordo com uma distribuição exponencial;
- G: tempos de serviço ou atendimento são distribuídos de forma arbitrária;
- 4: há quatro servidores;
- 50: a capacidade do sistema (tamanho da fila+atendimento) é limitada a 50 clientes (4 sendo servidos e, no máximo, 46 aguardando na fila);
- 2000: a fonte de origem dos clientes tem capacidade limitada para 2000 clientes;

- LIFO: a disciplina de serviço é "o último a chegar é o primeiro a ser atendido".

3.3.1 Modelo M/M/1

Neste tópico são apresentados os cálculos dos parâmetros para o modelo de filas com as seguintes características:

- Tempos entre chegada de clientes e tempos de serviços do servidor são descritos por distribuição exponencial (processos de Poison) M: *memoryless)*;
- Fila de um único servidor;
- Não há limitação para tamanho de fila (capacidade) nem de população;
- A disciplina de atendimento é FIFO (*First in, first out* – primeiro que entra, primeiro que sai).

Os dados deste sistema são:

- Taxa de chegada de clientes por unidade de tempo (λ):

$$\lambda = \frac{C}{T}$$

Onde:

- C: quantidade de clientes ou usuários que chegaram em um determinado intervalo de tempo;
- T: intervalo de tempo.

O tempo médio entre chegadas (T_{mc}) é calculado da seguinte forma:

$$T_{mc} = \frac{1}{\lambda}$$

- Taxa de demanda de serviço do servidor por unidade de tempo (μ)

$$\mu = \frac{A}{T}$$

Onde:

A: Quantidade de clientes atendidos durante um período de tempo;

T: intervalo de tempo.

O tempo médio de atendimento (T_{ma}) é calculado da seguinte forma:

$$T_{ma}\colon\ \frac{1}{\mu}$$

Exemplo: Quais as taxas de chegada e de atendimento de um sistema no qual, ao longo de 120 minutos, chegaram 90 clientes e foram atendidos 84 clientes? Quais os tempos médios de chegada e de atendimento?

Nesta situação, tem-se a taxa de chegada:

$$\lambda = \frac{C}{T} = \frac{90\ clientes}{120\ minutos} = 0{,}75\ {}^{cliente}/_{minuto}$$

O tempo médio entre chegadas é:

$$T_{mc} = \frac{1}{\lambda} = \frac{1}{0{,}75} = 1{,}33 \ldots minutos$$

A taxa de atendimento é:

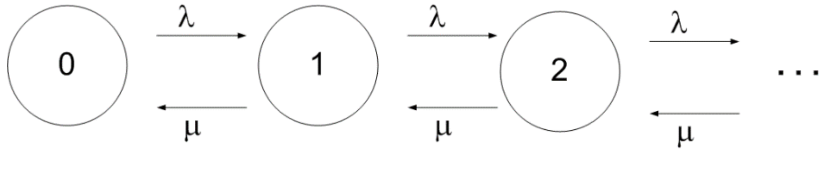
$$\mu = \frac{A}{T} = \frac{84}{120} = 0{,}7\ {}^{cliente}/_{minuto}$$

Assim, o tempo médio de atendimento é:

$$T_{ma}\colon\ \frac{1}{\mu} = \frac{1}{0{,}7} = 1{,}428\ minutos$$

Para o modelo M/M/1, os processos de vida e morte que alteram o estado do sistema seguem mudanças de acordo com as taxas de chegada e de atendimento, conforme a Figura 3.3:

Figura 3.3: processos de vida e morte com um servidor.



Fonte: o autor.

Os círculos representam possíveis estados de clientes na fila. Os cálculos dos parâmetros da fila são:

- ρ: Intensidade de tráfego:

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

- A condição de estabilidade é: $\rho < 1$. Se $\rho > 1$ significa que a fila cresce indefinidamente e os demais cálculos não são possíveis de serem realizados.

- p_i: probabilidade de o sistema estar no estado *n* (conter *n* clientes)

$$p_n = (1-\rho).\rho^n$$

- Probabilidade de haver n ou mais clientes na fila:

$$P = \rho^n$$

- Número médio de clientes no sistema:

$$E(n) = \frac{\rho}{1-\rho}$$

- Número médio de clientes na fila:

$$E(n_q) = \frac{\rho^2}{1-\rho}$$

- Tempo médio no sistema (ou de resposta):

$$E(r) = \frac{1}{\mu(1-\rho)}$$

- Tempo médio de espera

$$E(w) = \frac{\rho}{\mu(1-\rho)}$$

Exemplo 1: Em um guichê de venda de ingressos para jogo de futebol, a taxa média de chegada é de 4,25 torcedores por minuto. A taxa de atendimento do guichê é de 4,5 torcedores por minuto.

a. Qual a intensidade do tráfego? O sistema é estável?
b. Qual a probabilidade de haver 3 torcedores no sistema?
c. Qual a probabilidade de haver 10 ou mais torcedores na fila?
d. Qual o número médio de torcedores no sistema?
e. Qual o número médio de torcedores na fila?
f. Qual o tempo médio no sistema?
g. Qual o tempo médio de espera?

Identifica-se, no enunciado, os valores básicos para o desenvolvimento do exercício:

- Taxa de chegada: λ = 4,25 torc/min
- Taxa de atendimento: μ = 4,5 torc/min

Desta maneira, a intensidade do tráfego é:

$$\rho = \frac{\lambda}{\mu} = \frac{4,25}{4,5} = 0,9444 \ldots$$

Portanto, como a intensidade de tráfego é menor do que 1 ($\rho < 1$), o sistema é estável. Continua-se, assim, com os demais cálculos.

A probabilidade de haver 3 torcedores (n=3) no sistema é calculada da seguinte forma:

$$P = (1-\rho).\rho^n = (1-0,944).0,944^3 = 0,0468\ ou\ 4,68\%$$

A probabilidade de haver 10 ou mais torcedores na fila:

$$p_n = \rho^n = 0,944^{10} = 0,5646\ ou\ 56,46\%$$

O número médio de torcedores no sistema (fila+atendimento):

$$E(n) = \frac{\rho}{1-\rho} = \frac{0,944}{1-0,944} = 17\ torcedores$$

O número médio de torcedores na fila:

$$E(n_q) = \frac{\rho^2}{1-\rho} = \frac{0,944^2}{1-0,944} = 16,06\ torcedores$$

O tempo médio no sistema:

$$E(r) = \frac{1}{\mu(1-\rho)} = \frac{1}{4,5.(1-0,944)} = 4\ minutos$$

O tempo médio de espera:

$$E(w) = \frac{\rho}{\mu(1-\rho)} = \frac{0,944}{4,5.(1-0,944)} = 3,7777 \ldots min\ ou\ 3\ min\ e\ 47\ segundos$$

3.3.2 Modelo M/M/m

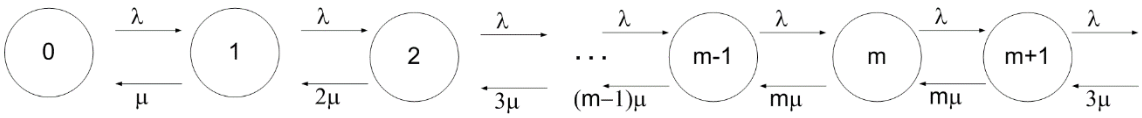
Neste tópico as características da fila são:

- Tempos entre chegada de clientes e tempos de serviços do servidor são descritos por distribuição exponencial (processos de Poison) M: *memoryless)*, da mesma forma que o modelo M/M/1;
- Dois ou mais servidores;
- Não há limitação para tamanho de fila (capacidade) nem de população

- A disciplina de atendimento é FIFO (*First in, first out* – primeiro que entra, primeiro que sai).

Para o modelo M/M/m, os processos de vida e morte que alteram o estado do sistema seguem mudanças de acordo com as taxas de chegada (aumento) e de atendimento multiplicado pela quantidade de atendentes ocupados, até o limite da quantidade de servidores disponíveis, conforme a Figura 3.4:

Figura 3.4: processos de vida e morte com um servidor.



Fonte: o autor.

Os círculos representam os possíveis estados de clientes na fila, que se alteram conforme descrito no parágrafo anterior.

A intensidade de tráfego é representada por ρ:

$$\rho = \frac{\lambda}{m\mu}$$

- Onde m é a quantidade de servidores/atendentes disponíveis.
- Condição de estabilidade: $\rho < 1$, da mesma forma que para M/M/1.

- Probabilidade de zero cliente no sistema

$$p_0 = \left[1 + \frac{(m\rho)^m}{m!\,(1-\rho)} + \sum_{n=1}^{m-1} \frac{(m\rho)^n}{n!}\right]^{-1}$$

A somatória do terceiro elemento da expressão indica que o valor de n varia de 1 atendente até a quantidade de 1 atendente a menos que o máximo. Por exemplo, se houver 7 atendentes, n variará de 1 até 6.

- p_i: probabilidade de o sistema estar no estado *n* (conter *n* clientes). Este cálculo depende se o estado *n* for menor do que a quantidade de servidores (*n*<*m*) ou maior (*n*>*m*):

$$p_n = \begin{cases} \dfrac{(m\rho)^n}{n!} p_0, & n < m \\ \dfrac{\rho^n m^m}{m!} p_0, & n \geq m \end{cases}$$

- Probabilidade de haver m ou mais clientes na fila:

$$P = \frac{(m\rho)^m}{m!\,(1-\rho)} p_0$$

- Número médio de clientes no sistema:

$$E(n) = m\rho + \frac{\rho.P}{1-\rho}$$

- Número médio de clientes na fila:

$$E(n_q) = \frac{\rho P}{1-\rho}$$

- Tempo médio no sistema (ou de resposta):

$$E(r) = \frac{1}{\mu}\left[1 + \frac{P}{m(1-\rho)}\right]$$

- Tempo médio de espera (fila)

$$E(w) = \frac{P}{m\mu(1-\rho)}$$

- Taxa de utilização média de cada servidor. Representa a fração do tempo que o servidor ficou ocupado.

$$U = \rho$$

Exemplo: Uma agência bancária possui cinco caixas automáticos (ATMs). Chegam 2 clientes por minuto para utilizar os caixas, formando uma fila única. O tempo médio de permanência no caixa por cliente é de 2 minutos. Calcule os parâmetros desta fila. Para o cálculo da probabilidade de o sistema estar no estado n, considerar n=5 e n=7.

- ρ: Intensidade de tráfego:

$$\rho = \frac{\lambda}{m\mu} = \frac{2}{5.\frac{1}{2}} = 0{,}8$$

Uma vez que $\rho < 1$, o sistema é estável. Portanto, pode-se prosseguir com os cálculos.

- Probabilidade de zero cliente no sistema

$$p_0 = \left[1 + \frac{(m\rho)^m}{m!\,(1-\rho)} + \sum_{n=1}^{m-1} \frac{(m\rho)^n}{n!}\right]^{-1}$$

$$= \left[1 + \frac{(5.0{,}8)^5}{5!\,(1-0{,}8)} + \frac{(5.0{,}8)^1}{1!} + \frac{(5.0{,}8)^2}{2!} + \frac{(5.0{,}8)^3}{3!} + \frac{(5.0{,}8)^4}{4!}\right]^{-1}$$

$$= 0{,}0130 \ ou \ 1{,}30\%$$

- Probabilidade de haver m ou mais clientes na fila:

$$P = \frac{(m\rho)^m}{m!\,(1-\rho)} p_0 = \frac{(5.\ 0{,}8)^5}{5!\,(1-0{,}8)} 0{,}013 = 0{,}5541 \ ou \ 55{,}41\%$$

- Número médio de clientes no sistema:

$$E(n) = m\rho + \frac{\rho.P}{1-\rho} = 5.0{,}8 + \frac{0{,}8.0{,}5541}{1-0{,}8} = 6{,}22 \ clientes$$

- Número médio de clientes na fila:

$$E(n_q) = \frac{\rho P}{1-\rho} = \frac{0{,}8.0{,}5541}{1-0{,}8} = 2{,}22 \ clientes$$

- Tempo médio no sistema (ou de resposta):

$$E(r) = \frac{1}{\mu}\left[1 + \frac{P}{m(1-\rho)}\right] = \frac{1}{0{,}5}\left[1 + \frac{0{,}5541}{5(1-0{,}8)}\right] = 3{,}11 \ minutos$$

- Tempo médio de espera:

$$E(w) = \frac{P}{m\mu(1-\rho)} = \frac{0{,}5541}{5.0{,}5(1-0{,}8)} = 1{,}11 \ minutos$$

- Taxa de utilização média de cada servidor

$$U = \rho = 0{,}8$$

3.3 Modelo Geral de Little

Modelo estabelecido por John Dutton Conant Little (1928-atual), apresenta expressões gerais para cálculos dos parâmetros de filas estáveis (ρ<1), populações infinitas e atendimento por ordem de chegada (FIFO). As expressões são apresentadas no Quadro 3.1 (PRADO, 2017; CORRÊA; CORRÊA, 2019):

Quadro 3.1

	MODELO 1	MODELO 2
Tipo de atendimento	Exponencial	Constante
Número médio de clientes na fila (L_q)	$\frac{\lambda^2}{\mu(\mu-\lambda)}$	$\frac{\lambda^2}{2\mu(\mu-\lambda)}$
Tempo médio de espera (W_q)	$\frac{\lambda}{\mu(\mu-\lambda)}$	$\frac{\lambda}{2\mu(\mu-\lambda)}$
Probabilidade de n clientes no sistema (P_n)	$(1-\rho).\rho^n$	$(1-\rho).\rho^n$

Fonte: (CORRÊA; CORRÊA, 2019; PRADO, 2017)

3.5 Exercícios

1. Em uma loja de materiais de construção os clientes retiram suas compras (pequenos volumes) em um balcão localizado no subsolo. O atendimento é único e consome um tempo médio de 8 minutos por cliente. Em média, a cada 10 minutos chega um cliente para fazer a retirada de sua compra.
    a. O sistema é estável?
    b. Na área de espera há cinco cadeiras. Qual a probabilidade de que haja clientes esperando em pé?
    c. Qual o número médio de clientes esperando para serem atendidos?
    d. Quanto tempo, em média, os clientes esperam para serem atendidos?

2. Um supermercado possui 8 caixas que atendem uma fila única de clientes com compras até 20 itens. A cada 0,7 minutos um cliente entra na fila. O tempo médio de atendimento em cada caixa é de 5 minutos. Calcular os parâmetros desta fila. Para o cálculo da probabilidade do sistema estar no estado n, considerar n=5 e n=10.

3. Uma companhia aérea possui um balcão com quatro guichês no aeroporto de Guarulhos. Foram coletadas as quantidades de chegadas de passageiros por minuto durante o período de 1 hora, conforme mostra a Tabela 1:

Tabela 1: Número de chegadas por minuto.

3	4	4	2	2	1
0	1	5	1	1	3
3	2	0	0	0	2
2	5	0	0	2	3
1	0	1	1	2	5
0	2	4	2	5	1
2	1	3	2	5	4
1	2	0	1	0	0
0	1	1	4	0	0
1	4	0	1	0	1

Os tempos de atendimento (minutos) são mostrados na Tabela 2:

Tabela 2: tempos de atendimento (minutos)

2,5	2,5	4,05	2,15	1,2	0,65
3,5	0,95	0,6	1,1	0,65	3,25
1,4	1,8	4,05	2,35	2,15	0,7
7	2,2	0,6	5,15	0,65	1,15
0,9	0,95	1,2	0,75	1,85	1
0,75	1,35	2,05	0,5	2,15	6,15
1,55	5,05	2,8	4,2	1,1	1,65
1,95	1,15	1,15	0,9	4,05	3,15
0,8	0,85	5,05	2,8	0,95	2,35
4,1	0,65	1,1	0,75	2,1	1,05

Calcular o tamanho média da fila, tempo médio de espera e o índice de utilização dos guichês.

Referências Bibliográficas

CORRÊA, H. L.; CORRÊA, C. A. **Administração de produção e operações.** 4. ed. Atlas. São Paulo. 2019.

ERLANG.COM.BR. **Conceito Erlang B**. Disponível em: http://www.erlang.com.br/erlangb.asp. Acessado em: 27/07/2018. 2007a

ERLANG.COM.BR. **Conceito Erlang C**. Disponível em: http://www.erlang.com.br/erlangc.asp. Acessado em: 27/07/20178. 2007bW

FOGLIATTI, M. C.; MATTOS, N. M. C. **Teoria das filas.** Editora Interciência. Rio de Janeiro. 2007.

HILLIER, F.S.; LIEBERMAN, G. J. Introdução à pesquisa operacional. 9. ed. Bookman. Porto Alegre. 2013.

PRADO, D. Teoria das filas e simulação. 6. ed. Falconi Editora. Nova Lima. 2017.

RAGSDALE, C. T. **Modelagem de planilha e análise de decisão**: uma introdução prática a business analytics. Cengage Learning. São Paulo. 2014.

TAHA, H. A. Pesquisa operacional; 8. ed. Pearson Prentice Hall. São Paulo. 2008.

WIKIPEDIA. **Agner Krarup Erlang**. Disponível em: https://pt.wikipedia.org/wiki/Agner_Krarup_Erlang. Acessado em 27/07/2018. 2018.

WIKIPEDIA. **Andrei Markov**. Disponível em: https://pt.wikipedia.org/wiki/Andrei_Markov . Acessado em: 02/08/2018. 2018.