Otimização II

Prof. Dr. Paulo Roberto Maia
Paulo.maia@inatel.br

P108 – Otimização II

Agenda

- Modelo M/M/1
- Modelo M/M/s>1
- Modelo M/M/1/K
- Modelo M/M/s>1/K
- Modelo M/M/1/N
- Modelo M/M/s>1/N

Introdução

Conceito

- Teoria das filas ou teoria de congestão inicialmente motivada por aplicação em sistemas telefônicos, é o ramo da Pesquisa Operacional que estuda as relações entre as demandas em um sistema e os atrasos sofridos pelos usuários deste sistemas.

Introdução

Filas de espera aparecem em qualquer sistema de produção, particularmente em sistemas de serviços:

- Bancos
- Supermercados
- Correios
- Postos de gasolina

Sistemas de manufatura: produtos aguardando processamento em máquinas ou estação de trabalho.

Sistemas de transporte:

- aviões
- navios

Introdução

Sistemas computacionais:

- tarefas aguardando processamento em computadores
- pacotes de dados aguardando transmissão em rede.

As formações de filas ocorrem porque a procura pelo serviço é maior do que a capacidade do sistema de atender a esta procura. As possíveis razões para não aumentar a capacidade de atendimento são basicamente duas: inviabilidade econômica e limitação de espaço.

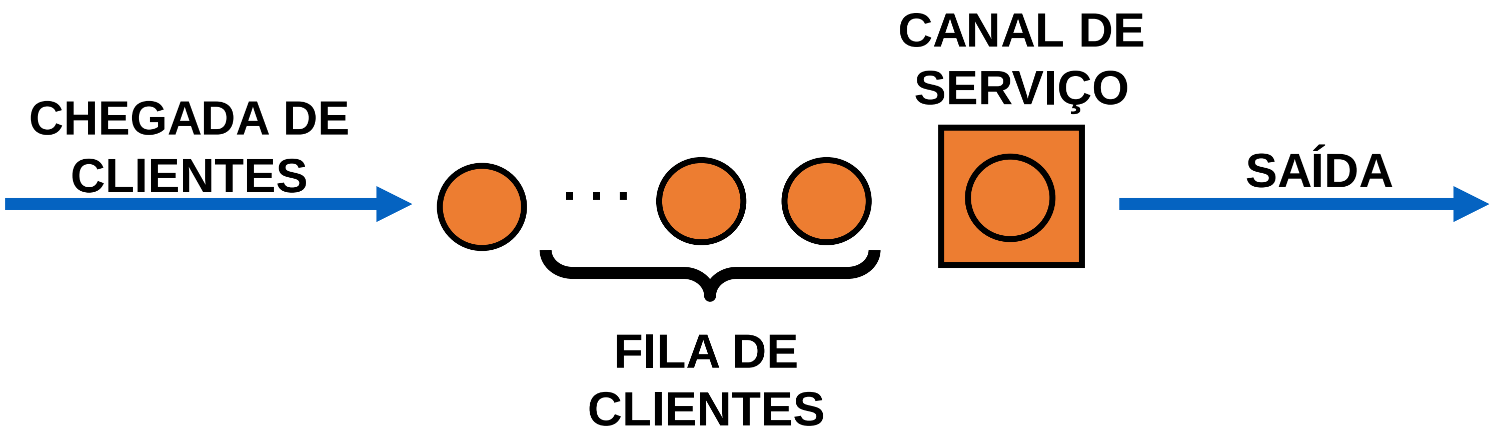
Introdução

Dessa forma, a Teoria das Filas tenta através de análises matemáticas detalhadas encontrar um ponto de equilíbrio que satisfaça o cliente e seja viável economicamente para o provedor do serviço.

A teoria das filas foi desenvolvida para prover modelos que retratem previamente o comportamento de um sistema que forneça serviços que possuam demandas que aumentem aleatoriamente.

Descrição do Problema de Filas

Um sistema de filas pode ser descrito como clientes chegando, esperando pelo serviço, se não forem atendidos imediatamente, e saindo do sistema após serem atendidos. O termo cliente é usado de maneira geral e não implica necessariamente num cliente humano.



Sistema de 1 fila e 1 canal

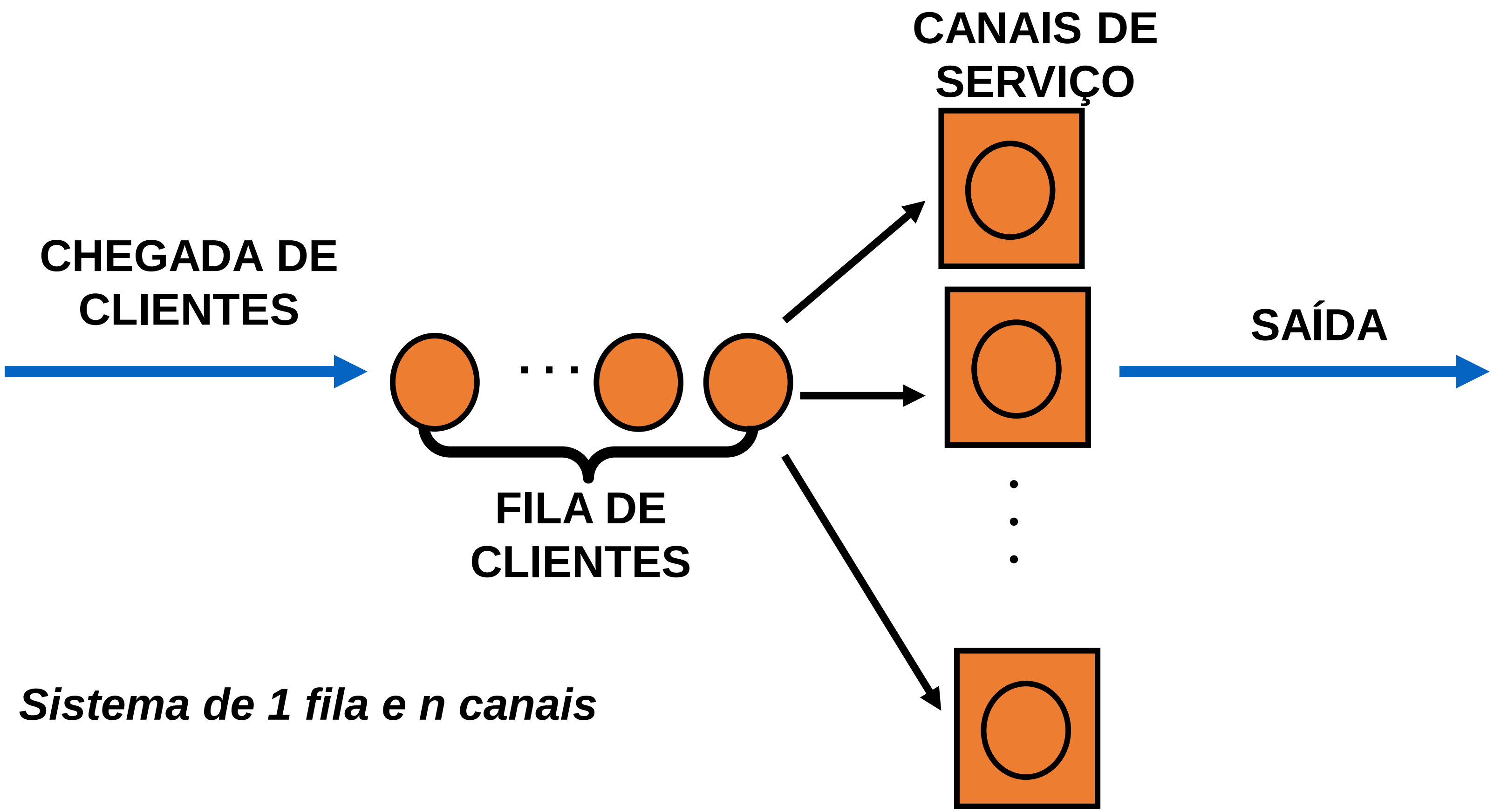
Descrição do Problema de Filas

Na maioria dos casos, seis características básicas de processos de filas fornecem uma descrição adequada de um sistema de filas:

(1) padrão de chegada dos clientes
(2) padrão de serviço dos servidores
(3) disciplina de filas
(4) capacidade do sistema
(5) número de canais de serviço
(6) número de estágio de serviços.

Características dos Processos de Filas

Processo de chegada, disciplina da fila e processo de serviço.



Sistema de 1 fila e n canais

Características dos Processos de Filas

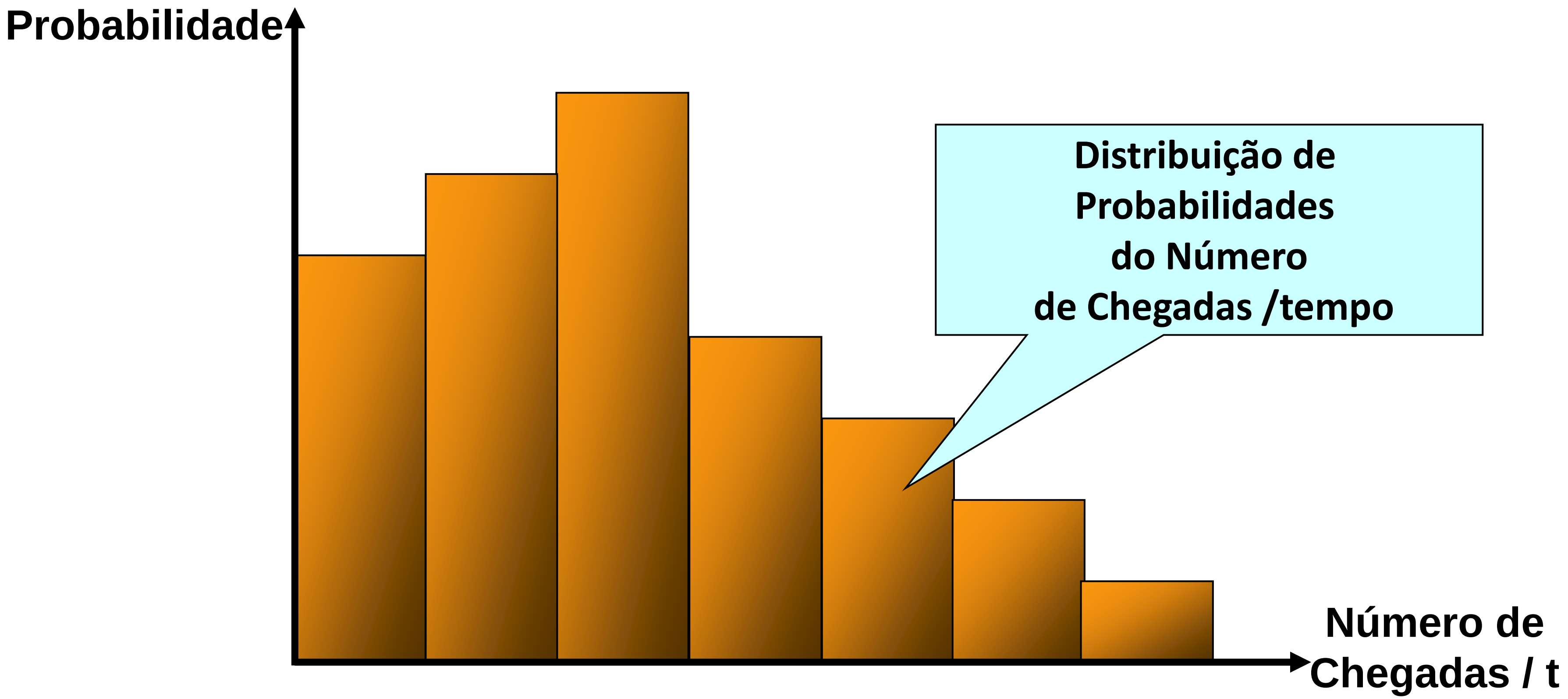
Padrão de chegada dos clientes

Nos processos de filas comuns, os processos de chegadas são estocásticos, ou seja, desenvolvem-se no tempo e no espaço conforme leis de probabilidade. Assim, é necessário conhecer a distribuição de probabilidade descrevendo os tempos entre as sucessivas chegadas dos clientes (tempos de interchegada). Também é necessário saber se os clientes podem chegar simultaneamente (chegada batch).

Características dos Processos de Filas

Padrão de chegada dos clientes

As chegadas de clientes a um sistema ocorrem de forma aleatória.



Características dos Processos de Filas

Padrão de serviço dos servidores

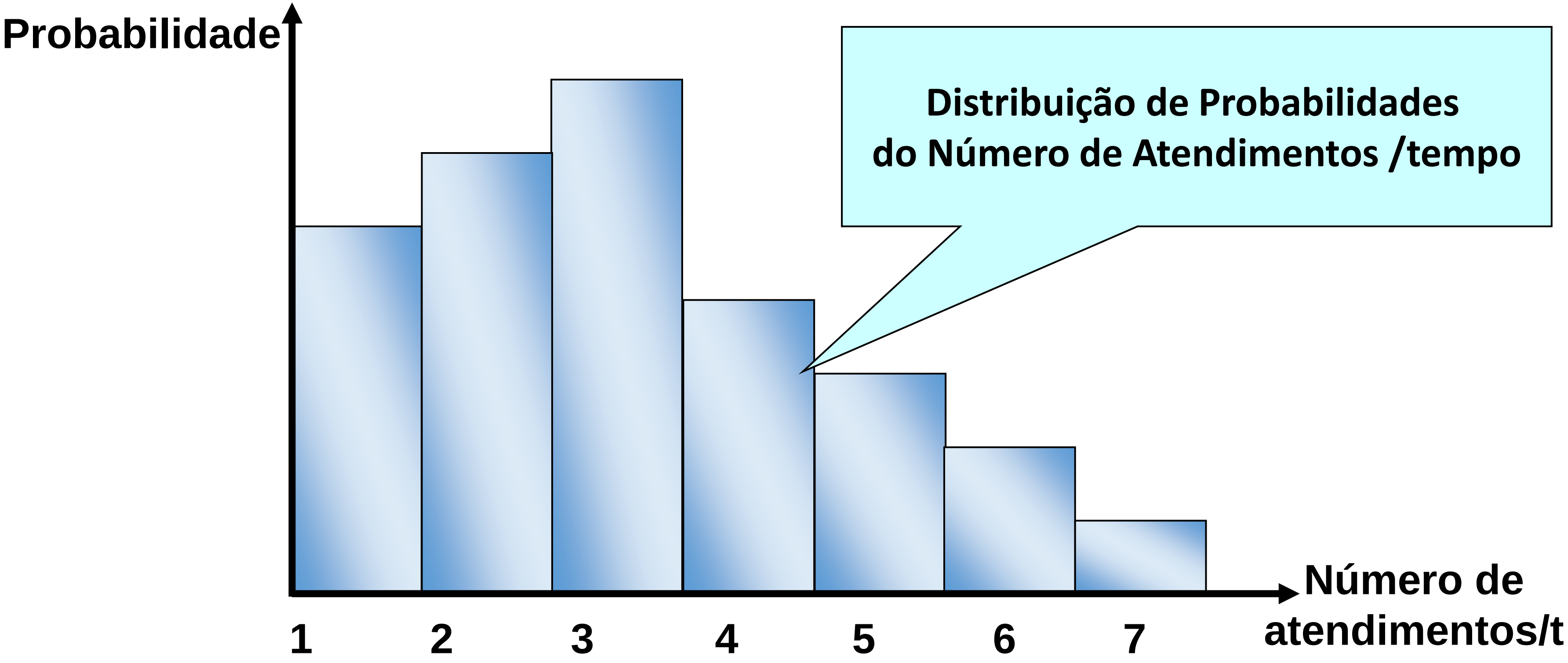
O processo de serviço pode depender do número de clientes esperando pelo serviço. Um servidor pode trabalhar mais rápido se a fila estiver aumentando, ou, caso contrário, pode ser tornar confuso é ficar mais lento. A situação na qual o serviço depende do número de clientes na fila é conhecida como serviço dependente do estado.

Características dos Processos de Filas

Padrão de atendimentos

O atendimento de clientes ocorre de forma aleatória.

LEVANTAMENTO ESTATÍSTICO: determinar a distribuição de probabilidades do número de atendimentos ou da duração de cada atendimento.



Características dos Processos de Filas

Disciplina de filas

A disciplina de filas refere-se a maneira como os clientes são escolhidos para entrar em serviço após uma fila ser formada. A maioria das disciplinas comuns que podem ser observadas na vida diária é FCFS (First-Come-First-Served), ou seja, o primeiro a chegar é o primeiro a ser servido. Entretanto, existem outras disciplinas, tais como, LCFS(Last-Come-First-Served), aplicável em sistemas de controle de estoque onde o item mais recente é mais fácil de ser apanhado, e diversas outras disciplinas baseadas em esquemas de prioridade.

Características dos Processos de Filas

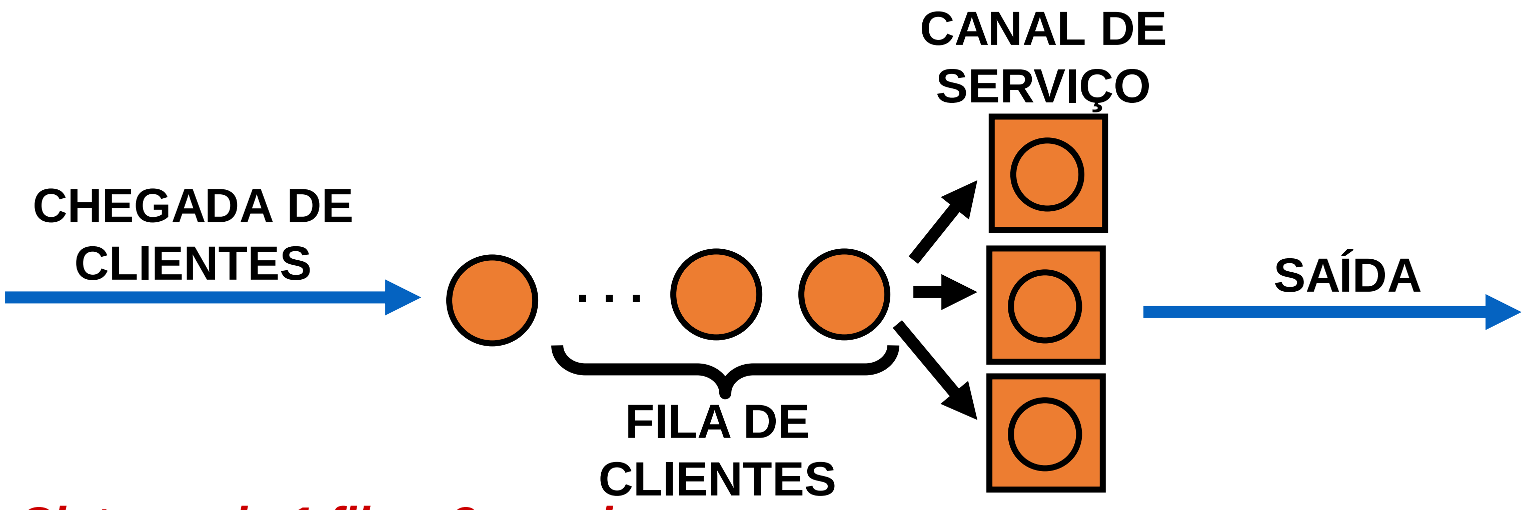
Capacidade do sistema

Em alguns processos de filas existe uma limitação física da quantidade de espaço na fila, de modo que, se as filas alcançarem um certo comprimento, nenhum novo cliente poderá entrar no sistema até que espaço disponível seja obtido com o atendimento de um cliente e a consequente diminuição do tamanho da fila. Estas situações são referidas como sistemas de filas finitos, ou seja, existe um limite finito do tamanho máximo do sistema.

Características dos Processos de Filas

Número de canais de serviço

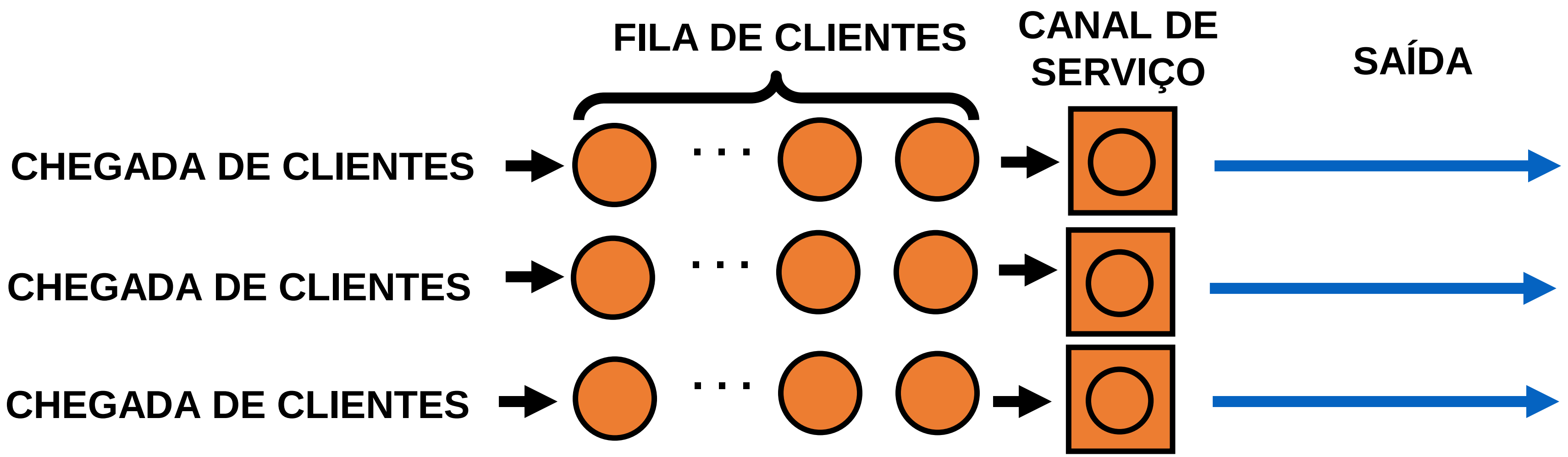
Quando o número de canais de serviço são definidos, tipicamente estão sendo determinados o número de estações de serviços paralelos que podem servir os clientes simultaneamente.



Sistema de 1 fila e 3 canais

Características dos Processos de Filas

Número de canais de serviço



Sistema de 3 fila e 3 canais

1- Sistema com canal simples

2- Sistema Multicanal com Fila única

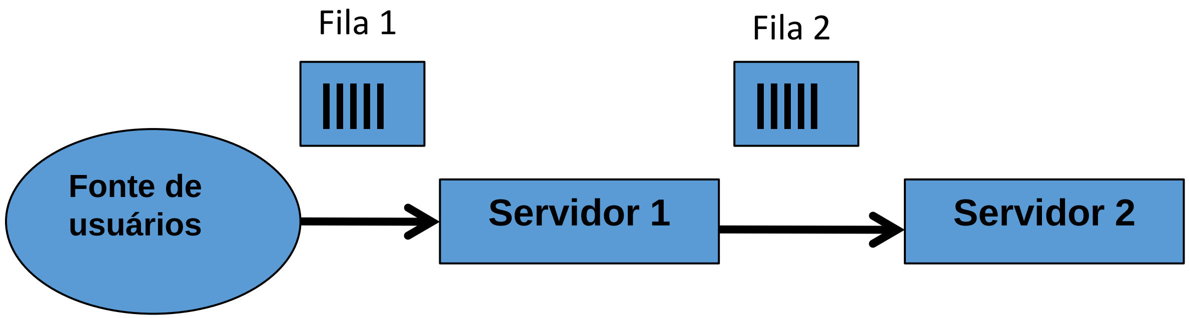
3- Sistema Multicanal com Fila individual

Descrição do Problema de Filas

Número de estágios de serviço

Um sistema de filas pode ter um único estágio de serviço, como no caso da barbearia, ou pode ter vários estágios. Um sistema de multi-estágio pode ser exemplificado como um procedimento de exame físico, onde cada paciente passa por diversos exames, tais como: sangue, vista, urina e etc.

Sistema de multi-estágio



Outro exemplo: reciclagem no processo de manufatura

Notação de Kendall-Lee

A notação de processos de filas mais utilizada atualmente foi proposta por Kendall, em 1953, e é descrita por um série de símbolos, tais como, A/B/m/C/k/N.

A primeira característica A refere-se à distribuição de probabilidade do processo de chegada dos clientes, podendo ser:

A=D: determinística, não é regido por fenômenos aleatórios.
A=M: distribuição exponencial, sem memória;
A= Ep: distribuição de Erlang com parâmetro de forma
A=G: distribuição genérica.

A segunda característica B refere-se ao o padrão de serviço de acordo com a distribuição de probabilidade para o tempo de serviço, com as mesmas variações de A

Notação de Kendall-Lee

A terceira característica m refere-se ao número de canais de serviços paralelos (servidores).

C especifica a disciplina da fila, por exemplo C= FCFS (first come, first served), LCFS (last come, first served) ex. elevador lotado ou SIRO (service in random order) ex. embarque metrô.

k indica o número máximo de usuários do sistema (incluindo usuários em serviço ou em fila) exemplo: K=10

N indica o tamanho da população, exemplo N=100. Podemos ter K= ∞ (sistema de capacidade ilimitada) e N= ∞ (sistema de população infinita).

Notação de Kendall-Lee

Caracteristicas	Símbolo	Explicação
Distribuição de Tempo de chegada (**A**) Distribuição de Tempo de Serviço (**B**)	M	Exponencial (markoviana)
	D	Degenerada (tempos constantes)
	Ek	Erlang (parâmetro de forma =k)
	G	Genérica ou GI (genérica independente)
Número de servidores em paralelo (**m**)		
Disciplina da Fila (**C**)	FCFS	First Come First Served
	LCFS	Last Come First Served
	SIRO	service in random order
Número máximo de usuários (**K**)	1,2,...,∞	
Tamanho da população (**N**)	1,2,...,∞	

Por exemplo, a notação M/G/3/FCFS/10/100 indica:

- Sistema de filas com intervalo de tempo entre chegadas exponencial;
- Tempo de serviço com distribuição genérica;
- Três servidores idênticos em paralelo;
- Disciplina da fila ao chegar FCFS
- Capacidade limitada a 10 (max. 3 servidores + 7 usuários)
- Tamanho da população finita = 100

Notação de Kendall-Lee

Use a notação de Kendall para classificar as seguintes situações:

- Corredor único para a lavagem automática de carros;
- Padaria padrão;
- Travessia do rio por balsas;
- Elevador;
- Bancos;
- Supermercados;
- Serviço de atendimento ao cliente
- Posto de gasolina;
- Call center
- Consultório médico

Notação de Kendall-Lee

De acordo com a notação de Kendall, decodifique o sistema de filas abaixo:

D	D	3	FIFO	10	50
G	Ep	15	SIRO	7	150
M	M	70	FIFO	∞	∞
Ep	D	1	LIFO	10	500

Parâmetros de sistemas de filas

Em muitos sistemas de filas, a chegada de usuários ocorre de maneira totalmente aleatória, ou seja, a chegada de um usuário não é influenciada pelo instante atual ou pelo tempo decorrido desde a última chegada ou o último término do serviço. Nestes casos, esses intervalos de tempos X totalmente aleatórios são descritos pela distribuição exponencial.

$$f_T(t) = \lambda e^{-\lambda t}, \qquad para\ t \geq 0$$

$$E(X) = \frac{1}{\lambda} e\ variância\ V(X) = 1/\lambda^2$$

Parâmetros de sistemas de filas

O número de chegada de usuários durante um intervalo de tempo de 0 a t tem distribuição de Poisson

$$p_n(t) = \frac{(\lambda t)^x e^{-\lambda t}}{n!}, \qquad para\ x = 1, 2, \dots$$

$$E\big(X(t)\big) = \lambda t\ e\ variância\ V\big(X(t)\big) = \lambda t$$

Medidas de efetividade de um sistema

- Percentagem do tempo ocioso ou ocupado
- Tempo médio que cada cliente gasta na fila de espera
- Tempo médio gasto pelo cliente no sistema
- Número médio de clientes na fila
- Número médio de clientes no sistema
- Probabilidade de existir um número n de clientes no sistema.

Parâmetros de sistemas de filas

Os usuários ou clientes chegam ao sistema a uma razão, que é determinada pela quantidade de usuários dividida pelo intervalo de tempo de observação.

A taxa de chegada é representada pela letra "lâmbda" e é calculada conforme fórmula abaixo:

$$\lambda = \frac{número\ de\ usuários\ que\ chegam}{intervalo\ de\ tempo}$$

Obs: Taxa de chegada de usuários no sistema é descrito pelo intervalo de tempo entre chegadas sucessivas de usuários. Normalmente os clientes chegam um a um ou lote.

Parâmetros de sistemas de filas

A frequência ou a velocidade com a qual os usuários são atendidos ou recebem o serviço é denominada de taxa de atendimento.

A taxa de atendimento é representada pela letra μ "mi" e é calculada pela expressão abaixo:

$$\mu = \frac{número\ de\ usuários\ atendidos}{intervalo\ de\ tempo}$$

Obs: Taxa de atendimento é descrito pelo tempo de serviço por usuário. Cada servidor não precisa ser um indivíduo, mas pode ser um grupo de pessoas ou máquinas realizando simultaneamente um serviço.

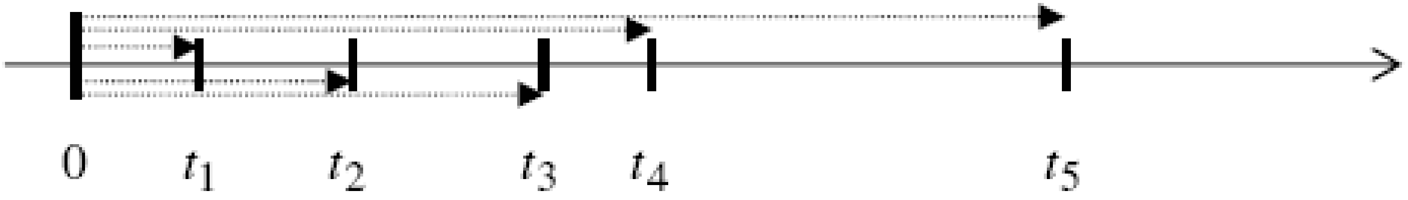
Parâmetros de sistemas de filas

Para tornar mais claras as definições dos parâmetros λ e μ, vamos resolver o seguinte exemplo.

Clientes chegam a um posto de atendimento e é observado que num intervalo de tempo de 5 minutos chegam 6 clientes. Qual é a taxa de chegada? Se no atendimento, em média, um usuário demanda 40 segundos, qual é a taxa de atendimento?

Parâmetros de sistemas de filas

A chegada de 5 usuários num sistema de filas hipotético. Suponhamos que foram anotados os instantes de chegada dos usuários, denotados por $t_i, com\ i = 1, 2, 3, 4\ e\ 5$, medidos a partir do instante zero.



1º usuário	$T_1 = t_1 - 0$
2º usuário	$T_2 = t_2 - t_1$
3º usuário	$T_3 = t_3 - t_2$
4º usuário	$T_4 = t_4 - t_3$
5º usuário	$T_5 = t_5 - t_4$

Parâmetros de sistemas de filas

$$\lambda = \frac{5}{T_1 + T_2 + T_3 + T_4 + T_5}$$

A taxa de chegada λ é o inverso da média dos tempos entre chegadas.

O TMC é o tempo médio entre chegadas (média dos tempos entre chegadas)

$$\lambda = \frac{1}{TMC}$$

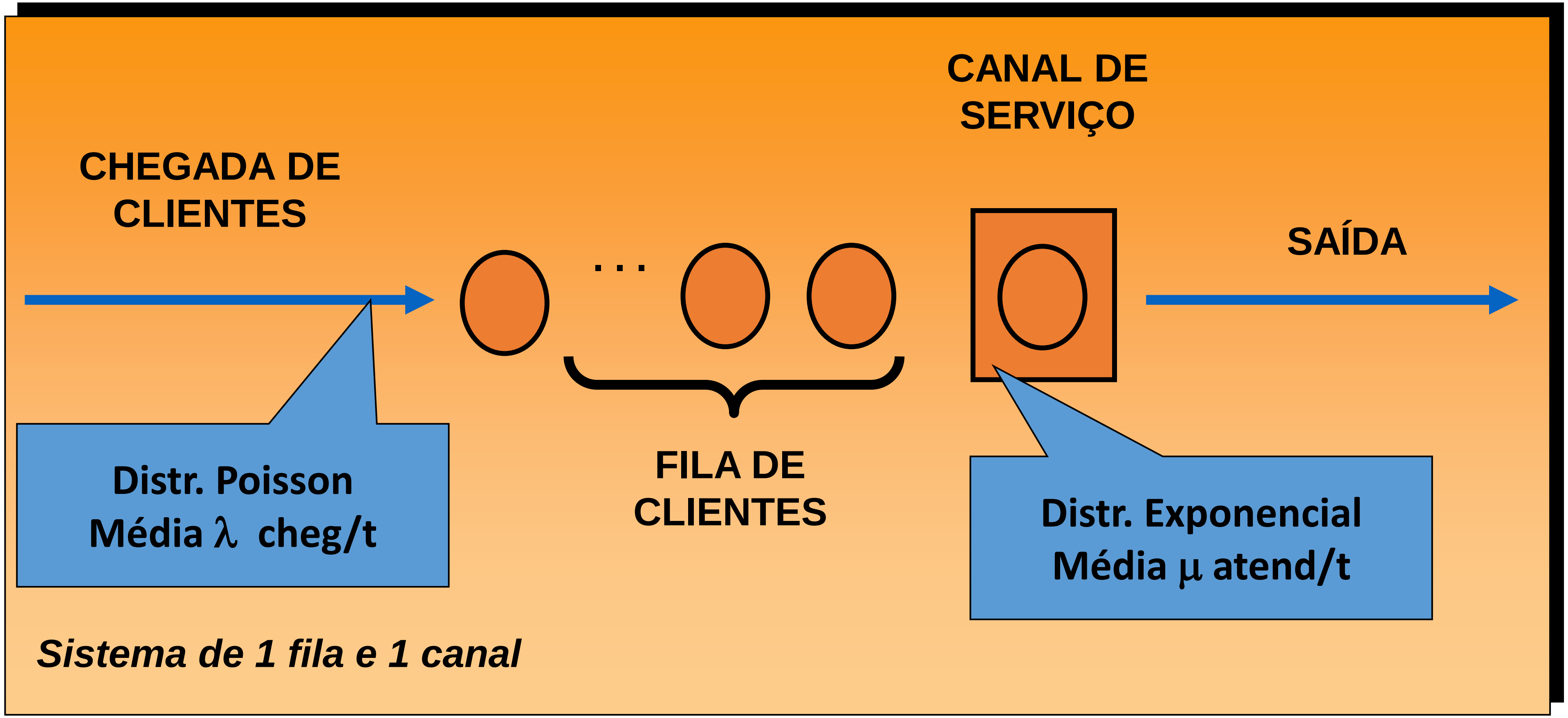
Parâmetros de sistemas de filas

Uma expressão análoga a expressão da taxa de chegada λ relaciona a taxa de atendimento μ e a média dos tempos de serviço, TMS.

$$\mu = \frac{1}{TMS}$$

Depois de levantados os dados λ e μ de um determinado sistema de filas, precisamos obter grandezas que representem medidas objetivas da situação operacional da fila em regime estacionário. Tais grandezas são designadas como medidas de efetividade dos sistemas de filas.

Modelo M/M/1


CANAL DE SERVIÇO
CHEGADA DE CLIENTES
SAÍDA
. . .
Distr. Poisson Média λ cheg/t
FILA DE CLIENTES
Distr. Exponencial Média μ atend/t
Sistema de 1 fila e 1 canal

Modelo M/M/1

Características gerais

- ✓ Chegadas: ocorrem segundo uma distribuição de Poisson com média λ chegadas/tempo.
- ✓ Tempos de atendimentos: seguem a distribuição exponencial com média $1/\mu$.
- ✓ Número de atendimentos: segue a distribuição de Poisson com média μ;
- ✓ O atendimento à fila é feito pela ordem de chegada.
- ✓ O número de possíveis clientes é suficientemente grande para que a população possa ser considerada infinita.
- ✓ Condição de estabilidade do sistema: $\lambda < \mu$.

Modelo M/M/1

Equação básica do sistema

Probabilidade de haver n clientes no sistema

$$P(n) = \left(\frac{\lambda}{\mu}\right)^n \cdot \left(\frac{\mu - \lambda}{\mu}\right) = (1 - \rho)\rho^n$$

$$P(0) = 1 - \rho$$

Probabilidade de que o número de clientes no sistema seja superior a um certo valor r:

$$P(n > r) = \left(\frac{\lambda}{\mu}\right)^{r+1}$$

Equação básica do sistema

Probabilidade de que sistema esteja ocioso:

$$P(n = 0) = \left(\frac{\mu - \lambda}{\mu}\right)$$

Probabilidade de que sistema esteja ocupado:

$$Taxa\ de\ ocupação(\rho) = P(n > 0) = \frac{\lambda}{\mu}$$

Modelo M/M/1

Equação básica do sistema

Probabilidade do tempo de espera no sistema (W) > t:

$$P(W > t) = e^{-\mu(1-\rho)t}, \qquad para\ t \geq 0$$

Probabilidade do tempo de espera na fila (W_q) > t:

$$P(W_q > t) = \rho e^{-\mu(1-\rho)t}, \qquad para\ t \geq 0$$

Medidas de efetividade

Número médio de clientes no sistema (L):

$$L = \frac{\lambda}{\mu - \lambda}$$

Número médio de clientes na fila (Lq):

$$L_q = \frac{\lambda^2}{\mu(\mu - \lambda)}$$

Modelo M/M/1

Medidas de efetividade

Tempo médio gasto no sistema por cliente(W):

$$W = \frac{1}{(\mu - \lambda)}$$

Tempo médio de espera na fila por cliente (W_q):

$$W_q = \frac{\lambda}{\mu(\mu - \lambda)}$$

Modelo M/M/1

Relações entre W_q, W, L_q e L

$$\boxed{L_q = \lambda \cdot W_q}$$

$$\boxed{L = \lambda \cdot W}$$

$$\boxed{L_q = L - \frac{\lambda}{\mu}}$$

$$\boxed{W_q = W - \left(\frac{1}{\mu}\right)}$$

Exemplo 1

Uma grande empresa siderúrgica dispõe de um conjunto numeroso de motores elétricos, instalados em seus galpões, que eventualmente apresentam defeitos oriundos de várias causas, inclusive choques mecânicos. Uma vez que o dano tenha sido constatado, a empresa imediatamente troca o equipamento por um reserva e recolhe o defeituoso para a seção de manutenção. Um levantamento estatístico do número de motores danificados por mês chegou aos dados da seguinte tabela:

Mês	Nº Equip. Danificados	Mês	Nº Equip. Danificados	Mês	Nº Equip. Danificados
1	9	11	9	21	11
2	8	12	12	22	12
3	11	13	11	23	11
4	10	14	13	24	13
5	8	15	10	25	9
6	11	16	13	26	11
7	13	17	10	27	13
8	13	18	12	28	10
9	12	19	11	29	10
10	12	20	12	30	10

Modelo M/M/1

Exemplo 1

A seção de manutenção da empresa opera atualmente quase em plena carga, tendo sido constatado em um levantamento que sua taxa de ocupação é de 88%, calcule:

a) A distribuição de frequência relativa para o número de equipamentos danificados.

b) Tempo médio que o motor permanece fora de serviço, em dias.

c) Número médio de motores aguardando o início dos trabalhos de reparo.

d) Tempo médio que o motor aguarda para início da manutenção.

e) Número médio de motores fora de serviço, em dias.

f) Probabilidade do sistema estar ocioso.

g) Probabilidade do sistema estar ocupado.

h) Probabilidade W>1.

i) Probabilidade Wq>1.

Modelo M/M/1

Exemplo 2

Clientes chegam a uma barbearia, de um único barbeiro, com tempo médio entre chegadas de 20 minutos. O barbeiro gasta em média 15 minutos com cada cliente.

a) Qual a probabilidade de um cliente não ter que esperar para ser servido? (0,25)

b) Qual o número esperado de clientes no salão de barbeiro? e na fila? (L = 3 clientes, Lq=2,25 clientes)

c) Quanto tempo em média, um cliente permanece no salão? (W=1 hora)

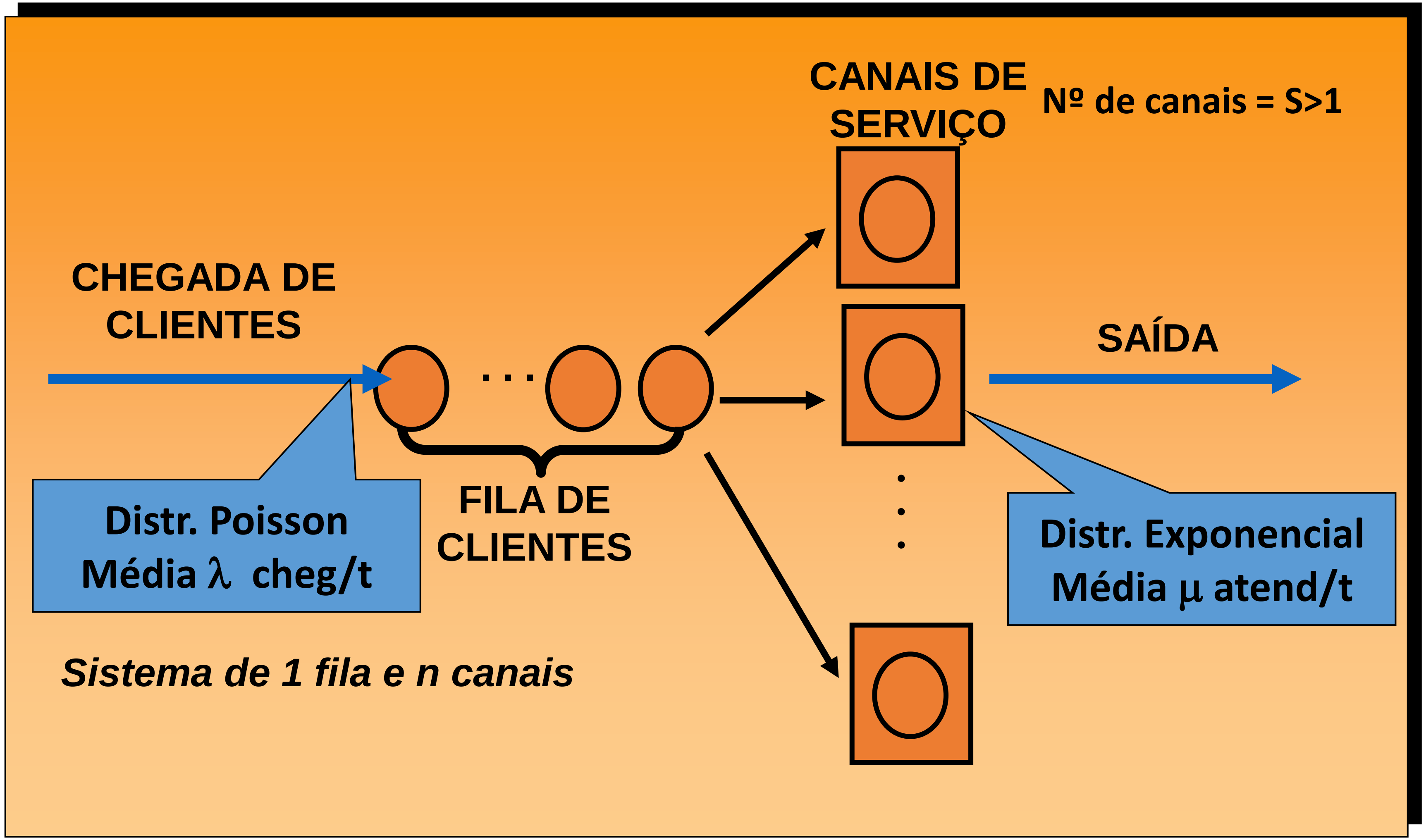
d) Quanto tempo em média, um cliente espera na fila? (Wq=45 minutos)

e) Probabilidade do cliente esperar mais de 1 hora para ser atendido.

f) Probabilidade do cliente ficar mais de 1 hora no salão.

e) O barbeiro está considerando a possibilidade de colocar um segundo barbeiro desde que o tempo de permanência médio de cada cliente no salão passe 1,25 horas. Para quanto teria que aumentar a taxa de chegada para que o segundo barbeiro ficasse justificado? (λ=3,2 clientes/hora)

Modelo M/M/s>1



CANAIS DE SERVIÇO
Nº de canais = S>1
CHEGADA DE CLIENTES
SAÍDA
Distr. Poisson
Média λ cheg/t
FILA DE CLIENTES
Distr. Exponencial
Média μ atend/t
Sistema de 1 fila e n canais

Modelo M/M/s>1

Características gerais

- ✓ Chegadas: ocorrem segundo uma distribuição de Poisson com média λ chegadas/tempo.
- ✓ Tempos de atendimentos: seguem a distribuição exponencial com média 1/µ.
- ✓ Número de atendimentos: segue a distribuição de Poisson com média µ;
- ✓ O atendimento à fila é feito pela ordem de chegada.
- ✓ Número de canais de serviço: S
- ✓ O número de possíveis clientes é suficientemente grande para que a população possa ser considerada infinita.
- ✓ Ritmo de serviço: µ.S
- ✓ Condição de estabilidade do sistema: λ < µ.S ou $\rho = \frac{\lambda}{s\mu} < 1$

Modelo M/M/s>1

Equações básicas do modelo

Probabilidade de haver 0 cliente no sistema

$$P_0 = 1 \Bigg/ \left[\sum_{n=0}^{S-1} \frac{\left({}^{\lambda}\!/_{\mu}\right)^n}{n!} + \frac{\left({}^{\lambda}\!/_{\mu}\right)^S}{s!} \times \frac{1}{1 - \lambda/(s\mu)} \right]$$

Probabilidade de haver n clientes no sistema:

$n \leq S$

$$P_n = \frac{(\lambda/\mu)^n}{n!} \times P_0$$

$n \geq S$

$$P_n = \frac{(\lambda/\mu)^n}{s!\, s^{n-S}} \times P_0$$

Modelo M/M/s>1

Equações básicas do modelo

Probabilidade do tempo de espera no sistema (W) > t:

$$P(W > t) = e^{-\mu t}\left[1 + \frac{P_0(\lambda/\mu)^s}{s!\,(1-\rho)}\left(\frac{1 - e^{-\mu t(s-1-\lambda/\mu)}}{s - 1 - \lambda/\mu}\right)\right]$$

Probabilidade do tempo de espera na fila (W_q) > t:

$$P\left(W_q > t\right) = \left[1 - P\left(W_q = 0\right)\right]e^{-s\mu(1-\rho)t}\text{, } em\ que$$

$$P\left(W_q = 0\right) = \sum_{n=0}^{s-1} P_n$$

Modelo M/M/s>1

Medidas de efetividade

Número médio de clientes na fila: $L_q = \frac{P_0(\lambda/\mu)^s \rho}{s!\,(1-\rho)^2}$

Tempo médio de espera na fila: $W_q = \frac{L_q}{\lambda}$

Número médio de clientes no sistema: $L = L_q + \frac{\lambda}{\mu}$

Tempo médio gasto no sistema: $W = \frac{L}{\lambda} = W_q + \frac{1}{\mu}$

Modelo M/M/s>1

Exemplo 1

Um sistema de uma fila e um canal está com sobrecarga de trabalho. Após análise estatística, o analista de pesquisa operacional descobriu que a média de chegadas de clientes ao sistema é de 30 por hora. O atendente tem capacidade para atender somente 20 clientes por hora. Dessa forma, ele está planejando criar mais uma seção de atendimento igual à primeira e passar a operar com dois canais e uma fila somente. Para essa nova situação, calcule:

a) Número médio de clientes na fila.(Lq)

b) Tempo médio que o cliente fica na fila.(Wq)

c) Número médio de clientes no sistema.(L)

d) Tempo médio que o cliente fica no sistema.(W)

e) Probabilidade W>12 min.

f) Probabilidade Wq>6 min.

g) Probabilidade de haver 3 clientes no sistema.

Modelo M/M/s>1

Exemplo 2

O fechamento de várias agências bancárias ao redor de uma determinada agência provoca a necessidade de contratar novos servidores, pois a taxa de chegadas de clientes é agora de $\lambda = 0{,}60$ clientes por minuto. O tempo de atendimento é prestado na ordem de chegada, com média igual a 2 minutos. O gerente decide colocar mais dois caixas. Considerando essa situação determine:

a) A taxa de ocupação do sistema. (0,4)

b) O número médio de clientes no sistema. (1,29)

c) O número médio de clientes na fila. (0,094)

d) A probabilidade de o sistema estar vazio. (0,294)

e) A probabilidade de ter mais do que cinco clientes no sistema. (0,0091)

f) O tempo médio de espera de um cliente na fila. (0,157)

g) O tempo médio de permanência de um cliente no sistema. (2,16)

h) Probabilidade W>3 min. (0,2506)

i) Probabilidade Wq>2 min. (0,0233)

Referencial Bibliográfico

- Frederick S. Hillier, Gerald J. Lieberman; Introdução à Pesquisa Operacional; 9ª Edição, Editora Mc Graw Hill; 2013.

- Marcos Arenales, Vinícius Armentano, Reinaldo Morabito, HoracioYanasse; Pesquisa Operacional; 6ª Edição, Editora Campus, 2007.

- Eduardo L. de Andrade, Introdução a Pesquisa Operacional; 4ª Edição; Editora LTC; 2009.

- Wagner, H.M., Pesquisa Operacional, 2a edição. Prentice-Hall do Brasil, 1986.

- Taha, H. A., Pesquisa Operacional, 8a edição. Pearson (Prentice-Hall), 2008

Sobre a disciplina

Dúvidas?

Obrigado!