

ANÁLISE DE DESEMPENHO DE SISTEMAS DE FILAS COM SIMULAÇÃO BASEADA EM EVENTOS DISCRETOS

Rafael M. Andrade, rafaelmarizandrade@ufmg.br¹

Gustavo M. Andrade, gustavo.andrade.1380099@sga.pucminas.br²

Eriky S. Gomes, eriky-tn@ufmg.br¹

Frederico R. B. Cruz, fcruz@ufmg.br¹

¹Universidade Federal de Minas Gerais, 31270-901 - Belo Horizonte - MG,

²Pontifícia Universidade Católica de Minas Gerais, 30535-901 - Belo Horizonte - MG

Resumo. Modelos de filas são amplamente utilizados em aplicações da engenharia mecânica, especialmente na análise de sistemas sujeitos a demandas por serviços com variabilidade elevada. Este trabalho apresenta o desenvolvimento de um simulador de eventos discretos para análise de desempenho de sistemas modelados por filas, com foco inicial em modelos markovianos. A simulação por eventos discretos foi implementada na linguagem R, um ambiente que permite uma abordagem eficiente e flexível. As métricas avaliadas incluem a intensidade de tráfego (ρ), o tamanho médio da fila de espera (L_q) e o tempo médio de espera no sistema (W_q). Os resultados simulados foram comparados com valores analíticos conhecidos, demonstrando boa concordância e validando a proposta. A principal contribuição está na versatilidade da ferramenta, que pode ser facilmente adaptada para modelar sistemas mais complexos e não markovianos, para os quais não há solução analítica. A possibilidade de incorporar novas métricas e estruturas de fila amplia seu potencial de aplicação.

Palavras chave: Simulação a eventos discretos. Teoria de filas. Filas markovianas. Desempenho. Programação em R.

Abstract. Queueing models are widely used in Mechanical Engineering applications, particularly for analyzing systems subject to high variability in service demand. This work presents the development of a discrete-event simulator for performance analysis of systems modeled by queues, with an initial focus on Markovian models. The simulation was implemented in the R programming language, offering an efficient and flexible environment. The performance metrics analyzed include traffic intensity (ρ), average queue length (L_q), and average waiting time in the system (W_q). Simulated results were compared with known analytical values, showing good agreement and validating the proposed simulator. The main contribution lies in the versatility of the tool, which can be easily adapted to model more complex and non-Markovian systems, for which no analytical solutions exist. The ability to incorporate new performance metrics and queue structures further extends its applicability.

Keywords: Discrete-event simulation, queueing theory, Markovian queues, performance, R programming.

1. INTRODUÇÃO

Modelos de filas estão presentes em diversas situações práticas da engenharia mecânica, desempenhando um papel fundamental tanto em áreas mais evidentes, como a análise de sistemas de manufatura (Govil & Fu, 1999), quanto em contextos menos imediatos, como a gestão de grandes estruturas organizacionais, incluindo centros de dados (Kooile & Mandelbaum, 2002), portos e aeroportos (Tas *et al.*, 2013) e instituições de saúde (Almehdawe *et al.*, 2016). Esses modelos permitem analisar o comportamento de sistemas sujeitos a demandas por serviços com alta variabilidade, auxiliando na identificação de gargalos e na proposição de estratégias para mitigá-los, por meio de um melhor dimensionamento de recursos (Shurtle *et al.*, 2018).

Vamos tomar emprestado um exemplo extraído de Anderson *et al.* (2013), que descreve a situação de uma empresa de manufatura que utiliza um conjunto de seis máquinas idênticas, cada uma operando, em média, por 20 horas entre falhas. Assim, a taxa de chegada de pedidos de manutenção para cada máquina é $\lambda = 1/20 = 0,05$ por hora. Como as falhas ocorrem de forma aleatória, o processo de chegada das solicitações de reparo pode ser modelado por uma distribuição de Poisson. Um engenheiro de manutenção do departamento é o único responsável pelos reparos dessas seis máquinas. Os tempos de reparo seguem uma distribuição exponencial com média de duas horas por máquina, ou seja, a taxa de serviço é $\mu = 1/2 = 0,50$ máquinas por hora. Com essas taxas de chegada e de serviço, 0,05 e 0,50, respectivamente, podemos utilizar equações conhecidas para calcular as características operacionais do sistema (veja, por exemplo, Anderson *et al.*, 2013). Tais características fornecem ao gestor informações valiosas sobre o desempenho do sistema de manutenção. De fato, pode-se mostrar que, nesse caso, uma máquina com defeito aguarda, em média, 1,279 horas até o início do reparo,

e que 51,55% das falhas precisam aguardar na fila por atendimento. Esses dados indicam que talvez seja necessário adotar um sistema com dois engenheiros de manutenção para melhorar o serviço. Assim, uma contribuição importante dos modelos de filas de espera é justamente sua capacidade de revelar características operacionais indesejáveis, mesmo quando a taxa de serviço é maior que a taxa média de chegada ($\mu > \lambda$), o que à primeira vista pode parecer satisfatório. Este exemplo ilustra bem como métricas como o tempo médio de espera e a proporção de ocupação podem apontar para a necessidade de reconfiguração do sistema, ainda que sua estrutura aparente estar adequada. Esse é o tipo de problema que os modelos de filas discutidos a seguir ajudam a diagnosticar e resolver.

Diversos estudos têm se dedicado ao desenvolvimento de equações analíticas para uma ampla gama de sistemas de filas. Além do livro de Anderson *et al.* (2013), um material particularmente abrangente é o livro de Shortle *et al.* (2018), que vem sendo atualizado periodicamente para incorporar os resultados mais recentes da área. Essas equações podem ser úteis na análise de desempenho de sistemas, embora haja críticas quanto às simplificações envolvidas nos modelos matemáticos, como a suposição de independência entre chegadas e o pressuposto de que o sistema opera em regime permanente, que são condições raramente observadas na prática, em que são comuns chegadas em grupos e dependência temporal. Ainda assim, os modelos básicos de filas têm seu valor, funcionando como uma primeira aproximação, para análises posteriores mais realistas.

Buscando reduzir o grau de simplificação, surgem modelos mais sofisticados para os quais o tratamento analítico ainda não está disponível. Nesse contexto, ganham destaque as abordagens computacionalmente intensivas, especialmente os modelos baseados em simulação, que permitem maior flexibilidade ao abandonar premissas como independência, estacionariedade, falta de memória e chegadas unitárias. Também nessa vertente, há contribuições importantes, como a biblioteca `queuecomputer` (Ebert *et al.*, 2020), desenvolvida para o ambiente R, amplamente utilizado em estatística computacional e análise de dados (R Core Team, 2025).

Por que então propor mais um simulador de filas, como neste trabalho? A motivação está na oportunidade de revisar e aplicar conceitos fundamentais de computação científica, em especial, simulação a eventos discretos e estatística computacional via simulação Monte Carlo, ao mesmo tempo em que se desenvolve uma ferramenta aberta, com fácil possibilidade de extensão e adaptação.

O restante deste trabalho está organizado da seguinte forma. A Seção 2 apresenta a metodologia adotada, com destaque para as equações que descrevem as principais medidas de desempenho em modelos de filas mais simples. Em seguida, são discutidos na Seção 3 os resultados obtidos a partir das simulações realizadas, com uma análise crítica de seu comportamento frente aos valores teóricos esperados. Por fim, a Seção 4 reúne as conclusões do estudo, bem como observações finais e sugestões de tópicos que podem orientar pesquisas futuras na área.

2. METODOLOGIA

Na teoria das filas estudam-se diversos modelos básicos de filas, que são aproximações bastante razoáveis para situações reais. Um aspecto muito importante da teoria de filas é a simulação computacional, que se apresenta como uma ferramenta para analisar o desempenho de um determinado sistema de filas. Neste trabalho, desenvolve-se um sistema de simulação a eventos discretos para análise de desempenho em sistemas de filas. A abordagem por eventos discretos é reconhecida por sua eficiência computacional (Fishman, 2001), pois simula apenas os momentos em que há mudança de estado no sistema (por exemplo, chegada ou saída de clientes), ignorando períodos em que nada ocorre.

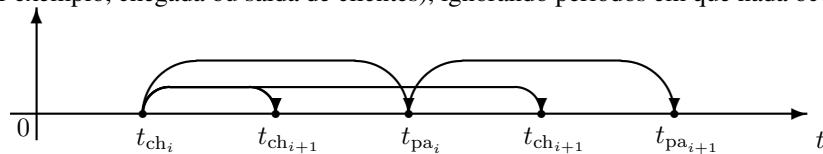


Figura 1. Instantes relevantes em uma simulação a eventos discretos de um sistema de filas

A Fig. 1 ilustra os instantes relevantes em uma simulação a eventos discretos aplicada a um sistema de filas. O tempo de chegada da i -ésima entidade arbitrária é denotado por t_{ch_i} . Nas chegadas, dois eventos futuros são determinados: o tempo de partida dessa entidade, t_{pa_i} , e o tempo da próxima chegada, $t_{ch_{i+1}}$. Ambos são definidos com base em números aleatórios gerados segundo distribuições previamente especificadas no modelo. A posição relativa entre t_{pa_i} e $t_{ch_{i+1}}$ é incerta, dependendo dos valores sorteados. Quando $t_{ch_{i+1}} < t_{pa_i}$, ocorre uma nova chegada antes da partida da entidade atual, o que implica um aumento no número de usuários presentes no sistema. Por outro lado, se $t_{pa_i} < t_{ch_{i+1}}$, a partida antecede a nova chegada, reduzindo momentaneamente o número de entidades no sistema. Essa dinâmica entre chegadas e partidas é fundamental para caracterizar o comportamento do sistema ao longo do tempo. Finalmente, quando a simulação se depara com um evento de partida, apenas um novo evento discreto é gerado, e isso apenas se houver usuários na fila de espera. Nesse caso, o próximo evento de partida é agendado no tempo $t_{pa_{i+1}}$, também definido com o auxílio de um gerador de números aleatórios apropriado.

O algoritmo proposto é apresentado em pseudocódigo na Fig. 2 e pode ser implementado em qualquer linguagem de programação. As operações de inserção e remoção no algoritmo podem ser facilmente implementadas usando estruturas

```

algorithm
(01) read  $t_{\max}, \rho$ ;
(02)  $\text{users} \leftarrow 0$ ;  $\lambda \leftarrow 1.0$ ;  $\mu \leftarrow \lambda/\rho$ ;
    /* generate last and first last events, enlist them, and run simulation */
(03)  $\text{event.time} \leftarrow t_{\max}$ ;
(04)  $\text{insert}('end', \text{event.time})$ ;
(05)  $\text{event.time} \sim \text{GI}(\text{rate} = \lambda)$ ;
(06)  $\text{insert}('arrival', \text{event.time})$ ;
(07) while ( $\text{event.type} \neq 'end'$ ) do
(08)   if ( $\text{event.type} = 'arrival'$ ) then
(09)      $\text{users} \leftarrow \text{users} + 1$ ;
    /* if server is available, generate departure, and enlist */
(10)     if ( $\text{users} = 1$ ) then
(11)        $\delta \sim \text{G}(\text{rate} = \mu)$ ;
(12)        $\text{insert}('departure', \text{event.time} + \delta)$ ;
(13)     end if
    /* generate next arrival and enlist it */
(14)      $\alpha \sim \text{GI}(\text{rate} = \lambda)$ ;
(15)      $\text{insert}('arrival', \text{event.time} + \alpha)$ ;
    /* remove earliest event from list */
(16)      $\text{remove}(\text{event.type}, \text{event.time})$ ;
(17)   else if ( $\text{event.type} = 'departure'$ ) then
(18)      $\text{users} \leftarrow \text{users} - 1$ ;
    /* if queue is not empty, generate departure, and enlist */
(19)     if ( $\text{users} \neq 0$ ) then
(20)        $\delta \sim \text{G}(\text{rate} = \mu)$ ;
(21)        $\text{insert}('departure', \text{event.time} + \delta)$ ;
(22)     end if
    /* remove earliest event from list */
(23)      $\text{remove}(\text{event.type}, \text{event.time})$ ;
(24)   end if
(25) end while
end algorithm

```

Figura 2. Algoritmo para simulação de um sistema de filas $GI/G/1/\infty$ por eventos discretos

de dados simples, como listas. Conforme ilustrado na Fig. 2, para simular as filas, são gerados os tempos entre chegadas (linhas 05 e 14), segundo qualquer distribuição geral e independente (GI) e os tempos de serviço (linhas 11 e 20), também de acordo com qualquer distribuição geral (G).

Vale destacar que o algoritmo descrito neste artigo pode servir como uma estrutura versátil para simulação de diferentes tipos de filas. Isso pode ser obtido apenas modificando-se as distribuições de probabilidade utilizadas para gerar os tempos entre chegadas e de serviço. Esse algoritmo já foi testado com sucesso em outros estudos, para simular coleta de dados e a boa qualidade das amostras geradas foi confirmada por meio vários experimentos (Cruz & Singh, 2021).

2.1 Filas $M/M/1$

Em particular, pode-se simular o modelo clássico de fila com chegadas segundo um processo de Poisson (markoviano), tempos de serviço exponenciais (também markoviano) e um único servidor, modelo conhecido, na notação de Kendall, como $M/M/1$. Trata-se de um dos modelos mais simples e fundamentais da teoria de filas, ilustrado na Figura 3. Nesse modelo, o número k de clientes que chegam por unidade de tempo segue uma distribuição de Poisson, cuja conhecida função de probabilidade é dada por:

$$p(k) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad k = 0, 1, 2, \dots, \quad (1)$$

em que λ é a taxa de chegada (ou seja, λ é o número médio de chegadas por unidade de tempo), enquanto os tempos de serviço t seguem uma distribuição exponencial, com a conhecida função densidade de probabilidade:

$$g(t) = \mu e^{-\mu t}, \quad t \geq 0, \quad (2)$$

em que λ representa a taxa de serviço (isto é, $1/\mu$ é o tempo médio de serviço). A partir desses parâmetros, define-se a intensidade de tráfego do sistema como $\rho = \lambda/\mu$.

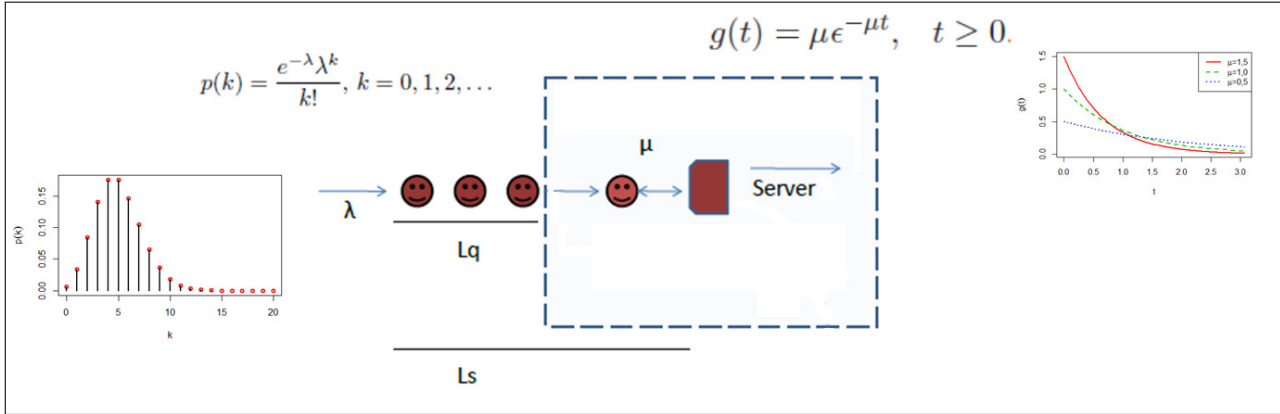


Figura 3. Representação de uma fila $M/M/1$

Sob a suposição de que a fila é observada por um longo período de tempo, até atingir um regime estacionário, e considerando que a intensidade de tráfego satisfaz $\rho < 1$, condição necessária para evitar o crescimento indefinido do número de clientes em espera, pode-se demonstrar que a distribuição estacionária do número de clientes (N) no sistema é dada por (Shortle *et al.*, 2018):

$$P(N = n) \equiv p_n = (1 - \rho)\rho^n, \quad (3)$$

em que p_n representa a probabilidade de se encontrar n clientes no sistema em um dado instante. A probabilidade de o sistema estar vazio é dada por $P(N = 0) \equiv (1 - \rho)$, o que reforça a interpretação de que a intensidade de tráfego (ρ) corresponde à proporção de tempo em que o servidor permanece ocupado. A partir de ρ , também se podem definir importantes medidas de desempenho, como o tamanho médio da fila de espera (L_q) e o tempo médio de espera na fila (W_q), antes de receber o serviço. Essas medidas são dadas, respectivamente, por:

$$L_q = \sum_{n=2}^{\infty} (n-1)p_n = \frac{\rho^2}{1-\rho}, \quad (4)$$

$$W_q = \frac{L_q}{\lambda} = \frac{\rho}{\lambda(1-\rho)}, \quad (5)$$

em que a Eq. (5) é também conhecida como a Lei de Little (veja detalhes em Little, 2011).

2.2 Filas $M/D/1$

As filas $M/D/1$ são um importante modelo de filas em sistemas industriais de fabricação. Nesse sistema de filas as chegadas seguem um processo de Poisson com taxa λ e os serviços são finalizados em um período fixo $1/\mu$. Pode-se mostrar que as medidas de desempenho são dadas por (Shortle *et al.*, 2018):

$$\rho = \frac{\lambda}{\mu}, \quad (6)$$

$$L_q = \frac{1}{2} \frac{\rho^2}{1-\rho}, \quad (7)$$

$$W_q = \frac{1}{2} \frac{\rho^2}{\lambda(1-\rho)}. \quad (8)$$

Se considerarmos o tempo médio de espera na fila ou o comprimento médio da fila como medidas de performance inversamente proporcionais à eficiência do sistema, pode-se notar que o sistema $M/D/1$ é duas vezes mais eficiente que o sistema $M/M/1$. Isto é, conforme esperado, a redução da variabilidade de serviço impacta positivamente a oferta de serviço.

2.3 Filas $M^{[X]}/M/1$

O modelo $M^{[X]}/M/1$ é uma importante generalização das filas $M/M/1$. Nesse modelo markoviano, as entradas ocorrem de acordo com um processo Poisson de taxa λ , mas o número de chegadas em cada entrada segue uma variável aleatória X , ao invés de corresponder a uma única chegada como no modelo $M/M/1$. Quanto aos tempos de serviço, eles também seguem uma distribuição exponencial de taxa μ .

Nesse sistema, não há uma expressão matemática geral para a probabilidade de se encontrar n clientes no sistema em um dado instante do regime estacionário. Ainda assim, as equações para as importantes medidas de desempenho ρ , L_q e W_q são conhecidas (Shortle *et al.*, 2018) e dadas, respectivamente, por:

$$\rho = \frac{\lambda \mathbb{E}[X]}{\mu}, \quad (9)$$

$$L_q = \frac{\rho(1 + \mathbb{E}[X^2] / \mathbb{E}[X])}{2(1 - \rho)} - \rho, \quad (10)$$

$$W_q = \frac{L_q}{\lambda \mathbb{E}[X]}, \quad (11)$$

em que se supõe que a intensidade de tráfego é $\rho < 1$, condição necessária para a estabilidade da fila.

3. RESULTADOS E DISCUSSÃO

O modelo foi implementado na linguagem R (R Core Team, 2025), uma plataforma amplamente utilizada em ciência de dados, o que permite fácil customização e incorporação de instrumentação para coleta de informações relevantes ao analista. Devido a restrições de espaço, o código desenvolvido para o R não foi incluído neste manuscrito. No entanto, ele está disponível mediante solicitação, diretamente com os autores.

Foram coletados dados simulados para cálculo de métricas clássicas de desempenho, mais especificamente a intensidade de tráfego ρ (proporção de tempo em que os servidores estão ocupados), o tamanho médio da fila de espera L_q e o tempo médio de espera na fila W_q . Inicialmente, analisou-se a convergência do algoritmo proposto tomando como base os resultados analíticos conhecidos para filas $M/M/1$, $M/D/1$ e $M^{[X]}/M/1$.

As Fig. 4 a 6 apresentam a convergência em função do tempo de simulação (em número de unidades da taxa de chegada), para as medidas de desempenho consideradas, para diversos valores no espaço paramétrico, e demonstram, de forma consistente, a aproximação dos resultados empíricos aos valores teóricos esperados. Para todos os modelos testados, observa-se clara concordância entre as métricas obtidas via simulação e aquelas derivadas analiticamente, mesmo para tamanhos amostrais moderados. Esses resultados indicam que a ferramenta desenvolvida é capaz de reproduzir com precisão o comportamento dos sistemas de filas modelados, validando seu uso em contextos similares. Além disso, os resultados sugerem que o simulador possui potencial para aplicação em cenários mais gerais, com diferentes configurações de chegada e atendimento.

Em seguida, investigamos o impacto do quadrado do coeficiente de variação do tempo de serviço T (definido por $\text{cv}^2[T] = \mathbb{V}[T] / \mathbb{E}[T]^2$) nas métricas de desempenho. Os modelos de filas $M/G/1$ simulados consideram que as chegadas são um processo de Poisson com taxa unitária ($\lambda = 1$) e tempos de partida seguindo distribuições gama com os quadrados dos coeficientes de variação pré-especificados em $\text{cv}^2[T] \in \{0, 1, 2\}$, que refletem os modelos de filas considerados, respectivamente, $M/D/1$, $M/M/1$ e $M/G/1$. O parâmetro $t_{\max}$ foi fixado em 20.000, o que implica em aproximadamente 20.000 entidades simuladas em cada iteração. Consideraram-se 100 replicações Monte Carlo do algoritmo para o cálculo das métricas de desempenho. Os resultados estão apresentados na Tab. 1.

Nota-se que as estimativas da intensidade de tráfego (ρ) simuladas são próximas do valor teórico. Todas as métricas apresentaram estabilidade numérica, com erros padrão da média inferiores a 2% do valor da estimativa. Além disso, os resultados indicam que o aumento do quadrado do coeficiente de variação dos tempos de serviço implica no aumento do tamanho médio da fila de espera L_q , do tempo médio de espera na fila W_q e da variabilidade das estimativas. Essa observação permite entender um importante aspecto gerencial desses sistemas de filas: fixada uma intensidade de tráfego ρ , a qualidade do serviço, avaliada pelas métricas de desempenho L_q e W_q , está inversamente relacionada com a variabilidade dos tempos de serviço.

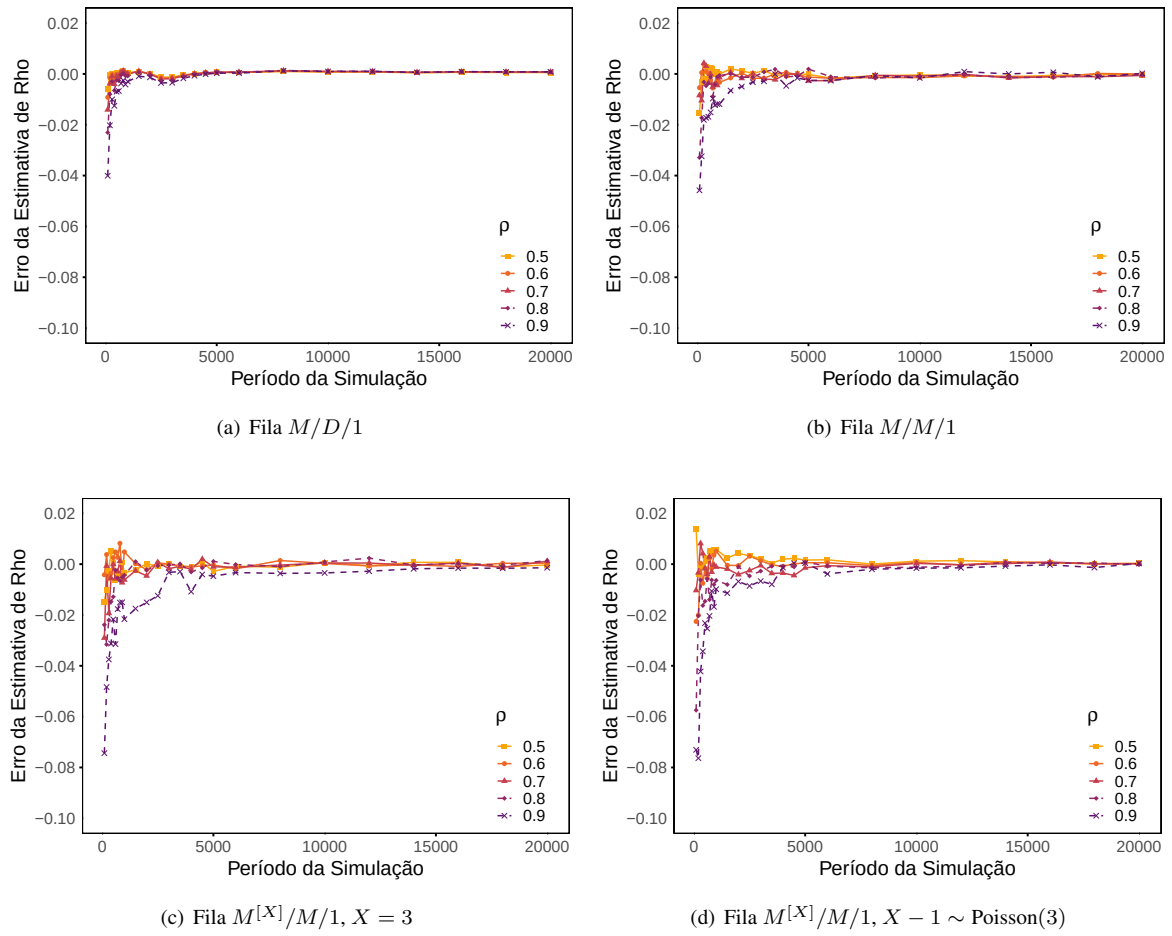
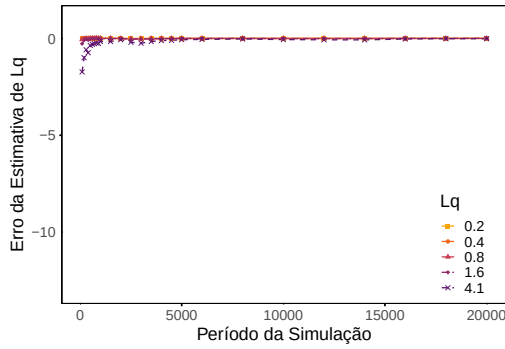


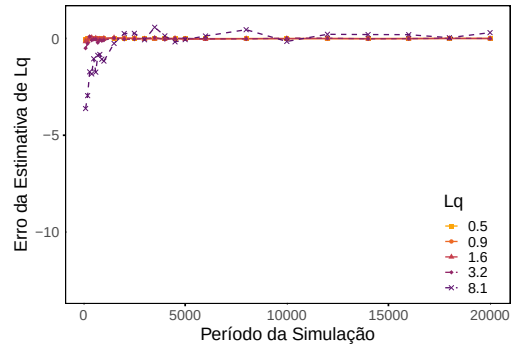
Figura 4. Convergência do algoritmo para a métrica ρ teórica

Tabela 1. Valores médios simulados das métricas de desempenho e seus respectivos erros padrão

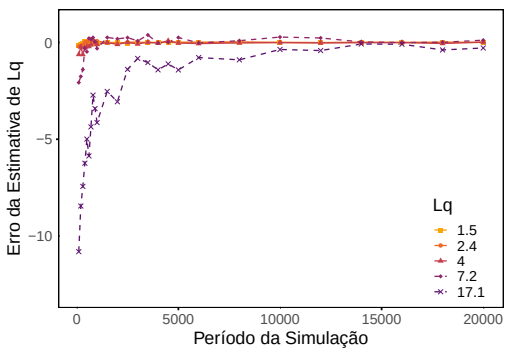
ρ	Modelo	$cv^2 [T]$	métricas estimadas					
			$\hat{\rho}$	$epm(\hat{\rho})$	$\hat{L}_q$	$epm(\hat{L}_q)$	$\hat{W}_q$	$epm(\hat{W}_q)$
0,5	$M/D/1$	0	0,5006	0,0004	0,2504	0,0008	0,2500	0,0007
	$M/M/1$	1	0,4997	0,0005	0,4985	0,0023	0,4983	0,0021
	$M/G/1$	2	0,4996	0,0005	0,7480	0,0034	0,7472	0,0033
0,7	$M/D/1$	0	0,7008	0,0005	0,8195	0,0035	0,8185	0,0030
	$M/M/1$	1	0,6994	0,0007	1,6285	0,0095	1,6282	0,0089
	$M/G/1$	2	0,7001	0,0007	2,4351	0,0153	2,4319	0,0150
0,9	$M/D/1$	0	0,9008	0,0007	4,0368	0,0434	4,0309	0,0414
	$M/M/1$	1	0,9002	0,0009	8,3997	0,1639	8,3899	0,1599
	$M/G/1$	2	0,8993	0,0010	12,2489	0,2218	12,2479	0,2197



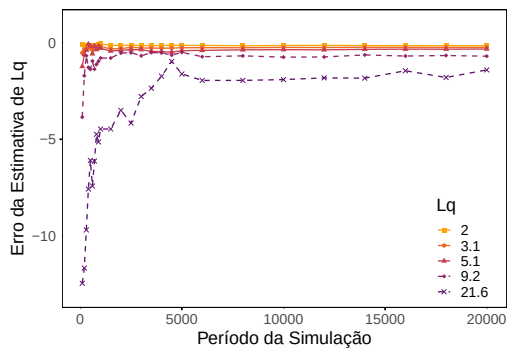
(a) Fila $M/D/1$



(b) Fila $M/M/1$

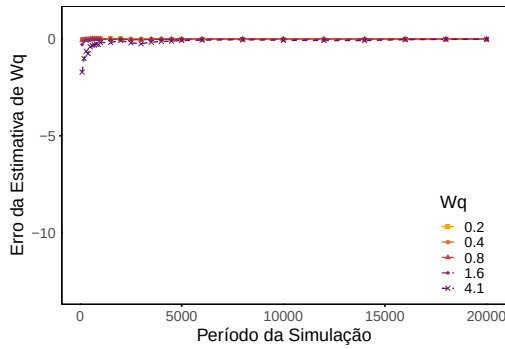


(c) Fila $M^{[X]}/M/1, X = 3$

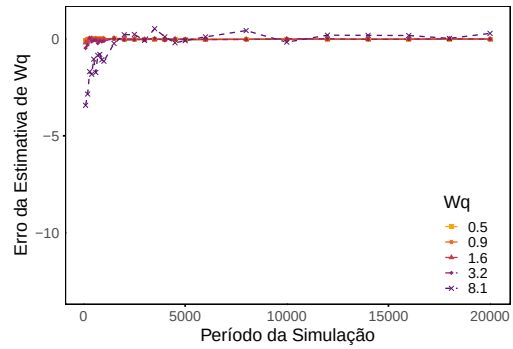


(d) Fila $M^{[X]}/M/1, X - 1 \sim \text{Poisson}(3)$

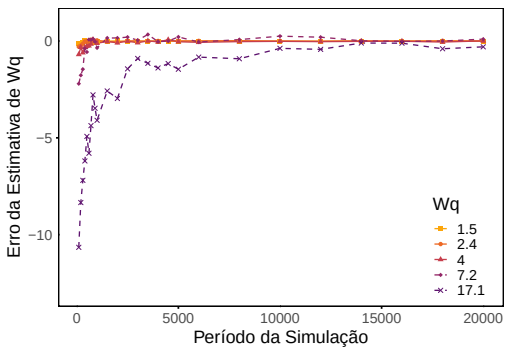
Figura 5. Convergência do algoritmo para a métrica L_q teórica



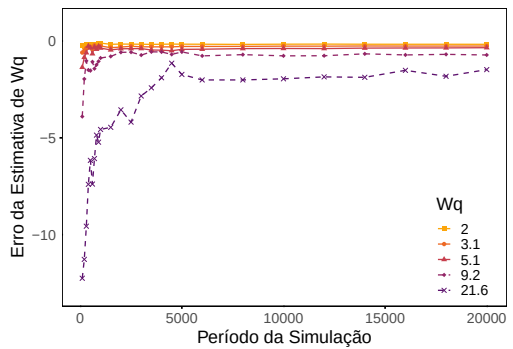
(a) Fila $M/D/1$



(b) Fila $M/M/1$



(c) Fila $M^{[X]}/M/1, X = 3$



(d) Fila $M^{[X]}/M/1, X - 1 \sim \text{Poisson}(3)$

Figura 6. Convergência do algoritmo para a métrica W_q teórica

4. CONCLUSÕES E OBSERVAÇÕES FINAIS

Este trabalho apresentou o desenvolvimento e a validação de um simulador de eventos discretos voltado à análise de desempenho de sistemas de filas, com ênfase inicial em modelos markovianos. A implementação em R mostrou-se adequada e eficiente, permitindo o cálculo de métricas clássicas de desempenho, como ρ , L_q e W_q . Os valores simulados foram comparados a resultados analíticos da literatura, com boa concordância, o que valida o simulador desenvolvido. A ferramenta destaca-se pela flexibilidade e potencial de expansão, oferecendo uma base sólida para o estudo de sistemas mais complexos, especialmente aqueles que não atendem às suposições dos modelos com solução analítica, tornando a simulação a principal alternativa de análise.

Também foram conduzidos experimentos variando o coeficiente de variação dos tempos de serviço, situação sem solução analítica, mostrando que, para uma mesma intensidade de tráfego ρ , maior variabilidade nos tempos de serviço implica aumentos em L_q , W_q e na variância das estimativas, evidenciando a relação inversa entre a qualidade do serviço e a variabilidade nos tempos de atendimento.

Tópicos de interesse para trabalhos futuros incluem o estudo de políticas de atendimento, análise de filas com prioridades, e a modelagem de sistemas com múltiplas estações de serviço ou com restrições de capacidade.

5. AGRADECIMENTOS

Os autores agradecem ao CNPq (processo 305442/2022-8), CAPES (processo 88887.823719/2023-00), FAPEMIG (processo CEX-PPM-00564-17) e PRPq-UFMG, pelo apoio parcial a esse trabalho.

6. REFERÊNCIAS

- Almehdawe, E., Jewkes, B. & He, Q.M., 2016. "Analysis and optimization of an ambulance offload delay and allocation problem". *Omega*, Vol. 65, pp. 148–158.
- Anderson, D.R., Sweeney, D.J., Williams, T.A., Camm, J.D., Cochran, J.J., Fry, M.J. & Ohlmann, J.W., 2013. *Quantitative methods for business*. South-Western, Cengage Learning, Mason, OH, 12^a ed.
- Cruz, G.M.B. & Singh, S.K., 2021. "A discrete event simulation algorithm to finite single-server Markovian queues". In *Anais do XXVIII Simpósio de Engenharia de Produção - XXVIII SIMPEP*. UNESP, Bauru, SP, pp. 1–11.
- Ebert, A., Wu, P., Mengersen, K. & Ruggeri, F., 2020. "Computationally efficient simulation of queues: The R package queuecomputer". *Journal of Statistical Software*, Vol. 95, No. 5, pp. 1–29.
- Fishman, G.S., 2001. *Discrete-event simulation: Modeling, programming, & analysis*. Springer, New York, NY.
- Govil, M.K. & Fu, M.C., 1999. "Queueing theory in manufacturing: A survey". *Journal of Manufacturing Systems*, Vol. 18, No. 3, p. 214.
- Koole, G. & Mandelbaum, A., 2002. "Queueing models of call centers: An introduction". *Annals of Operations Research*, Vol. 113, No. 1-4, pp. 41–59.
- Little, J.D.C., 2011. "Little's law as viewed on its 50th anniversary". *Operations Research*, Vol. 59, No. 3, pp. 536–549.
- R Core Team, 2025. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Shortle, J.F., James M. Thompson, D.G. & Harris, C.M., 2018. *Fundamentals of queueing theory*. John Wiley & Sons, Inc., New York, NY, quinta edition.
- Tas, D., Dellaert, N., van Woensel, T. & de Kok, T., 2013. "Vehicle routing problem with stochastic travel times including soft time windows and service costs". *Computers & Operations Research*, Vol. 40, No. 1, pp. 214–224.

7. RESPONSABILIDADE PELAS INFORMAÇÕES

Os autores são os únicos responsáveis pelas informações incluídas neste trabalho.