

Inferência e Correção de Vício em Filas Markovianas de Servidor Único

Luiz Carlos R. Pereira - *luizreal05@ufmg.br* - Universidade Federal de Minas Gerais

Márcio A. C. Almeida - *estatimarcio@gmail.com* - Universidade Federal do Pará

Frederico R. B. Cruz - *fcruz@est.ufmg.br* - Universidade Federal de Minas Gerais

Fernando L. P. Oliveira - *fernandolui@iceb.ufop.br* - Universidade Federal de Ouro Preto

As filas de espera estão presentes em diversas situações do nosso dia a dia, tais como na lanchonete, no supermercado, no banco e no posto de gasolina. Várias são as medidas que descrevem o desempenho das filas. Neste trabalho daremos foco à estimativa da intensidade do tráfego ρ (também denominado fator de utilização da estação de serviço), para o caso de filas $M/M/1$, que, na notação de Kendall, representa filas com tempos entre chegadas e de serviço exponenciais (markovianos), servidor único e capacidade de espera infinita. Experimentos computacionais apresentados mostram que o estimador de máxima verossimilhança para ρ apresenta vício para amostras pequenas e moderadas ($n < 200$). Adicionalmente, uma versão do estimador, corrigida pelo método *bootstrap* não paramétrico, foi testada. Assim, com um esforço computacional extra, o pesquisador pode alcançar, com amostras de tamanho $n = 5$, estimativas para ρ com erro médio equivalente a estimativas com amostras de tamanho $n = 20$. Esta redução pode ser bastante importante do ponto de vista prático, pela diminuição que pode acarretar nos custos e no tempo para obtenção das amostras.

Palavras-chave: *Bootstrap, desempenho, filas markovianas, inferência, vício.*

Inferência e Correção de Vício em Filas Markovianas de Servidor Único

Luiz Carlos R. Pereira - *luizreal05@ufmg.br* - Universidade Federal de Minas Gerais

Márcio A. C. Almeida - *estatimarcio@gmail.com* - Universidade Federal do Pará

Frederico R. B. Cruz - *fcruz@est.ufmg.br* - Universidade Federal de Minas Gerais

Fernando L. P. Oliveira - *fernandoluiz@iceb.ufop.br* - Universidade Federal de Ouro Preto

1 Introdução

As filas de espera estão muito presentes no nosso dia a dia, em situações tais como na lanchonete, no supermercado, no banco e no posto de gasolina. Comum a todos esses casos é que as filas ocorrem quando a procura por um determinado serviço é maior que a capacidade do sistema de prover este serviço. Conforme Wagner [6], você pode caracterizar um sistema de filas pelo seu processo de chegada, seu processo de saída, a disciplina da fila, seu mecanismo de serviço, entre outras formas. Essas características estão presentes em todos os sistemas de filas e são muito importantes para analisarmos cada tipo de fila.

Várias são as medidas que podem atestar o desempenho de um sistema de filas. Neste trabalho daremos foco à estimativa da intensidade do tráfego ρ (às vezes chamado de fator de utilização da estação de serviço), para o caso de filas $M/M/1$, que, na notação de Kendall [4], representa filas com tempos entre chegadas exponenciais (markoviano), tempos de serviço também exponenciais, servidor único e área de espera infinita.

Este artigo está organizado com se segue. Na Seção 2 é apresentada a metodologia utilizada. Os resultados computacionais e sua análise são discutidos na Seção 3. Finalmente, a Seção 4 encerra o artigo com conclusões finais e tópicos para futuras pesquisas na área.

2 Metodologia

2.1 Filas $M/M/1$

As filas $M/M/1$, apesar de serem as mais simples do ponto de vista matemático, estão entre aquelas com mais aplicações na prática. Elas aproximam diversas situações reais e são a base de modelos mais complexos. Segundo a notação de Kendall [4], são filas com:

$$\begin{aligned} \text{densidade exponencial para os tempos entre chegadas: } & f(t) = \lambda e^{-\lambda t}, \quad t \geq 0, \\ \text{densidade exponencial para os tempos de serviço: } & g(t) = \mu e^{-\mu t}, \quad t \geq 0. \end{aligned} \quad (1)$$

Assumindo que o sistema de filas atinge um regime estacionário (também denominado equilíbrio estatístico), isto é, que ele foi observado após um prolongado período de funcionamento, pode-se calcular a intensidade de tráfego ρ , definido como:

$$\rho \equiv \frac{\lambda}{\mu}, \quad (2)$$

que, para garantir que o sistema atinge o equilíbrio, deve obedecer a $\rho < 1$.

Associada a esta intensidade de tráfego, pode-se encontrar uma distribuição de probabilidade geométrica para o número de clientes no sistema no momento da partida, M , que, após o equilíbrio, pode ser escrita como:

$$P(M = m) = \begin{cases} \rho^m(1 - \rho), & m = 0, 1, 2, \dots, \\ 0, & \text{caso contrário.} \end{cases} \quad (3)$$

A partir da Eq. (3), importantes características de desempenho podem ser encontradas em função de ρ , tais como a probabilidade de servidor ocioso, $P(M = 0) = 1 - \rho$, a esperança do número de clientes no sistema, $L = \rho/(1 - \rho)$, e o tamanho médio da fila, $L_q = \rho^2/(1 - \rho)$.

2.2 Estimador de Máxima Verossimilhança para ρ

Para obtenção de uma estimativa para ρ , baseada em uma amostra \mathbf{X} de tamanho n , pode-se utilizar a Eq. (3) e definir-se a seguinte função de verossimilhança:

$$L(\rho, \mathbf{X}) = (1 - \rho)^n \rho^y, \quad (4)$$

em que $y = \sum_{i=1}^n x_i$.

Pode-se mostrar que o estimador de máxima verossimilhança (EMV), para ρ , é dado por:

$$\hat{\rho} = \frac{y}{n + y}. \quad (5)$$

2.3 Correção *Bootstrap* de Vício

Um método utilizado para correção de vício em estimadores é o *bootstrap* [3]. Na sua versão não paramétrica, para estimar o vício de um parâmetro de interesse, β , realizamos diversas reamostragens (com reposição) e estimamos o parâmetro para cada reamostragem. Tiramos então a média dessas estimativas, $\hat{\beta}_{(\bullet)}$ e o vício pode ser estimado da seguinte forma:

$$\text{vício} = \hat{\beta}_{(\bullet)} - \hat{\beta}. \quad (6)$$

A versão corrigida do estimador, pelo método *bootstrap* não paramétrico, é dada por:

$$\tilde{\beta}_B = 2\hat{\beta} - \hat{\beta}_{(\bullet)}. \quad (7)$$

Este método foi usado por diversos pesquisadores no passado, com resultados bastante satisfatórios (ver, por exemplo [1]). Trabalhos recentes mostram o uso do método *bootstrap* para estimação de intervalos de confiança (ver [2]). A seguir, são apresentados resultados computacionais e as principais conclusões obtidas.

3 Resultados e Discussão

Todos os algoritmos anteriormente descritos foram codificado em R [5] e estão disponíveis a pedido diretamente com os autores. Inicialmente foram realizadas simulações de amostras com intensidades de tráfego $\rho \in \{0,01; 0,10; 0,20; 0,50; 0,90; 0,99\}$ e tamanhos $n \in \{5; 10; 20; 50; 100; 200; 500\}$. Para todos esses cenários foram calculadas estimativas médias de ρ pelo EMV, Eq. (5), com os respectivos erros-padrão da média (EPM), para 10.000 replicações Monte Carlo. Os resultados podem ser vistos na Tabela 1 e nas Figuras 1 e 2.

Tabela 1: Estimativas médias de ρ com respectivo erro padrão entre parênteses

ρ	n						
	5	10	20	50	100	200	500
0.01	0,0080 (0,0004)	0,00910 (0,00029)	0,00968 (0,00021)	0,00987 (0,00014)	0,00994 (0,00010)	0,01000 (0,00007)	0,010025 (0,000044)
0.10	0,0849 (0,0011)	0,09279 (0,00084)	0,09668 (0,00062)	0,09820 (0,00039)	0,09940 (0,00028)	0,09953 (0,00020)	0,100027 (0,000127)
0.20	0,1725 (0,0015)	0,18718 (0,00109)	0,19372 (0,00079)	0,19710 (0,00050)	0,19892 (0,00035)	0,19923 (0,00025)	0,199806 (0,000160)
0.50	0,4528 (0,0017)	0,47665 (0,00116)	0,48838 (0,00081)	0,49509 (0,00050)	0,49801 (0,00035)	0,49868 (0,00025)	0,499603 (0,000158)
0.90	0,8784 (0,0006)	0,88979 (0,00036)	0,89500 (0,00023)	0,89801 (0,00014)	0,89895 (0,00010)	0,89940 (0,00007)	0,899811 (0,000043)
0.99	0,9875 (0,0001)	0,98887 (0,00004)	0,98945 (0,00002)	0,98978 (0,00001)	0,98990 (0,00001)	0,98994 (0,00001)	0,989976 (0,000004)

Podemos constatar que o EMV realmente apresenta vício, para amostras pequenas e moderadas. Para amostras grandes ($n \geq 200$), o comportamento é bastante satisfatório, conforme visto na Figura 1, que representa graficamente os valores apresentados na Tabela 1. Portanto, as medidas de desempenho da fila estimadas a partir de ρ também apresentarão vício, o que deve ser uma preocupação do pesquisador. Outra conclusão importante é que os EPM são maiores para intensidades de tráfego $\rho \approx 0,50$ e menores para intensidades extremas, $\rho \approx 0,99$ e $\rho \approx 0,01$, o que sugere que estes são os valores mais difíceis de estimar na prática. A Figura 2 mostra claramente este fenômeno.

Além disso, testes preliminares foram realizados com a versão corrigida do EMV para ρ , Eq. (5). Os valores médios obtidos, para diversos tamanhos de amostra, para 10.000 replicações Monte Carlo e $\rho = 0,5$, são apresentados na Figura 3. É possível notar pelos resultados que, com um esforço computacional extra, o pesquisador pode alcançar com amostras de tamanho $n = 5$ estimativas para ρ com o mesmo erro médio que para amostras de tamanho $n = 20$. Esta redução pode ser bastante importante do ponto de vista prático, pela diminuição que pode acarretar nos custos e no tempo para obtenção das amostras.

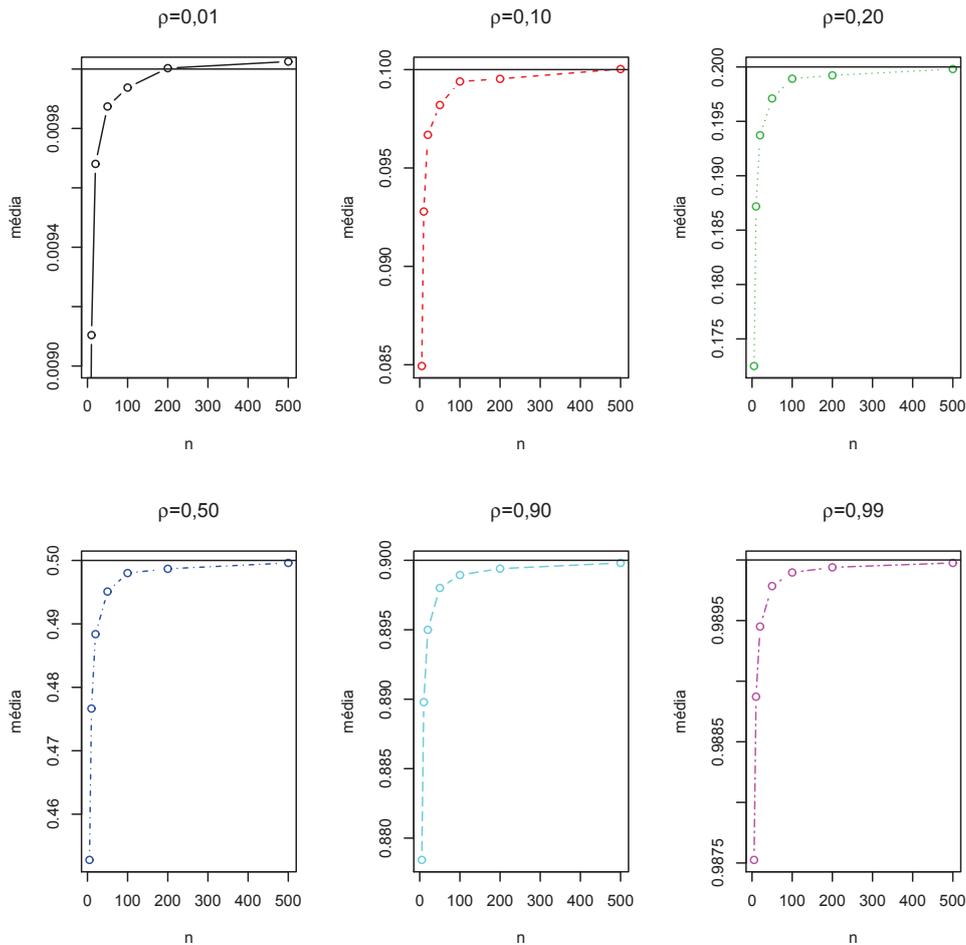


Figura 1: Estimativas médias de ρ em função do tamanho da amostra n

4 Conclusões e Observações Finais

Dada a importância das filas no dia a dia das pessoas, este trabalho abordou o problema de estimação de um dos seus parâmetros, a intensidade de tráfego ρ . O foco foi o tipo mais simples de filas, $M/M/1$, com chegadas e partidas markovianas e servidor único, que é uma aproximação acurada para vários problemas práticos e base para modelos mais completos e abrangentes. Foi constatada a presença de vício no estimador de máxima verossimilhança para ρ e investigada a utilização do conhecido método *bootstrap* não paramétrico para sua correção. Resultados preliminares promissores apontaram para uma correção efetiva do vício, possibilitando redução no tamanho de amostra, com conseqüente diminuição nos custos e tempo para sua obtenção.

Investigações futuras incluem a avaliação da qualidade da correção pelo *bootstrap* em outras medidas de desempenho da fila, tais como a probabilidade de sistema vazio, no tamanho médio da fila e número médio de usuários no sistema. Também outros tipos de filas devem ser investigados, tais como filas markovianas infinitas multi-servidoras, $M/M/c$, filas finitas, $M/M/c/K$, e assim por diante.

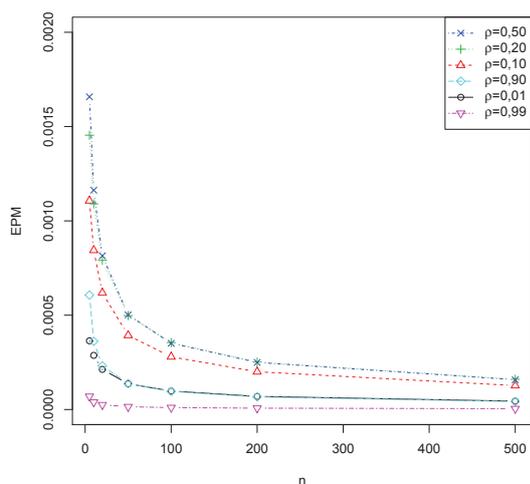


Figura 2: Erro-padrão das estimativas médias

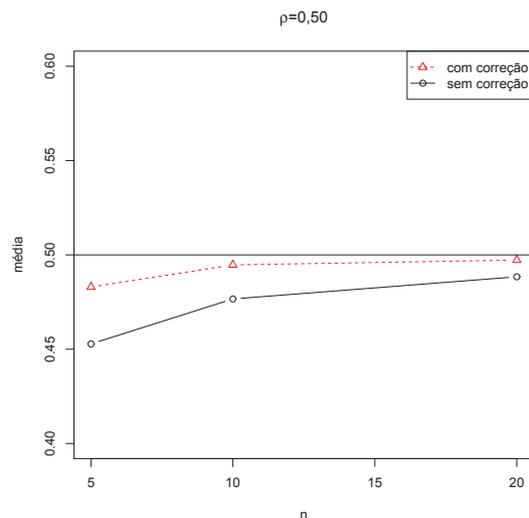


Figura 3: Correção das estimativas médias

Agradecimentos

Os autores agradecem ao CNPq e à Fundação de Amparo à Pesquisa do estado de Minas Gerais (FAPEMIG), pelo auxílio financeiro a esta pesquisa, bem como à UFOP e à UFPA.

Referências

- [1] CRUZ, F. R. B.; COLOSIMO, E. A.; MACGREGOR SMITH, J. Sample Size Corrections for the Maximum Partial Likelihood Estimator. **Communications in Statistics - Simulation and Computation**. Taylor & Francis, v. 33, n. 1, p. 35-47, 2004.
- [2] DOMINGUES, K. M.; OLIVEIRA, F. L. P.; CRUZ, F. R. B.; BESSEGATO, L. F. Estimação de Intervalos de Confiança via Reamostragem *Bootstrap*. **I Encontro Interacadêmico de Engenharia de Produção - EINEPro 2015**. São João da Barra, RJ, p. 1-12, 2015.
- [3] EFRON, B., TIBSHIRANI, R. **An introduction to the bootstrap**. London, UK: Chapman e Hall, 1993. 456 p.
- [4] KENDALL, D. G. Stochastic processes occurring in the theory of queues and their analysis by the method of imbedded Markov chains. **Annals of Mathematical Statistics**. Institute of Mathematical Statistics, v. 24, p. 338-354, 1953.
- [5] R CORE TEAM. **R: A Language & Environment for Statistical Computing**. R Foundation for Statistical Computing, Vienna, Austria, 2013. <http://www.R-project.org/>
- [6] WAGNER, H. M. **Pesquisa Operacional**. Prentice-Hall do Brasil Ltda., Rio de Janeiro, 2 ed., 1985. 851 p.