

## Estimação da Intensidade de Tráfego em Filas Markovianas de Servidor Único via SIR

Victor Brito Quinino<sup>1</sup>

UFMG, Belo Horizonte, MG

Frederico R. B. Cruz<sup>2</sup>

UFMG, Belo Horizonte, MG

Filas markovianas de servidor único, que são representadas na notação de Kendall por  $M/M/1$ , são de interesse na engenharia, pois podem ser vistas como um primeiro passo para um processo mais elaborado de análise, com modelos de filas mais sofisticados. Uma vez escolhido o modelo apropriado, a tarefa seguinte é a estimação estatística de seus parâmetros.

O foco deste artigo é apresentar a utilização do método de amostragem/reamostragem de importância (ou SIR, do inglês *sampling/importance resampling*), para a estimação da intensidade de tráfego em filas  $M/M/1$ , ou seja, a fração de tempo em que a fila fica ocupada, definida como a razão entre a taxa de chegada e a taxa de serviço.

Suponha que  $X_1, X_2, \dots, X_i, \dots$ , sejam variáveis aleatórias independentes e identicamente distribuídas, em que  $X_i$  denota o número de usuários no sistema no  $i$ -ésimo momento aleatório de uma fila  $M/M/1$ . Assim, a função de probabilidade no ponto da variável aleatória  $X_i$  é dada por  $P(X = x) = \rho^x(1 - \rho)$ ,  $x = 0, 1, 2, \dots$ , em que  $\rho = \lambda/\mu > 0$  é a intensidade de tráfego. A função de verossimilhança, baseada em amostra aleatória, é dada por:

$$L(\rho, \mathbf{x}) = \rho^{x_1}(1 - \rho) \times \rho^{x_2}(1 - \rho) \times \dots \times \rho^{x_n}(1 - \rho) = \rho^y(1 - \rho)^n, \quad (1)$$

em que  $y = \sum_{i=1}^n x_i$ , o que produz o seguinte estimador de máxima verossimilhança (EMV) para  $\rho$ ,  $\hat{\rho}_{MLE} = \frac{y}{y+n}$ .

O método não iterativo denominado “*sampling/importance resampling*” (SIR), discutido em detalhes por Smith and Gelfand [2] e Rubin [1], é utilizado para realização de inferência Bayesiana e é descrito de forma breve a seguir.

Uma amostra aleatória de tamanho  $k$  é selecionada da função densidade de probabilidade a priori  $\pi(\rho|\boldsymbol{\alpha})$ , em que  $k \geq 5.000$ . Os valores amostrais são designados como  $\rho_{\text{prior},i}$ ,  $i = 0, 1, \dots, k$ . Uma função peso  $W_i$ ,  $i = 0, 1, \dots, k$ , é calculada para cada ponto  $\rho_{\text{prior},i}$ , em que  $W_i$  é proporcional à função de verossimilhança. Por último, uma amostra aleatória também de tamanho  $k$  é selecionada dos valores a priori de  $\rho_{\text{prior},i}$ ,  $i = 0, 1, \dots, k$ , com probabilidades proporcionais aos  $W_i$ . Essa nova amostra, denominada  $\rho_{\text{post},i}$ ,  $i = 0, 1, \dots, k$ , pode ser considerada como proveniente da distribuição a posteriori de  $\rho$  e pode ser utilizada para a estimativa bayesiana para  $\rho$  via SIR, conforme a seguir.

$$\hat{\rho}_{\text{SIR}} = \frac{\sum_{i=1}^k \rho_{\text{post},i}}{n}. \quad (2)$$

Os experimentos computacionais apresentados objetivam comparar, via simulação Monte Carlo e utilizando-se a raiz do erro quadrático médio, a efetividade do método SIR em relação ao MVE,

<sup>1</sup>victorquinino@ufmg.br.

<sup>2</sup>fcruz@est.ufmg.br.

quando utilizados para estimar  $\rho$  em uma fila  $M/M/1$ , com formas de seleção amostral idênticas. Para tal, foram utilizados tamanhos amostrais  $n \in \{10, 20, 50, 80, 100, 200\}$ , com 1000 replicações Monte Carlo. As amostras da quantidade de clientes no sistema foram geradas por uma distribuição geométrica com parâmetro  $1 - \rho$ . O tamanho gerado da distribuição a priori de  $\rho$  foi fixado em  $k = 5000$ . Para uma comparação em situações diversas, foram criados três casos diferentes de distribuição a priori: (i) utilizando-se uma distribuição uniforme; (ii) utilizando-se uma distribuição triangular com moda localizada exatamente no valor simulado de  $\rho$ , que é o melhor caso possível; e (iii) utilizando-se uma distribuição triangular com moda localizada em um ponto distante do valor simulado de  $\rho$ , que é o pior caso possível.

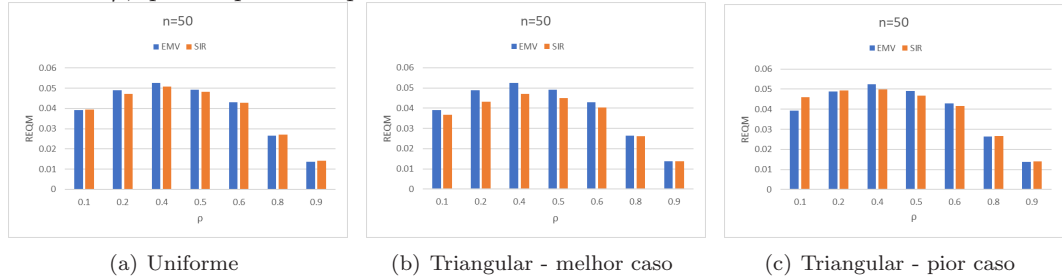


Figura 1: Comparação do desempenho dos métodos SIR e EMV para três distribuições, com  $n = 50$ . Para os três casos, o método SIR apresentou um desempenho superior (um menor erro quadrático médio) na região central do espaço paramétrico. Para os casos (i) e (iii) o método EMV se mostrou mais eficiente nas fronteiras do espaço paramétrico. No caso (ii) o método SIR se mostrou mais eficiente em todo espaço paramétrico. Deve-se notar, também, que, para um maior tamanho amostral, ( $n \geq 100$ ) os resultados dos dois métodos tendem a apresentar desempenhos equivalentes.

Podemos concluir que o SIR mostra-se, de forma geral, superior ao EMV, para valores de  $\rho$  na região central do espaço paramétrico, ou seja, para valores de  $\rho$  próximos de 0.5. Por outro lado, o EMV mostra-se melhor nas fronteiras do espaço paramétrico, ou seja, para  $\rho = 0.1$  ou  $\rho = 0.9$ . Além disso, conforme o tamanho amostral  $n$  aumenta, os dois métodos se aproximam e convergem para o valor simulado.

Podemos observar também que, caso a distribuição a priori seja feita com um bom conhecimento prévio de  $\rho$ , o método SIR se mostra superior ao EMV, em todo o espaço paramétrico, o que é uma situação possível na prática visto que, em vários sistemas, os gestores do sistema têm uma ideia razoável do valor do  $\rho$ , por meio do conhecimento de quantas pessoas há em média no sistema.

Futuros trabalhos podem ser feitos utilizando outros métodos de coleta amostral, por exemplo contando o número de chegadas durante o serviço do  $i$ -ésimo usuário. Além disso, novos trabalhos podem ser desenvolvidos para estimativas via SIR em outros tipos de filas, p.e.,  $M/M/s$  e outras.

## Agradecimentos

Os autores agradecem ao CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico, processos 305515/2018-7 e 160974/2020-8) e à FAPEMIG (Fundação de Amparo à Pesquisa do Estado de Minas Gerais, processo CEX-PPM-00564-17).

## Referências

- [1] Rubin, D. B. Using the SIR algorithm to simulate posterior distributions. *Bayesian Statistics*, 3:395–402, 1988.
- [2] Smith, A. F. M. and Gelfand, A. E. Bayesian statistics without tears: A sampling-resampling perspective. *The American Statistician*, 46(2):84–88, 1992.