UNIVERSIDADE FEDERAL DE MINAS GERAIS
Programa de Pós-Graduação em Estatística

Estimação e Intervalo de Confiança Bootstrap para a Intensidade de Trafego em filas M/M/1/K usando o Método do Mínimo Qui-Quadrado

Monografia apresentada ao
Curso de Especialização em Estatística com ênfase em Indústria e Mercado
Programa de Pós-Graduação em Estatística da
Universidade Federal de Minas Gerais

Autor: Vinícius Melo de Araújo Rocha
Orientador: Prof. Roberto Quinino

Belo Horizonte
Outubro / 2018

Vinícius Melo de Araújo Rocha

Estimação e Intervalo de Confiança Bootstrap para a Intensidade de Trafego em filas M/M/1/K usando o Método do Mínimo Qui-Quadrado

Orientador: Prof. Roberto Quinino

Monografia apresentada ao curso de Especialização em Estatística do Instituto de Ciências Exatas da Universidade Federal de Minas Gerais, como requisito parcial para obtenção do título de Especialista em Estatística.

Belo Horizonte
Instituto de Ciências Exatas da UFMG

Agosto / 2018

A Deus por nos amparar em todos os momentos de nossas vidas e a minha família, que sempre esteve ao meu lado e me deu força, coragem e suporte para seguir até o fim desta jornada.

"O sucesso nasce do querer, da determinação e persistência em se chegar a um objetivo. Mesmo não atingindo o alvo, quem busca e vence obstáculos, no mínimo fará coisas admiráveis."

José de Alencar

AGRADECIMENTOS

Agradeço a UFMG, quem me deu esta oportunidade de poder estudar, a minha família, a qual agradeço pelo apoio incondicional e pela base de minha educação. Agradeço ao Professor Roberto Quinino que com sua didática e paciência torna o estudo muito mais interessante e instigante.

Resumo

Teoria de filas é um campo vasto de estudos. O processo de prestação de serviços no qual se exige formação de filas de atendimento é objeto de estudo de diferentes pesquisadores. Analisar o processo de formação das filas, desempenho no atendimento e as variáveis que impactam em seu processo de formação como intensidade de trafego, definida como sendo a razão entre a taxa de chegada e a taxa de atendimento se fazem de grande importância a pesquisadores que analisam a eficiência do dimensionamento da estrutura de atendimento. Todavia, sabe-se que na maioria das vezes os parâmetros envolvidos no processo não são conhecidos e precisam ser estimados. Esta monografia visa obter intervalos de confiança *bootstrap* para a intensidade de trafego em filas markovianas finitas com um único servidor, denominadas M/M/1/K, na notação de Kendall. Para tanto utilizaremos o método de estimação do mínimo qui-quadrado. Um exemplo numérico ilustra a eficiência do procedimento.

Palavras-chave: Filas markovianas, servidor único, intensidade de trafego, estimação por mínimo qui-quadrado, bootstrap.

Abstract

Queuing theory is a vast research field. The service process in which queues are formed is the object of study by different researchers. Analyzing the queuing process, service performance and the variables that impact on its arrangement, such as the traffic intensity, defined as the ratio between the arrival rate and the service rate, are of great importance to researchers who analyze efficiency on dimensioning the service structure. However, it is known that in most cases the parameters involved in the process are unknown and need to be estimated. The monography aims to obtain bootstrap confidence interval for the traffic intensity in finite Markovian queues with a single server, known as M / M / 1 / K, in Kendall notation. For this we will use the estimation method of the chi-square minimum. A numerical example illustrates the efficiency of the procedure.

Keywords: Markovian queues, single server, traffic intensity, chi-square minimum estimation, bootstrap.

LISTA DE FIGURAS

LISTA DE TABELAS

SUMÁRIO

1. Introdução

A utilização de filas está no cotidiano da sociedade. Filas são utilizadas tanto em ambientes físicos como supermercados, padarias, postos de gasolina, como em ambientes virtuais como canais de atendimento telefônico, canais de atendimento de *chats* via internet ou qualquer outro lugar onde a estrutura de atendimento seja finita e o fluxo de chegada seja maior que o fluxo de saída. Segundo Almeida & Cruz (2015) determinado sistema de filas pode ser descrito por seis componentes, sendo: (i) o processo de chegada, (ii) a distribuição do tempo de serviço, (iii) o número de servidores, (iv) a capacidade do sistema, (v) a população de usuários e (vi) a disciplina de atendimento.

De maneira geral, sistemas de filas possuem características diversas, todavia, ainda que em diferentes setores, sua forma de funcionamento é similar. O objeto de estudo na formação e comportamento de filas visa tornar eficiente a estrutura de atendimento, minimizando custo e maximizando o desempenho de atendimento.

Desta forma, conhecer algumas das características dos componentes determinadores da fila, tais como a taxa de chegada λ, o tempo médio $1/\mu$ que um servidor leva para executar o serviço, a intensidade do trafego ρ, definida como sendo a razão entre λ e μ, a esperança do número de clientes no sistema L e o tamanho da fila L_q, se fazem necessários para posterior estudo. Nesta monografia trataremos da estimação da intensidade de trafego ρ em filas finitas com chegada Poisson e taxa de atendimento exponencial. O termo finito indica que estamos supondo que o número de clientes no sistema não pode ultrapassar uma quantidade especificada (neste estudo representaremos por K). Em geral isto é encontrado na maioria das situações reais como supermercados, bancos, etc. Sua representação usual, segundo Kendall (1953), é M/M/1/K, sendo que o primeiro M indica chegadas Poisson, o segundo M taxa de atendimento exponencial, o terceiro termo indica um servidor e o último que o sistema possui capacidade máxima de K clientes no sistema.

Atualmente existe vasta quantidade de estudos a respeito de diversas características das filas, tanto por métodos bayesianos, quanto por métodos inferenciais clássicos. Para intensidade de trafego, Choudhury and Borthakur (2008) apresentam estimadores bayesiano e de máxima verossimilhança para filas M/M/1. Já, via métodos clássicos, cita-se como exemplo o trabalho realizado por Clarke (1957), que apresenta estimadores de máxima verossimilhança para taxa de chegada e taxa de atendimento em filas M/M/a. Bhat et. Al. (1997) mostram que os estimadores de momentos podem ser usados para estimar a taxa de chegada e atendimento em filas M/M/1. O objetivo desta monografia visa obter intervalo de confiança *bootstrap* para a intensidade de trafego em filas markovianas finitas com um único servidor, que será feito a partir da apresentação da estimativa para intensidade de trafego ρ em filas M/M/1/K pelo método do mínimo qui-quadrado avaliando o seu desempenho. Considera-se que o método do mínimo qui-quadrado é mais intuitivo que o de máxima verossimilhança (Berkson, 1980). Não encontramos na literatura a avaliação do método proposto em sistema de filas.

O trabalho está organizado como se segue. Na seção 2, detalhamos o modelo probabilístico para filas M/M/1/K e o método de estimação do mínimo qui-quadrado. Na seção 3 realizamos um detalhamento da metodologia de

Bootstrap. Na seção 4 apresentamos um exemplo numérico e conclusões e sugestão de trabalhos futuros são descritas na seção 5.

2. Modelo Probabilístico e o Método do Mínimo Qui-Quadrado

De acordo com Gross et al. (2009), a distribuição estacionária do número de clientes (*M*) no sistema em uma fila M/M/1/K, considerando ρ <1, é dado por (2.1).

$$P_m \equiv P(M=m) = \begin{cases} \dfrac{1-\rho}{1-\rho^{K+1}}\rho^m, \; m=0,1,\ldots,K \end{cases} \qquad (2.1)$$

Seja uma amostra aleatória do sistema em que se observa x_i, $i=1,2,\ldots,n$ a quantidade de clientes no sistema. Considere que $\boldsymbol{X} = \{x_1, x_2, \ldots, x_n\}$ representa a amostra aleatória de tamanho *n*. A frequência observada para os valores de *m*, *m*=0,1, ..., *K*, é indicada por O*j*, *j*=1,2, ..., *K*+1. Ou seja, na amostra **X** observamos O_1 resultados com *m*=0; O_2 resultados com *m*=3 até O_{k+1} resultados com *m*=*K*. Já o valor esperado de cada uma das K+1 categorias será dada por $E_j = n\pi_j$ em que π_j é dado por:

$$\pi_j = \begin{cases} \dfrac{1-\rho}{1-\rho^{K+1}}\rho^{j-1}, \; \mathrm{j}=1,\ldots,K+1 \end{cases} \qquad (2.2)$$

Considere a estatística definida pela expressão (2.3). Dado a amostra X, o valor de ρ que minimiza (2.3), $\hat{\rho}$, será o estimador denominado mínimo qui-quadrado.

$$\chi^2(\rho) = \sum_{j=1}^{k+1} \frac{(O_j - E_j)^2}{E_j} \qquad (2.3)$$

De acordo com e Berkson (1980) estimadores baseado no método do mínimo qui-quadrado são consistentes e assintoticamente eficientes.

3. Metodologia Bootstrap

Todos os algoritmos foram implementados no MATLAB (versão MATLAB® R2008a) e o código está disponível mediante solicitação para fins de ensino e pesquisa diretamente com o autor ou no anexo A.

Em estatística, bootstrap é um método de reamostragem proposto por Efron (1979). É utilizado para aproximar a distribuição de uma estatística baseado em reamostragens de uma amostra aleatória e permite, assim, estimarmos o erro padrão de estatísticas bem como construir intervalos de confiança e realizar testes de hipóteses sobre parâmetros de interesse.

Segundo Efron (1979) a premissa base da metodologia de bootstrap é que, na ausência de informações sobre a distribuição P, a amostra observada contém todas as informações disponíveis acerca da distribuição em estudo, sendo assim, reamostrar a amostra é a melhor forma para se encontrar os resultados esperados sobre o comportamento de reamostragem da distribuição alvo.

A significancia de Efron (1979) é descrita no trabalho de Davison e Hinkley (1997) que cita:

"A primeira publicação de Bradley Efron (1979), publicado acerca da metodologia de bootstrapping foi um evento de grande relevância em estatística, onde, de uma única vez, sintetizou novas ideias sobre reamostragem e estabeleceu novas formas de trabalho para a produção de simulações baseadas em analises estatísticas. A ideia de reamostragem, até então, complicava e se mostrava como um fator que contribuía com falta de precisão, levando a vieses, variâncias, e outras medidas de incertezas uma vez utilizadas simulações computadorizadas. Estes fatores, de incremento de imprecisão, ocupavam tempo e esforço de pesquisadores teóricos e usuários da analise estatística."

Maiores detalhes sobre a metodologia bootstrap podem ser encontrados em Efron & Hastie (2016).

Segundo Davison e Hinkley (1997) a linha tradicional de estudos para inferência estatística se baseia na utilização de modelos já idealizados e assunção de premissas. Com frequência fórmulas para medição de precisão como erro padrão são baseadas em teoria assintótica e não são utilizáveis para pequenas amostras. Uma moderna alternativa a linha tradicional é a inferência via bootstrap, metodologia introduzida por Efro (1979). Bootstrapping é um forma computacional de reamostragem, cuja aplicação se dá em um vasto campo de aplicação e permite o tratamento de distribuições desconhecidas em modelos mais realísticos.

Frequentemente os estudos se interessam apenas em uma característica da distribuição do estimador, como por exemplo o erro padrão ou o viés. Estimativas para estas medidas podem ser eficientemente obtidas da aplicação da metodologia de bootstrap.

O procedimento bootstrap utilizado nesta monografia utilizou as seguintes etapas:

Etapa 1: Considere uma amostra aleatória de *n* observações do sistema que se deseja estimar a intensidade de trafego ρ.

Etapa 2: Com os dados da Etapa 1, encontre o estimador baseado no mínimo qui-quadrado de ρ, seja: $\hat{\rho}$.

Etapa 3: Com a estimativa $\hat{\rho}$ obtida na Etapa 2, determine a distribuição de probabilidade das possíveis quantidades presentes no sistema (0 a k). Use a expressão (2.2). Faça uma variável contadora *b* ser igual a zero e o número de reamostragem bootstrap ser *B*=50000.

Etapa 4: Gere *n* observações dos valores compreendidos entre 0 e *k* com probabilidades definidas na Etapa 3. Quantifique a frequência de cada um dos valores compreendidos entre 0 e k e aloque no vetor O. Encontre o valor de ρ que minimiza (2.3) e defina como $\hat{\rho}^B$. Arquive esta estimativa na b+1 linha do vetor M.

Etapa 5: Se $b < B$ então vá para Etapa 4 e caso contrário vá para Etapa 6.

Etapa 6: Calcule o percentil 2,5% e 97,5% de *M*. Estes valores definem o 95% percentile method confidence interval.

4. Exemplo Numérico

Considere um sistema de atendimento eletrônico via chat com apenas um canal de atendimento muito requisitado. Os pedidos de conexão chegam de acordo com um processo de Poisson a uma taxa λ por hora. A duração de cada atendimento apresenta distribuição exponencial com parâmetro μ minutos. Se a conexão com o atendente estiver ocupada, os pedidos de conexão seguintes são colocados em espera até um máximo de dez pedidos pendentes, sendo que os pedidos excedentes serão simplesmente perdidos (K = 11). Uma amostra aleatória de 100 observações do sistema é realizada com objetivo de estimar a intensidade de trafego (ρ). Cada observação quantifica o número de pedidos de conexão no atendimento eletrônico. Tomando-se como base experiências anteriores, a amostra foi coletada tendo a as observações feitas com espaços de 10 minutos de diferença, tornando aceitável a independência das observações.

Na tabela abaixo, as possíveis categorias (CAT) e suas frequências observadas (O) são apresentadas na Tabela 2, com a finalidade de estimar a intensidade de trafego ρ.

CAT	O	CAT	O
0	13	6	9
1	13	7	9
2	9	8	8
3	8	9	2
4	10	10	7
5	5	11	7

Tabela 1: Categorias versus frequências observadas

Neste sentido a estimativa da intensidade de trafego será o valor de ρ que minimiza (2.3). O valor encontrado para estimativa da intensidade de trafego foi $\hat{\rho} = 0.94.$ O intervalo bootstrap de 95% apresentou os limites superior e inferior dados respectivamente por 0,9930 e 0.8870.

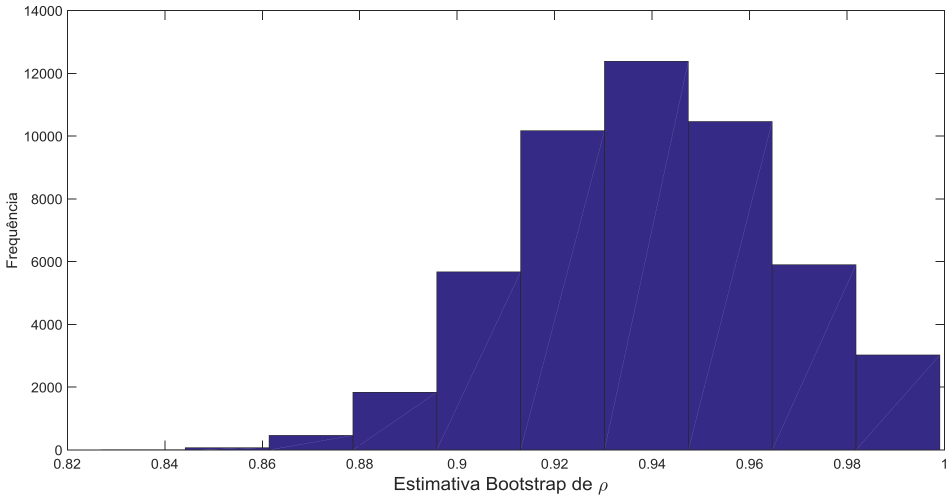

Figura 1: Estimativas Bootstrap de ρ.

5. Conclusões e Sugestão de Estudos Futuros

Intervalos de confiança Bootstrap para intensidade de trafego utilizando o procedimento de mínimo qui-quadrado foi utilizado em filas M/M/1/K. O procedimento apresentado mostrou-se simples e foi desenvolvido um script em Matlab que pode ser usado sem grandes dificuldades.

Considerando que o procedimento de estimação por mínimo-quadrado é consistente e assintoticamente eficiente, propriedade também do estimador de máxima verossimilhança, uma continuidade da pesquisa seria compará-los em amostras pequenas por meio da amplitude dos intervalos.

Anexo A: Programa Matlab

```
clear all
tic
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Estimativa Incicial do Bootstrap
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

R=[13 13 9 8 10 5 9 9 8 2 7 7]; %dados observados na coleta
R=R';

Ta=100;
k=11; %Capacidade do Sistema
Dados=0:k;
cont=1;
for rho=0.75:0.001:0.999

P0=(1-rho)/(1-rho^(k+1));

cont1=1;
for n=0:k

Pn(cont1,1)=P0*(rho^n);
cont1=cont1+1;
end
O=R;
E=n*Pn;
Qui(cont,1)=sum(((O-E).^2)./E);
Qui(cont,2)=rho;
cont=cont+1;

end

[M,j]=min(Qui(:,1));
rhoe=Qui(j,2)

toc

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Início do Bootstrap
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

corridas=50000; %Réplicas Bootstrap

P0=(1-rhoe)/(1-rhoe^(k+1));

cont2=1;
for n=0:k
Pnf(cont2,1)=P0*(rhoe^n);
cont2=cont2+1;
end

for i=1:corridas;
R1=randsample(Dados,Ta,true,Pnf);
cont3=1;
for j=0:k
Ob(cont3,1)=sum(R1==j);
cont3=cont3+1;
```

```
end
cont4=1;
for rhob=0.7:0.001:0.999
P0=(1-rhob)/(1-rhob^(k+1));
cont5=1;
for n=0:k
Pnb(cont5,1)=P0*(rhob^n);
cont5=cont5+1;
end
Eb=Ta*Pnb;
Quib(cont4,1)=sum(((Ob-Eb).^2)./Eb);
Quib(cont4,2)=rhob;
cont4=cont4+1;
end
[Mb,jb]=min(Quib(:,1));
Saida(i,1)=Quib(jb,2);
end

LS = prctile(Saida(:,1),97.5)
LS = prctile(Saida(:,1),2.5)
```

Referências Bibliográficas

Almeida, M. A. C., Cruz, F. R. B. (2015) "Análise de Desempenho em Filas M/M/1 Usando uma Abordagem Bayesiana", Proceeding Series of the Brazilian Society of Applied and Computational Mathematics, 3 (2), 1-7.

Bhat, U. N., Miller, G. K., Rao, S. S. (1997) "Statistical Analysis of Queueing Systems", Frontiers in Queueing: Models and Aplications in Science and Engineering. CRC Press, Probability and Stochastics Series. Edited by Jewgeni H. Dshalalow.

Clarke, A. B. (1957) "Maximum likelihood estimates in a simple queue", *The Annals of Mathematical Statistics*", vol. 28, 1036-1040.

Gross, D.; Shortle, J. F.; Thompson, J. M. & Harris, C. M. (2009). *Fundamentals of queueing theory*, 4th ed., Wiley-Interscience, New York, NY.

Kendall, D. G. (1953). "Stochastic processes occurring in the theory of queues and their analysis by the method of imbedded Markov chains", *Annals of Mathematical Statistics*, vol. 24, pp. 338-354.

Davison, A. C.; Hinkley, D. V. Bootstrap methods and their application. Cambridge: Press Syndicate of the University of Cambridge, 1997.

Berkson, J. 1980. 'Minimum Chi-Square, Not Maximum Likelihood'. The Annals of Statistics

MATLAB® Release R2008a, The MathWorks™, Inc., Natick, Massachusetts, USA.

Bradley Efron (1979). Bootstrap Methods: Another Look at the Jackknife. The Annals of Statistics. 7 (1): 1–26

Efron, B. & Hastie, T. (2016). Computer Age Statistical Inference: Algorithms, Evidence, and Data Science. New York, NY: Cambridge University Press.