

Como os *softwares* simulam dados?

Daiane de Oliveira Gonçalves¹

Discente, Licenciatura em Matemática, Bolsista PROBIC/UNIFAL, UNIFAL-MG, Alfenas, MG

Andréia do Carmo de Oliveira²

Discente, Licenciatura em Matemática, UNIFAL-MG

Ludmila Cristine Araújo de Oliveira³

Discente, Licenciatura em Matemática, UNIFAL-MG

Mayra Camila de Oliveira Silva⁴

Discente, Licenciatura em Matemática, UNIFAL-MG

Fabricio Goecking Avelar⁵

Instituto de Ciências Exatas, UNIFAL-MG

1 Introdução

Segundo [1], simulação é o ato de tratar um problema real através da sua reprodução em um ambiente controlado pelo pesquisador. Este ambiente pode também ser reproduzido por um sistema computacional, através de geração de dados aleatórios que descrevem uma situação controlada. A simulação de dados de distribuições discretas e contínuas tem sido utilizada em diversas áreas do conhecimento tais como Matemática Aplicada e Estatística. Geralmente os *softwares* estatísticos e matemáticos possuem comandos prontos para a simulação de dados das distribuições mais comuns. Mas, e se for necessária a simulação de dados de uma distribuição cuja função geradora de dados não esteja implementada? Nesse caso é imprescindível o conhecimento das técnicas de simulação de dados de distribuições discretas e contínuas. Portanto, o objetivo desse trabalho é apresentar uma forma de simular dados de distribuições discretas e contínuas.

Segundo [2], a maioria das técnicas de simulação de dados de distribuições contínuas e discretas é baseada no Método Congruencial, que consiste em gerar dados x_1, x_2, \dots, x_n a partir de um valor x_0 com a forma recursiva $x_n \equiv ax_{n-1} \pmod{m}$ em que a e $m \in \mathbb{Z}_+$. Dividindo-se cada x_i por m obtém-se uma amostra aleatória de tamanho n de uma variável aleatória U que possui distribuição Uniforme Contínua no intervalo $[0, 1]$ ($U \sim U_c[0, 1]$). O Método da Transformação Inversa para simular dados de uma distribuição discreta

¹daiane_og@hotmail.com

²andreaia_areado88@hotmail.com

³lud.araujo_@hotmail.com

⁴mayra_camila12@hotmail.com

⁵fabricio@unifal-mg.edu.br

consiste, basicamente, em descobrir em qual intervalo $[F(x_{j-1}), F(x_j))$ o valor da variável U está, em que F é a função de distribuição acumulada da variável X da qual se deseja simular os dados e x_j , $j = 1, \dots$ são os valores assumidos pela variável X . Já para se simular dados de uma variável contínua X basta observar que $X = F^{-1}(U)$, em que F^{-1} é a função inversa da função de distribuição F da variável X . A demonstração de tal fato pode ser encontrada em [2].

2 Materiais e Métodos

O Método da Transformação Inversa para o caso discreto foi utilizado para se simular dados da distribuição Geométrica e o mesmo método para o caso contínuo foi utilizado na simulação de dados da distribuição Exponencial. Foram simulados dados de diversos tamanhos amostrais e diferentes valores para os parâmetros do modelo. Para a distribuição Geométrica foram escolhidos os valores do parâmetro do modelo: 0, 1, 0, 3, 0, 5, 0, 7 e 0, 9. Para a distribuição Exponencial, foram escolhidos os seguintes valores para o parâmetro do modelo: 0, 1, 1, 2, 6 e 10. Para cada um desses valores escolhidos para o parâmetro dos modelos Geométrico e Exponencial foram simuladas amostras aleatórias de tamanho 10, 30, 100 e 1000. Em todos os cenários, a amostra aleatória da variável U foi gerada a partir do método congruencial com $m = 2^{31} - 1$, $a = 7^5$ e $x_0 = \frac{2^{31}}{2}$. Todas as rotinas foram desenvolvidas no *software* computacional R.

3 Resultados e Discussão

Comparando-se os valores das probabilidades calculadas a partir dos dados simulados com os valores das probabilidades teóricas, verificou-se que o Método da Transformação Inversa simulou, de forma adequada, dados das distribuições Geométrica e Exponencial em todos os cenários estudados.

4 Conclusões

O Método da Transformação Inversa é adequado para a simulação de dados aleatórios das distribuições Geométrica e Exponencial.

Agradecimentos

Ao PROBIC/UNIFAL-MG pela concessão da bolsa de Iniciação Científica.

Referências

- [1] D. Gamerman. *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. Chapman & Hall: Boca Raton, 1997, 245 p.
- [2] S.M. Ross. *Simulation*. 4. ed. Burlington: Academic Press, 2006, 298 p.