Avaliação de Modelos Probabilísticos via MMC (Método Monte Carlo)

Lorí Viali[1] e Adão L. Hentges

Instituto de Matemática, Departamento de Estatística UFRGS
90616-900 - Av. Bento Gonçalves, 9500 - Porto Alegre, RS
e-mail: viali@mat.ufrgs.br, hentges@mat.ufrgs.br

[1] Faculdade de Matemática (FAMAT) - Departamento de Estatística - PUCRS
90619-900 - Av. Ipiranga, 6681 - Porto Alegre, RS
e-mail: viali@pucrs.br

Na estatística, especialmente na estimação, um dos problemas básicos é a determinação da forma funcional da variável aleatória X. O conhecimento apenas do modelo não é suficiente para muitas aplicações. Na grande maioria das aplicações é, ainda, necessário à determinação de suas duas características básicas, quando existentes, o valor esperado e a variabilidade. O valor esperado, anotado por μ ou E(X) é o ponto de equilíbrio da distribuição e a variabilidade σ informa o quanto a variável se concentra em torno deste valor.

Assim se a variável aleatória X tem um comportamento dado por f_X precisamos conhecer, isto é, determinar:

$$\mu = \int_{X(S)} x\, f_X(x)dx$$

$$\sigma^2 = \int_{X(S)} x^2 f_X(x)dx - \left(\int_{X(S)} x\, f_X(x)dx \right)^2$$

ou

$$\sigma^2 + \mu^2 = \int_{X(S)} x^2 f_X(x)dx$$

onde X(S), representa o conjunto de valores da variável X, normalmente um intervalo real e S o espaço amostral do experimento associado.

O problema é que muitos dos modelos probabilísticos utilizados na Estatística não apresentam uma forma fechada, isto é, as integrais acima só podem ser avaliadas por métodos numéricos como a regra do trapézio.

No entanto, a disponibilidade computacional atual permite explorar formas alternativas de solução como a simulação pelo MMC (Método Monte Carlo). Simulações pelo MMC dependem da geração eficaz de uma amostra aleatória dos valores de X.

Neste artigo é apresentada uma revisão da geração de números aleatórios via computacional (números pseudo-aleatórios) e sugerida duas alternativas para estimar o valor de θ onde θ pode representar tanto o valor esperado quanto a variabilidade do modelo considerado. As opções consideradas são amostrar o vetor de "n" valores (X_1, X_2, ..., X_n) diretamente de f_X ou avaliar as integrais acima a partir da transformação de valores gerados através da uniforme no intervalo [0; 1].

Os dois procedimentos são comparados e analisados quanto a convergência e tendenciosidade.

São apresentados exemplos baseados em algumas densidades f_X de forma simples e outros de formas mais complexas. Nas simulações apresentadas o tamanho amostral "n" é elevado de forma a possibilitar uma comparação da precisão das duas alternativas.

Referências

[1] BRATLEY, Paul, FOX, Bennett, L., SCHRAGE, Linus E. A Guide to Simulation. New York: Springer-Verlag, 1987. 377 p.

[2] LAW, Averill M., KELTON, W. David. Simulation Modeling and Analysis. New York: McGraw-Hill, 2000. 760 p.

[3] L'ECUYER, P. Random number generation. Draft for a chapter of the forthcoming Handbook of Computational Statistics. (J. E. Gentle, W. Haerdle and Y. Mori eds.). New York: Springer-Verlag, 2004.