Proceeding Series of the Brazilian Society of Applied and Computational Mathematics, Vol. 2, N. 1, 2014.
Trabalho apresentado no CMAC-Sul, Curitiba-PR, 2014.

Estudo de Distribuições de Probabilidade: Simulação e Aplicação

Raí Oliveira Bueno da Silva*
Depto de Físico Química, IQ, UNESP.
14800-900, Araraquara, SP
E-mail: raibueno@ymail.com

Jorge M. V. Capela **Marisa V. Capela**
Depto de Físico Química, IQ, UNESP.
14800-900, Araraquara, SP
E-mail: marisavc@iq.unesp.br

RESUMO

Muitos fenômenos têm associados a eles uma ou mais variáveis aleatórias, cujos valores possíveis são descritos por uma distribuição de probabilidades [6]. A distribuição de probabilidades de uma variável aleatória especifica um modelo probabilístico ou estocástico para o fenômeno em questão. Então, para entender o comportamento de uma variável aleatória precisamos definir a forma de sua distribuição de probabilidade.

Uma variável aleatória discreta tem distribuição de Bernoulli quando ela representa um experimento cujo resultado pode ser um sucesso (se ocorrer o evento de interesse) ou um insucesso (o evento de interesse não ocorre). A probabilidade de sucesso é p e a de insucesso é q=1-p. Nessas condições se a variável aleatória X tem distribuição de Bernoulli a sua função de probabilidade é dada por [5]:

$$P(X = x) = p^x q^{1-x}.$$

Uma variável tem distribuição Binomial quando representa a execução de n vezes um experimento de Bernoulli, sendo cada execução independente da outra:

$$P(X = k) = C_{n,k}\, p^x q^{1-x}.$$

A distribuição mais simples para uma variável aleatória é a distribuição Uniforme. A distribuição Normal ou Gaussiana é a mais familiar das distribuições de probabilidade contínua e também uma das mais importantes cuja função densidade de probabilidade é dada por [5]:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}.$$

O objetivo deste trabalho é estudar e compreender as distribuições de probabilidade discretas e contínuas, suas propriedades e particularidades empregando simulação.

A simulação computacional de qualquer fenômeno aleatório envolve a geração de variáveis aleatórias com distribuições pré-definidas. Uma vez que um modelo de distribuição de probabilidade tenha sido escolhido, um algoritmo para geração da variável aleatória deve ser utilizado. Podemos empregar o método de Monte Carlo de simulações [4,5] para retirar amostras de distribuições de probabilidades e deduzir distribuições de probabilidades amostrais de parâmetros de interesse, como média, desvio padrão, proporções específicas.

Quanto à simulação de Monte Carlo de valores das distribuições de probabilidades, a planilha de cálculo [3] fornece um conjunto de ferramentas para análise de dados — denominado Ferramentas de análise. A ferramenta de análise *Geração de números aleatórios* preenche um intervalo com números aleatórios independentes tirados de uma, dentre várias distribuições. Entretanto, o número de distribuições disponíveis é pequeno e, além disso, para o estudo de distribuições amostrais, precisam ser gerados um conjunto muito grande de amostras. O que se propõe neste trabalho é a possibilidade de gerar números de uma determinada distribuição por fórmulas, portanto, mais prático.

Com um gerador de números aleatórios, na planilha será usada a função ALEATÓRIO(), podemos, em princípio, gerar números aleatórios de outras variáveis usando a correspondente função de distribuição acumulada.

* IC - PROPe/CDC-Fundunesp

Assim, para gerar valores x de uma distribuição de Bernoulli com probabilidade de sucesso *p*, basta gerar um número aleatório u. Se u<p, então x =0, caso contrário, x = 1. O número de sucessos num experimento de Bernoulli com n repetições e probabilidade de sucesso p tem distribuição Binomial de parâmetros *n* e *p*. Logo, a distribuição Binomial pode ser construída repetindo-se o procedimento de construção da distribuição de Bernoulli.

A distribuição contínua mais importante é a distribuição Normal ou de Gauss. Há vários métodos para gerar variáveis aleatórias normais. Um método eficiente é o de Box-Muller [1]. Nesse método são geradas duas variáveis aleatórias independentes com distribuição normal de média 0 e variância 1:

$$z_1 = \text{raiz}(-2.\text{Ln}(u)).\cos(2.\ \pi.v) \text{ e } z_2 = \text{raiz}(-2.\text{Ln}(u)).\text{sen}(2.\pi.v), \qquad (1)$$

onde u e v são variáveis aleatórias com distribuição Uniforme no intervalo de 0 a 1. A distribuição Uniforme é gerada facilmente pela função ALEATÓRIO(). Valores da variável x com distribuição normal de média μ e variância σ^2 são obtidos por $x_1= \mu+z_1.\sigma$ ou $x_2= \mu+z_2.\sigma$.

O estudo das variáveis aleatórias foi útil e importante para a construção dos modelos probabilísticos de algumas situações experimentais [2] e também para a consequente estimação de seus parâmetros. Conclui-se ainda que o emprego de planilhas de cálculo em conjunto com a simulação foi uma alternativa viável e interessante para representar situações reais, ou ainda para descrever experimentos aleatórios.

Palavras-chave*: Simulação, Planilha de Cálculo, Distribuições de probabilidade.*

Referências

[1] Box, G.E.M.; Muller, M.E. A note on the generation of random normal deviates. Ann. Math. Statist. n. 29, pp. 610-611, (1958).

[2]Capela, M. **V.** ; Acciari, H A ; Capela, J. M. V.; Carvalho, M T ; Melin, M C S . Repeatability of corrosion parameters for titanium-molybdenum alloys in 0,9% NaCl solution. Journal of Alloys and Compounds, vol. 465, pp. 479-483, (2008).

[3] Carey, P. Berk, K. Data analysis with Microsoft Excel. Duxbury Press, 2000.

[4] Grafen, A., Hails, R. Modern Statistics for the Life Sciences. New York: Oxford University Press, 2002.

[5]Morettin, L. G.A. Estatística Básica Probabilidade e Inferência. São Paulo, Pearson Prentice Hall, 2010.

[6] Triola, M.F. Introdução à Estatística, 7ed Livros Técnicos e Científicos Editora S.A.,2008.

