Proceeding Series of the Brazilian Society of Applied and Computational Mathematics, Vol. 3, N. 2, 2015.
Trabalho apresentado no III CMAC - SE, Vitória-ES, 2015.

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Resolução Computacional de Problemas de Probabilidade

Lúcio Souza Fassarella[1]
Departamento de Matemática Aplicada, UFES, São Mateus, ES

Resumo. Neste artigo discuto brevemente a essência de uma proposta para a resolução de problemas de probabilidade através de simulações computacionais de ensaios aleatórios, tendo em perspectiva sua aplicação no ensino de matemática. Além das definições básicas, apresento dois exemplos simples que ilustram a técnica.

Palavras-chave. Probabilidade, Simulação Computacional, Ensino de Matemática

1 Introdução

"*Os computadores liberam a matemática do mundo real do cálculo manual, permitindo que ela possa ir mais rápido e mais longe do que jamais alguém poderia ter imaginado. Agora, é vital que a educação matemática também incorpore essa automação.*"
Conrad Wolfram (Tradução livre)

Vivemos uma era em que o uso do computador está bastante difundido em diversos setores da sociedade, e mesmo em nossas próprias vidas. Dadas suas amplas potencialidades, é natural pensarmos no emprego do computador na educação, particularmente na educação matemática. Consonante, o uso da informática tem sido uma das fortes tendências em Educação Matemática e vem ganhando espaço cada vez maior nas práticas pedagógicas [1].

O uso do computador no ensino da matemática é especialmente promissor devido à Matemática ter como essência a *elaboração e resolução de problemas* – conforme disse Paul Halmos num famoso artigo em 1980:

> "Eu acredito que os problemas são o coração da Matemática e espero que nós professores, nas aulas e seminários e nos livros e artigos que escrevermos, enfatizemos isso cada vez mais, e que treinemos nossos estudantes a serem melhores elaboradores e solucionadores de problemas do que nós mesmos somos." [2] (Tradução livre)

Neste artigo discuto a abordagem algorítmica para resolução de problemas de probabilidade pela simulação computacional de ensaios aleatórios – denominada simplesmente

[1]lucio.fassarella@ufes.br

resolução computacional de problemas de probabilidade. A ideia está fundada diretamente na interpretação frequentista das probabilidades, o que suscita a expectativa de que esta abordagem pode realmente ajudar os estudantes a compreender melhor a teoria das probabilidades, além de ampliar sua capacidade para interpretar e resolver problemas específicos.

O uso do computador na resolução de problemas matemáticos pode ocorrer de duas formas diferentes, pelo menos: (i) como *ferramenta auxiliar* para executar algoritmos preestabelecidos utilizados numa estratégia de resolução concebida sem o subsídio do computador ou (ii) como *instrumento cognitivo* que nos permite conceber estratégias de resolução alternativas. Talvez não seja essencial distinguir essas duas formas de usar o computador na prática, até porque elas podem colaborar, se sobrepor ou se confundir em diversas situações; entretanto, ilustro as ideias com dois exemplos: o computador é usado como ferramenta auxiliar na *resolução numérica de uma equação de quinto grau através de um algoritmo genérico (tal como o método de Newton)*, enquanto ele é usado como instrumento cognitivo na *investigação das propriedades de uma configuração geométrica mediante um software de geometria dinâmica.* Chamo de *abordagem algorítmica* a tentativa de resolver um problema pela elaboração de um algoritmo específico, cuja execução gera a solução exata ou uma aproximação arbitrariamente próxima da solução exata. A abordagem algoritmica é um caso especial do uso do computador como instrumento cognitivo que geralmente também o aplica como ferramenta auxiliar (para implementar os algoritmos-solução obtidos).

Trabalhos recentes indicam a relevância do uso do computador ou da computação simbólica no ensino de Probabilidade:

> "*Graças à abundância de computadores de baixo custo, agora temos a oportunidade de melhorar o ensino de matemática em uma frente vastamente mais ampla e interessante, que irá melhorar a importância da matemática aprendida pelos alunos. Por exemplo, com a ajuda do método de Monte Carlo, e programas como o Mathematica, a Probabilidade pode assumir um lugar proeminente na Matemática Escolar. [Nesse caso,] eliminamos a opressora dificuldade com os cálculos combinatórios. [Além disso,] medidas importantes relacionadas com a distribuição normal (de Gauss) e outras distribuições exponenciais [também] podem ser abordadas pelo método de Monte Carlo, tornando essa parte da matemática não mais complicada do que [mera] contagem.*" [4, p.3] (Tradução livre)

Embora a abordagem pareça mais indicada para o Ensino Médio ou Superior, uma pesquisa "realizada com o objetivo de investigar as contribuições que a inserção da tecnologia pode trazer à educação estocástica" e que teve como questão descobrir "como os recursos tecnológicos podem ser úteis para a construção de novos conhecimentos da Estocástica no Ensino Fundamental" chegou à seguinte conclusão:

> "Tornou-se evidente que a inserção de tais recursos gera conhecimentos mais amplos e precisos, porém exige do professor um conhecimento teórico-

metodológico muito mais aprofundado sobre o assunto. Além disso, os resultados destacaram a importância da simulação e do processo de interação na educação estocástica." [3, p.6]

A estrutura do artigo é a seguinte: na Seção 2 defino genericamente a resolução computacional de problemas de probabilidade, na Seção 3 discuto dois exemplos e na Conclusão apresento alguns comentários.

2 Resolução Computacional de Problemas de Probabilidade

Preliminarmente esclareço que *simulação computacional de um ensaio aleatório* significa *usar um gerador de números pseudo-randômicos para selecionar um elemento de um conjunto finito dado*. Aqui, denoto por $Random\,(U)$ a simulação computacional de um ensaio aleatório em U.

A resolução computacional para o caso de um problema que se resume a *determinar a probabilidade de um evento E num espaço amostral finito U no qual todos os elementos são igualmente prováveis*, pode ser esquematizada em quatro etapas:

1. definir o espaço amostral U de modo adequado para realizar ensaios pseudo-ramdômicos;
2. caracterizar o evento E por um conjunto finito de propriedades em U;
3. realizar uma sequência finita N de simulações de ensaios aleatórios em U, contando o número n de vezes nos quais os ensaios resultam num elemento do evento E;
4. calcular a probabilidade de E pelo quociente
$$p\,(E) = \frac{n}{N}.$$

Naturalmente, pode-se escrever um pseudo-algoritmo para calcular a probabilidade de E utilizando a função característica de E (*viz.*, o teste que determina se um elemento de U pertence a E):

$$\varepsilon\,(u) = \begin{cases} 1 & , \; u \in E \\ 0 & , \; u \in U \setminus E \end{cases} ,$$

Explicitamente, o algoritmo-solução do problema é dado por (N é o número de ensaios a serem realizados):

- **Algoritmo-solucão**

  Entrada: $U,\ \varepsilon,\ N$;
  $n = 0$
  Para $i = 1$ **até** $i = N$ **faça**
      Se $\varepsilon\,(Random\,(U)) = 1$ **então** $n = n + 1$;
  Retorne: n/N

DOI: 10.5540/03.2015.003.02.0129

Destaco que a execução desse algoritmo-solução não resulta no valor exato da probabilidade de E, mas numa aproximação cujo erro pode ser arbitrariamente reduzido aumentando o número N de ensaios. Usando Estatística, pode-se determinar o número necessário de ensaios para que, com uma probabilidade arbitrariamente alta, o erro da estimativa seja menor do que uma margem previamente estabelecida. Entretanto, não desenvolvo esse tópico nesta breve exposição.

Nos exemplos que seguem ilustro concretamente a abordagem, utilizando com adaptações o algoritmo-solução genérico proposto acima. No primeiro exemplo calculo uma probabilidade simples, enquanto no segundo resolvo um problema de probabilidade condicional. Em ambos os casos, considero somente espaços de probabilidade finitos nos quais todos os elementos são igualmente prováveis.

3 Exemplos

3.1 O problema dos aniversários coincidentes

- **Problema**: Dado um inteiro $m \geq 2$, qual é a probabilidade de um grupo de m pessoas possuir pelo menos duas pessoas que fazem aniversário num mesmo dia (supondo que todas as possíveis datas de aniversário sejam igualmente prováveis)?

Resolução

Considere que o ano possui 365 dias e associe biunivocamente cada dia do ano a um número entre 1 e 365, o espaço amostral do problema é o conjunto U de todas as sequências de m números entre 1 e 365 – cada termo representando o dia do aniversário de uma determinada pessoa do grupo. O evento em questão é o subconjunto de todas as sequências de U que possuem pelo menos dois termos iguais. Denotando por $\varepsilon : U \to \{0, 1\}$ a função que para $u \in U$ retorna 0 se a sequência u não possui termos iguais e 1 se a sequência u possui pelo menos dois termos iguais, pode-se escrever o seguinte algoritmo-solução:

- **Algoritmo-solucão**

1: **Entrada**: $m,\ N$;
2: $U = \{\text{sequências com } m \text{ números entre 1 e 365}\}$
3: $n = 0$;
4: $\varepsilon : U \to \{0, 1\}\,,\ \varepsilon(u) = \begin{cases} 0\,, & \text{se } u \text{ não possui termos iguais,} \\ 1\,, & \text{se } u \text{ possui pelo menos dois termos iguais;} \end{cases}$
5: **Para** $i = 1$ **até** $i = N$ **faça**
6: **Se** $\varepsilon\,(Random\,(U)) = 1$ **então** $n = n + 1$;
7: **Retorne**: n/N

Este algoritmo foi programado no software *Mathematica* e retornou os seguintes resultados para $N = 10000$ simulações em cada caso:

- para $m = 2$, a probabilidade computada foi 0.0022 (o valor exato é $1/365 \approx 0.0027$);

- para $m = 3$, a probabilidade computada foi 0.0089 (o valor exato é $1/365 + 2 \times 364/365^2 \approx 0.0082$);
- para $m = 10$, a probabilidade computada foi 0.1152 (o valor exato fica para o leitor determinar...).

3.2 O problema da retirada de bolas nas gavetas

- **Problema**: Um móvel tem três gavetas iguais. Em uma gaveta há duas bolas brancas, em outra há duas bolas pretas, e na terceira há uma bola branca e outra preta. Abrimos uma gaveta ao acaso e tiramos uma bola ao acaso sem olhar a segunda bola que está na gaveta. A bola que tiramos é branca. Qual é a probabilidade de que a segunda bola que ficou sozinha na gaveta seja também branca?

Resolução

Represente uma bola preta por 0 e uma bola branca por 1 e considere que as três gavetas têm igual probabilidade de serem escolhidas (antes de realizar a primeira retirada); assim, defina o espaço amostral com elementos igualmente prováveis por

$$U = \{a, b, c\}, \ a = \{(0,0)\}, \ b = \{(0,1), (1,0)\}, \ c = \{(1,1)\}.$$

Pode-se pensar nos ensaios em duas etapas (como descreve o problema): primeiro escolha aleatoriamente uma das gavetas a, b ou c; se a escolha resultar em a ou c, não é necessário fazer nova escolha para determinar a cor da bola que será retirada primeiro (ela será necessariamente preta no caso de a e branca no caso de c); entretanto, se o resultado da escolha da gaveta for b, então realize uma escolha aleatória em $\{0,1\}$ para determinar qual é a cor da primeira bola retirada (e também da segunda, consequentemente). Assim, numa sequência de ensaios aleatórios em U, a informação de que a primeira retirada resultou numa bola branca é levada em conta simplesmente desprezando todos os ensaios que resultam na gaveta a ou resultam na gaveta b com o subsequente ensaio aleatório em $\{0,1\}$ resultando em 0. Naturalmente, é necessário contar separadamente o número de ensaios válidos (aqueles que correspondem à primeira bola retirada ser branca) dos ensaios que resultam no evento em questão (aqueles que correspondem à escolha da gaveta c).

- **Algoritmo-solucão**[2]

Entrada: N;
$n = 0; m = 0$;
Para $i = 1$ **até** $i = N$ **faça**
  $x = Random\,(U)$,
  Se $x = b$ **e** $Random\,\{0,1\} = 1$ **então** $m = m + 1$,
  Se $x = c$ **então** $m = m + 1, \ n = n + 1$;
Retorne: n/m

[2] Destaco que este algoritmo não utiliza a função ε definida na Seção 2; entretanto, algo similar está implícito nas condições de contagem dos índices m e n.

DOI: 10.5540/03.2015.003.02.0129

Este algoritmo foi programado no software *Mathematica* e retornou os seguintes resultados para $N = 1000$, $N = 10000$ e $N = 100000$ simulações, respectivamente: 0.6736, 0.6792 e 0.6642. A resposta exata do problema é $2/3 \approx 0.6666$.

4 Conclusão

Neste artigo, apresento uma abordagem para resolver problemas de probabilidade usando o computador que alinha-se à tendencia atual de usar a computação simbólica na educação matemática e que pode facilmente ser integrada às metodologias de ensino denominadas Resolução de Problemas e Investigação Matemática.

Naturalmente, a resolução computacional não substitui a resolução analítica dos problemas de probabilidade, mas constitui uma abordagem alternativa que eventualmente é mais fácil ou mais atraente de ser implementada. Destaco duas vantagens da resolução computacional: a primeira consiste dela contornar problemas intermediários de combinatória; a segunda, consiste numa padronização das resoluções dos problemas de probabilidade, dada pelo algoritmo-solução genérico definido na Seção 2.

Dentre as perspectivas futuras para o desenvolvimento deste trabalho, está a realização de pesquisas de campo para avaliar o efeito da aplicação da abordagem no aprendizado de probabilidade.

Referências

[1] M. de C. Borba, R.S.R. da Silva e G. Gadanidis, Fases das Tecnologias Digitais em Educação Matemática, Editora Autêntica, (2014).

[2] P. Halmos, The Heart of Mathematics, The American Mathematical Monthly, vol. 87(7), 519–524, (1980).

[3] L. O. Souza, A Educação Estatística no Ensino Fundamental e os recursos tecnológicos, Dissertação de Mestrado em Ensino de Ciências e Matemática, Universidade Cruzeiro do Sul, (2009).

[4] J.J. Uhl and D. Woods, The crisis we face and how to try to deal with it, Symbolic Computation and Education, Chapter 1, World Scientific, (2007).

DOI: 10.5540/03.2015.003.02.0129