

# Resolução de Integrais com Simulações de Monte Carlo: Comparação Com Métodos Analíticos e Numéricos

Pedro Henrique da Rosa Silva  
*Universidade Federal do ABC (UFABC), Santo André, Brasil*  
(Dated: 2025)

Eventos incertos permeiam nosso cotidiano e alguns podem ser mais problemáticos do que outros. Nas áreas da Matemática e da Física encontramos ferramentas para a construção de previsões e, portanto, para uma melhor tomada de decisão. O método de Monte Carlo é uma dessas ferramentas que, ao fazer uso da aleatoriedade, retorna possíveis resultados para diferentes problemas. Neste trabalho, realizamos simulações do método de Monte Carlo para a resolução de integrais. Mostramos que o método, mesmo com variáveis aleatórias, retorna valores muito próximos dos analíticos. Assim, com o domínio do método, podemos replicar tais simulações para problemas sem solução analítica fechada, conhecendo-se apenas a teoria do problema.

## INTRODUÇÃO

Desde a antiguidade nos deparamos com fenômenos de natureza aleatória que influenciam a tomada de decisão e moldam cenários futuros. Em paralelo, a Matemática e a Física desenvolveram ferramentas que permitem formular previsões para eventos incertos. O método de Monte Carlo (MC) foi formalizado na década de 1940 por John von Neumann, Stanislaw Ulam e Nicholas Metropolis no contexto do Projeto Manhattan, em Los Alamos, e recebeu esse nome em alusão ao famoso distrito de cassinos em Mônaco. De modo geral, define-se a população de interesse e seus parâmetros (média, desvio padrão, etc.) e postula-se uma distribuição (Normal, Exponencial, etc.). Em seguida, amostras independentes e identicamente distribuídas são geradas computacionalmente e estatísticas de interesse são aproximadas por médias amostrais.

## METODOLOGIA

### Revisão sucinta de Probabilidade

Nessa etapa buscamos fazer um breve resumo da teoria que se acredita ser necessária para a realização das simulações. Associamos a cada experimento um espaço amostral  $\Omega$  e, para cada evento  $E \subseteq \Omega$ , uma probabilidade  $P(E)$  que satisfaz  $0 \leq P(E) \leq 1$ ,  $P(\Omega) = 1$  e

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i). \quad (1)$$

Uma variável aleatória  $X$  é uma função mensurável de  $\Omega$  em  $\mathbb{R}$ . Para  $X$  contínua, a função distribuição acumulada é  $F(x) = P(X \leq x)$ . Quando a variável aleatória  $X$  assumir um número finito de valores possíveis ou número infinito enumerável, vamos dizer que essa é uma variável aleatória discreta. De um modo geral, os valores de  $X$  pertencem ao conjunto dos inteiros de valores. Nesse caso, com  $X$  assumindo os valores  $x_1, x_2, \dots$ , podemos definir uma função de probabilidade  $p(x_i) = P(X = x_i)$ .

Já quando a variável aleatória puder assumir quaisquer valores dentro de um intervalo de números, vamos dizer que essa é uma variável aleatória contínua. Podemos dizer que elas não podem ser enumeradas (contadas). Aqui, definimos uma função de densidade de probabilidade  $f(x)$ , que descreve a verossimilhança de uma variável aleatória tomar um valor dado, e a função de distribuição acumulada será definida por

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(x) dx. \quad (2)$$

Para uma variável aleatória discreta  $X$  que assume os valores  $x_1, x_2, \dots, x_n$ , define-se o valor esperado  $X$ , representando o valor médio "esperado" de uma experiência se ela for repetida muitas vezes, por

$$E(X) = \sum_{i=1}^n x_i p(x_i) \quad (3)$$

Sendo  $X$  uma variável aleatória contínua, seu valor esperado será

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx. \quad (4)$$

Supondo  $X_1, X_2, \dots, X_n$  uma sequência de variáveis aleatórias cuja média é  $E[x] = \mu < \infty$ . Seja  $S_n = X_1 + \dots + X_n$  uma variável aleatória, a Lei dos Grandes Números nos diz que sequência de médias amostrais  $\frac{S_n}{n}$  converge para  $\mu$  quando  $n \rightarrow \infty$ . Em outras palavras, a média dos resultados de um experimento tende a se aproximar do valor esperado à medida que mais tentativas se sucedem. Então, ao realizar cada vez mais tentativas, mais a probabilidade da média dos resultados observados se aproxima da probabilidade real. Para essa lei existem diferentes versões. Teremos a versão Lei Forte dos Grandes Números quando  $X_1, X_2, \dots$ , satisfizerem

$$\frac{S_n - E(S_n)}{n} \rightarrow 0 \quad (5)$$

isto é

$$P\left(\lim_{n \rightarrow \infty} \frac{S_n}{n} = \mu\right) = 1 \quad (6)$$

### Método de Monte Carlo

Seja  $I = E(X)$  o valor esperado de alguma variável aleatória  $X$ , o valor o qual desejamos conhecer. Sendo possível gerar novos valores de variáveis aleatórias e independentes que possuam a mesma distribuição de probabilidade de  $X$ , vamos realizar uma simulação toda vez que um novo valor for gerado [5]. Com isso, em  $n$  simulações concluídas obtemos  $X_1, X_2, \dots, X_n$ . Assumindo

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (7)$$

como a média de  $X$ , pela Lei Forte dos Grandes Números  $\bar{X}$  pode ser usado para se estimar o valor de  $I$ , ou seja,

$$E(\bar{X}) = \frac{1}{n} \sum_{i=1}^n E(X_i) = I \quad (8)$$

### Desenvolvimento do código para integração unidimensional

Seja  $g(x) : \mathbb{R} \rightarrow \mathbb{R}$  uma função contínua, suponhamos uma integral unidimensional do tipo

$$\int_0^1 g(x) dx \quad (9)$$

com  $X \in [0, 1]$ . Considerando uma variável aleatória  $U$  uniformemente distribuída no intervalo  $[0, 1]$ , o valor esperado de  $g(U)$  será

$$E(g(U)) = \int_0^1 g(x) f_U(x) dx = \int_0^1 g(x) dx = I \quad (10)$$

onde  $f_U(x)$  é uma função de densidade de probabilidades definida por  $f_U(x) = 1$  se  $x \in [0, 1]$ , e  $f_U(x) = 0$  caso contrário.

Supondo então variáveis aleatórias uniformes independentes  $U_1, U_2, \dots, U_n$ , teremos que todas as variáveis aleatórias  $g(U_1), g(U_2), \dots, g(U_n)$  são independentes e identicamente distribuídas com média  $I$ , isto é,

$$\int_0^1 g(x) dx = E[g(U)] \cong \frac{1}{n} \sum_{i=1}^n g(U_i) \quad (11)$$

Portanto, pela Lei Forte dos Grandes Números, temos que, com probabilidade 1,

$$\frac{1}{n} \sum_{i=1}^n g(U_i) \rightarrow I = E[g(U)]. \quad (12)$$

Sendo assim, vamos aproximar  $I$  ao gerar uma grande quantidade de números aleatórios  $u_i$  e considerar como aproximação o valor médio de  $g(u_i)$ .

Sabendo que os limites de integração em diferentes problemas não necessariamente são iguais a 0 e 1, podemos realizar substituições de variáveis. Para calcular a integral

$$\int_a^b g(x) dx \quad (13)$$

escolhemos  $y = \frac{(x-a)}{(b-a)}$  e  $dy = \frac{dx}{(b-a)}$ . Assim, obtém-se

$$I = \int_0^1 g[a + (b-a)y](b-a) dy = \int_0^1 h(y) dy \quad (14)$$

onde  $h(y) = (b-a)g[a + (b-a)y]$ .

Com isso, vamos gerar números aleatórios e tomar o valor médio de  $h$ , calculado nesses números, como aproximação de  $I$ . As integrais que se enquadram nesse caso são:

$$\begin{aligned} & \int_1^3 \left(5 - x + \frac{x^2}{2} + \frac{x^3}{3}\right) dx \\ & \int_1^4 \left(\frac{1}{x^2} - \frac{2}{x^2}\right) dx \\ & \int_1^5 x \ln x dx \\ & \int_0^5 x e^{-x} dx \end{aligned} \quad (15)$$

Já para uma integral com forma

$$\int_0^\infty g(x) dx \quad (16)$$

escolhemos  $y = \frac{1}{(b-a)}$  e  $dy = -\frac{dx}{(b-a)^2} = -y^2 dx$ . Nesse caso, obtém-se

$$I = \int_0^\infty h(y) dy \quad (17)$$

onde  $h(y) = \frac{g(\frac{1}{y}-1)}{y^2}$ . Aqui também vamos gerar números aleatórios  $y_i$  e tomar o valor médio de  $h_i$  como aproximação de  $I$ :

$$= \frac{1}{n} \sum_{i=1}^n h(y_i). \quad (18)$$

As integrais para a análise nesse caso são:

$$\int_0^\infty \frac{1}{x^2 - 1} dx \quad (19)$$

$$\int_0^\infty e^{-x} \sin(x) dx$$

### Regra do Trapézio

Seja  $f(x) : \mathbf{R} \rightarrow \mathbf{R}$  uma função contínua, suponhamos uma integral unidimensional do tipo

$$\int_a^b f(x) dx \quad (20)$$

com  $x \in [a, b]$ . Suponhamos que o intervalo  $[a, b]$  seja dividido em  $J$  subintervalos de tamanho  $h$ , de modo que  $h = (b - a)/J$ . O lado direito do  $k$ -ésimo subintervalo fica em  $a + kh$ , e o lado esquerdo fica em  $a + kh - h = a + (k - 1)h$ . Assim, a área do trapézio para esta fatia é

$$A_k = \frac{h}{2} [f(a + (k - 1)h) + f(a + kh)] \quad (21)$$

Esta é a regra do trapézio. Ela nos dá uma aproximação trapezoidal para a área sob um subintervalo de nossa função. Para calcular uma aproximação a área entre o intervalo  $[a, b]$  basta somar todas as áreas dos trapézios entre  $[a, b]$ , isto é:

$$\int_a^b f(x) dx \cong \frac{h}{2} [f(a) + f(b) + 2 \times \sum_{k=1}^J f(a + kh)] \quad (22)$$

### Regra de Simpson

Seja  $f(x) : \mathbf{R} \rightarrow \mathbf{R}$  uma função contínua, suponhamos uma integral unidimensional do tipo

$$\int_a^b f(x) dx \quad (23)$$

com  $x \in [a, b]$ . Desta vez, por fins de argumentação, suponha que temos três pontos em  $x = -h, 0$ , e  $+h$ . Se ajustarmos uma quadrática  $Ax^2 + Bx + C$  a esses pontos, então por definição teremos:

$$\begin{aligned} f(-h) &= Ah^2 - Bh + C, \\ f(0) &= C, \\ f(h) &= Ah^2 + Bh + C. \end{aligned} \quad (24)$$

Resolvendo essas equações simultaneamente para  $A, B$ , e  $C$  obtemos:

$$\begin{aligned} A &= \frac{1}{2h^2} [f(-h) + 2f(0) + f(h)] \\ B &= \frac{1}{2h} f(h) - f(-h) \\ C &= f(0) \end{aligned} \quad (25)$$

e a área sob a curva  $f(x)$  entre  $-h$  e  $h$  é dada pela aproximação da área abaixo da parábola

$$\int_{-h}^h (Ax^2 + Bx + C) dx = \frac{h}{3} [f(-h) + 4f(0) + f(h)] \quad (26)$$

Como estamos integrando de  $x = a$  a  $x = b$  em subintervalos de largura  $h$ . Então, os três pontos que delimitam o primeiro par de subintervalos caem em  $x = a, a + h$  e  $a + 2h$ . Aqueles que delimitam o segundo par ficam em  $a + 2h, a + 3h, a + 4h$ , e assim por diante. Desta forma, o valor aproximado da integral completa é dado por:

$$\int_a^b f(x) dx \cong \frac{h}{3} [f(a) + f(b) + 4 \sum_{\substack{k=1 \\ \text{odd}}}^{J-1} f(a + kh) + 2 \sum_{\substack{k=2 \\ \text{even}}}^{J-2} f(a + kh)] \quad (27)$$

Note que o número total  $J$  de subintervalos deve ser um número par para que a Regra de Simpson calcule as aproximações corretamente. Para fins de simplicidade, tanto na Regra do Trapézio quanto na Regra de Simpson foram utilizados  $J = 10$  subintervalos.

Definidos todos os métodos numéricos de integração, podemos calcular as integrais através de códigos em Python, disponibilizados no [GitHub](#) do autor.

Para fins de comparação, após calcular as integrais numericamente através dos métodos citados acima vamos calcular o erro absoluto entre o valor analítico da integral numérica e o valor obtido numericamente em cada método de integração.

## RESULTADOS

Os algoritmos foram implementados em Python e aplicados a integrais com solução analítica conhecida. A Tabela I em apêndice compara os valores analíticos com as estimativas do método de Monte Carlo (MC) para diferentes tamanhos de amostra  $N$  juntamente com os

métodos da Regra do Trapézio (RT) e Regra de Simpson (RS).

Dados os valores da Tabela I, podemos calcular os erros absolutos entre as integrais analíticas e os métodos numéricos, isto é,  $E_A = |I_A - I_N|$ , sendo  $E_A$  o erro absoluto,  $I_A$  o valor da integral obtido analiticamente e  $I_N$  o valor da integral obtido numericamente. A tabela II em apêndice mostra os erros relativos citados acima. Partindo da tabela II, podemos observar a precisão dos valores obtidos através das integrais numéricas com os gráficos à seguir.

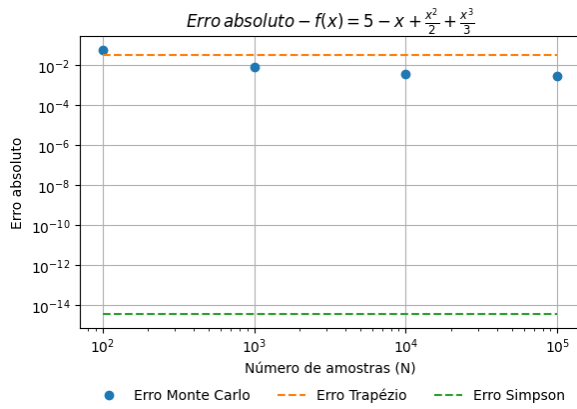


Figura 1:  
Erro absoluto em função do número N de amostras

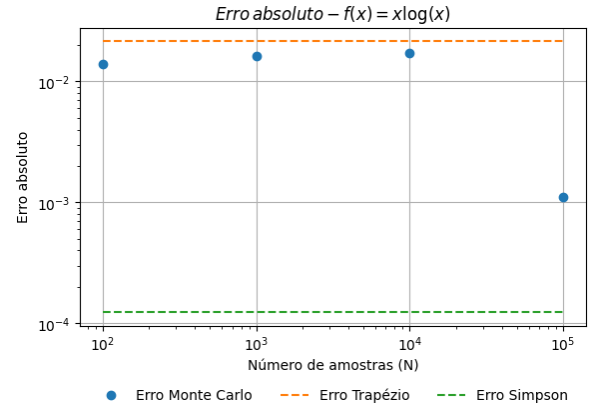


Figura 3:  
Erro absoluto em função do número N de amostras

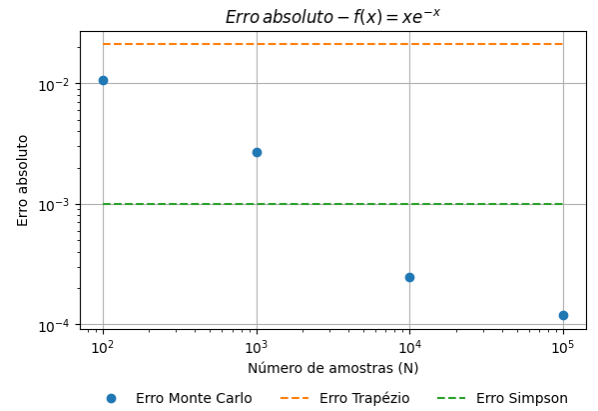


Figura 4:  
Erro absoluto em função do número N de amostras

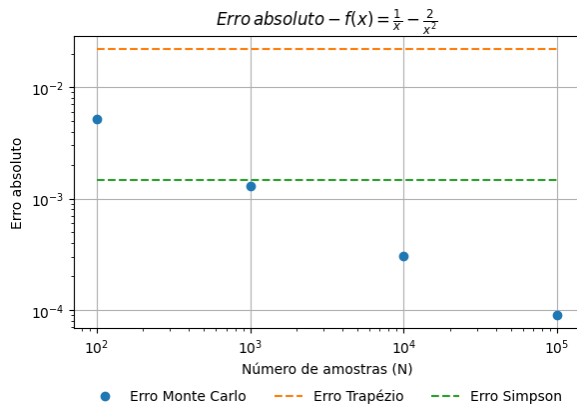


Figura 2:  
Erro absoluto em função do número N de amostras

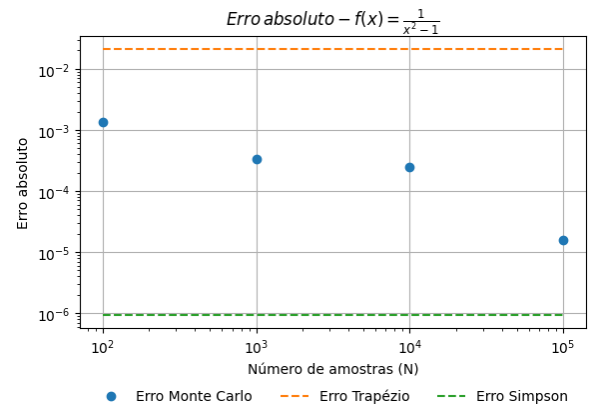


Figura 5:  
Erro absoluto em função do número N de amostras

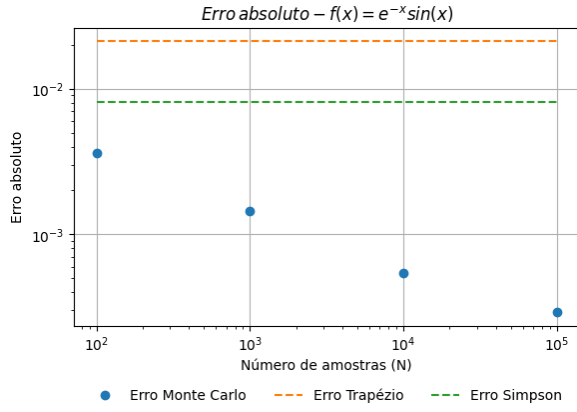


Figura 6:

Erro absoluto em função do número  $N$  de amostras

Os gráficos acima comparam o erro absoluto da Regra do Trapézio (Erro Trapézio) e o erro absoluto da Regra de Simpson (Erro Simpson) com o erro absoluto do método de Monte Carlo (MC) em função do número  $N$  de amostras (números aleatórios).

## DISCUSSÃO

Com auxílio da Tabela II, é possível observar que o método de Monte Carlo para integrais apresenta um erro absoluto em relação aos valores calculados analiticamente da ordem de  $10^{-2}$  para  $N = 10^2$  simulações e erros de ordem até  $10^{-5}$  para algumas funções com  $N = 10^5$  simulações, como mostra a tabela II. Com auxílio dos gráficos 1 à 6 é possível observar que o valor do erro absoluto do método de Monte Carlo é inversamente proporcional ao número de  $N$  de amostras (simulações), com exceção do gráfico 3, que demonstrou erros absolutos maiores para  $N = 10^3$  e  $N = 10^4$  em relação ao erro absoluto para  $N = 10^2$ ). Esse comportamento atípico em relação às demais integrais pode estar relacionado ao caráter aleatório das simulações.

Os erros absolutos da Regra do Trapézio e da Regra de Simpson serviram como referência para o método de Monte Carlo devido ao fato de que suas precisões não dependem do número  $N$  de simulações do método de Monte Carlo e apresentam erros relativamente baixos para  $J = 10$  subintervalos.

Sendo assim, o método de Monte Carlo demonstrou-se eficaz para aproximar integrais unidimensionais, retornando valores próximos aos analíticos mesmo para tamanhos de amostra moderados. Dado seu caráter flexível, o método não depende da dimensão das funções a serem calculadas e é uma excelente ferramenta numérica para problemas sem solução fechada ou de alta dimensionalidade.

- 
- [1] A. Gadelha, *Notas de Aula Teoria de Probabilidade I: Uma pequena história da probabilidade* (DME/IMU/UFRJ, 2004).
  - [2] C. Allen, *Introduction to Monte Carlo Methods* (Computational Science Education Project, Free Software Foundation, 1995).
  - [3] S. Raychaudhuri, “Introduction to Monte Carlo Simulation,” in *Proceedings of the 2008 Winter Simulation Conference* (IEEE, 2008).
  - [4] S. M. Ross, *Introduction to Probability Models*, 11th ed. (Academic Press, 2014).
  - [5] R. Y. Rubinstein and D. P. Kroese, *Simulation and the Monte Carlo Method*, 3rd ed. (Wiley, 2017).
  - [6] M. E. J. Newman, *Computational Physics* (CreateSpace Independent Publishing Platform, 2013).
  - [7] P. H. R. Silva, *Resolução de Integrais com Simulações de Monte Carlo*, (2024)

## APÊNDICE

Tabela I: Comparação entre soluções analíticas, métodos numéricos e método de Monte Carlo (MC).

Integral	Solução Analítica	RT	RS	MC			
				N=10 <sup>2</sup>	N=10 <sup>3</sup>	N=10 <sup>4</sup>	N=10 <sup>5</sup>
$\int_1^3 (5 - x + \frac{x^2}{2} + \frac{x^3}{3}) dx$	17	17,033	16,999	16,938	17,008	16,996	16,997
$\int_1^4 (\frac{1}{x} - \frac{2}{x^2}) dx$	-0,114	-0,135	-0,115	-0,119	-0,112	-0,114	-0,114
$\int_1^5 x \ln x dx$	14,118	14,139	14,118	14,131	14,134	14,101	14,117
$\int_0^5 x e^{-x} dx$	0,9595	0,9384	0,9585	0,949	0,957	0,960	0,959
$\int_0^\infty \frac{1}{x^2+1} dx$	1,5708	1,5674	1,5708	1,5694	1,5705	1,5710	1,5708
$\int_0^\infty e^{-x} \sin x dx$	0,5	0,4925	0,5081	0,5036	0,5014	0,5005	0,4997

Tabela II: Erros absolutos entre soluções analíticas, Regra do Trapézio (RT), Regra de Simpson (RS) e método de Monte Carlo (MC).

Integral	Erro Absoluto - $E_A =  I_A - I_N $					
	RT	RS	MC			
			N=10 <sup>2</sup>	N=10 <sup>3</sup>	N=10 <sup>4</sup>	N=10 <sup>5</sup>
$\int_1^3 (5 - x + \frac{x^2}{2} + \frac{x^3}{3}) dx$	0,03333	$3,55 \times 10^{-15}$	0,06216	0,00837	0,00386	0,00292
$\int_1^4 (\frac{1}{x} - \frac{2}{x^2}) dx$	0,02206	0,00145	0,00516	0,00130	0,00030	$9,04 \times 10^{-5}$
$\int_1^5 x \ln x dx$	0,02142	0,00012	0,01384	0,01610	0,01707	0,00110
$\int_0^5 x e^{-x} dx$	0,02113	0,00099	0,01061	0,00270	0,00025	0,00011
$\int_0^\infty \frac{1}{x^2+1} dx$	0,00333	$9,39 \times 10^{-7}$	0,00136	0,00033	0,00024	$1,59 \times 10^{-5}$
$\int_0^\infty e^{-x} \sin x dx$	0,00751	0,00815	0,00364	0,00145	0,00055	0,00029