Variações do método de Monte Carlo

Matheus Oliveira da Silva

Abril 2023

1 Introdução

A ideia por trás da integração por métodos de Monte Carlo é o cálculo da integral [1]

$$\int_{\omega} f(x)\, dx \tag{1}$$

através do uso de amostras aleatórias. Em sua forma básica, consiste na geração de N pontos $X_1, X_2, ..., X_N$ uniformemente distribuídos, para o cálculo da estimativa

$$F_N = \frac{1}{N}\sum_{i=1}^{N} f(x_i) \tag{2}$$

que aproxima o valor real da integral.

Dada a natureza aleatória do processo de estimação, foram desenvolvidas ao longo dos anos variações do método, de forma a reduzir a variância do estimador, diminuindo o número de iterações necessárias (N) para convergência da estimativa.

O presente trabalho visa apresentar a aplicação de quatro variações do método de Monte Carlo, comparando seus resultados e velocidades de convergência, para o cálculo da integral da função

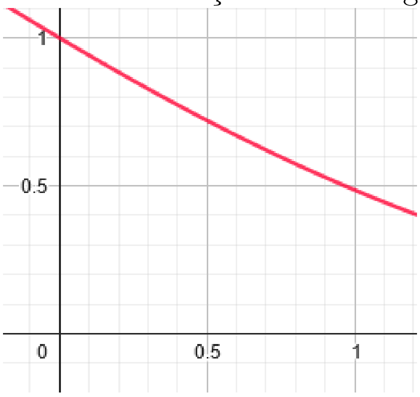
$$f(x) = e^{-ax} cos(bx) \tag{3}$$

onde $a = 0.592216676$ e $b = 0.49752833870$

2 Desenvolvimento

A função que será integrada, apresentada na fórmula 3, pertence a região do quadrado unitário, como pode ser observado na imagem abaixo

Figura 1: Gráfico da função sendo integrada em 3



Os métodos que serão estudados são

- Crude Monte Carlo
- Monte Carlo Hit or Miss
- Importance Sampling
- Control Variates

2.1 Calculo do tamanho amostral (N)

De forma a garantir um erro relativo $\epsilon \leq 0,05\%$, precisamos encontrar um tamanho amostral (N), para cada um dos métodos apresentados, considerando seus respectivos desvios. Isto é, encontrar N tal que

$$|\frac{y - \hat{y}}{y}| \leq 0,0005 \tag{4}$$

onde y é o valor real da integral e $\hat{y}$ sua estimativa

Assim, como apresentado por Bussab e Moretin em [2], pela Lei dos Grandes Números a média amostral $\hat{X}$ converge para o valor real da esperânça de uma variável aleatória X, enquanto que a distribuição amostral, pelo Teorema do Limite Central, é aproximadamente normal, com média $E[X]$ e variância σ^2.

Dessa forma, podemos estimar o tamanho amostral N utilizando a fórmula apresentada em [2] (capítulo 10, seção 11)

$$N = \frac{16\sigma^2 Z^2}{\epsilon^2} \tag{5}$$

onde Z é o valor de uma distribuição $N(0,1)$, que garante uma probabilidade de 99% da estimativa estar dentro do intervalo, para o erro buscado ϵ (considerando um intervalo de confiânça bilateral).

Como o valor real da variância populacional σ^2 e o erro relativo $0,0005 * y$ são desconhecidos, serão estimados valores iniciais para ambas as variáveis, considerando uma amostra inicial de tamanho 1000. Isto é, o tamanho amostral para cada método utilizado será estimado como

$$\hat{N} = \frac{16S^2Z^2}{(\hat{y} * 0,0005)^2} \tag{6}$$

onde S^2 é a variância amostral e $\hat{y}$ é o valor estimado para a integral, considerando o método utilizado.

2.2 Crude Monte Carlo

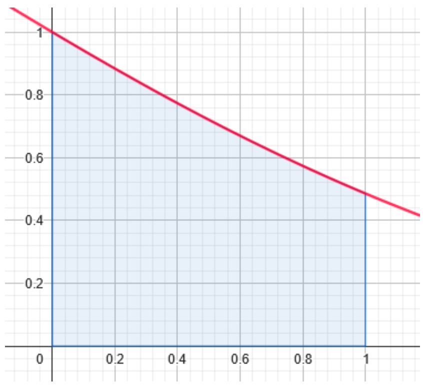
Para o método de Crude Monte Carlo, são geradas N variáveis aleatórias independentes e identicamente distribuidas $X_i \sim U(0,1)$, calculando-se a integral da função (equação 3) como sendo a média destes valores

$$\hat{\gamma} = \frac{1}{N}\sum_{i=1}^{N} f(x_i) \tag{7}$$

2.3 Monte Carlo Hit or Miss

O método de Monte Carlo Hit or Miss consiste no cálculo da proporção de pontos, sorteados aleatóriamente na região [0,1]x[0,1], que pertencem a área abaixo da curva da função (ver figura 2).

Figura 2: Área abaixo da curva da função



Dessa forma, são sorteados N pontos e a aproximação da integral da função buscada é dada por

$$\hat{\gamma} = \frac{\sum_{i=1}^{N} h(x_i, y_i)}{N} \tag{8}$$

onde $h(x, y) = Ind(y \leq f(x))$

2.4 Importance Sampling

O método de Amostragem por Importância considera que será calculada a estimativa para o valor de

$$\gamma = \int f(x)\, dx = \int \frac{f(x)}{g(x)} g(x)\, dx \tag{9}$$

onde as N variáveis aleatórias que serão geradas não possuem mais distribuição $U(0, 1)$, mas sim uma distribuição mais conveniente, com densidade $g(x)$.

Para a integral considerada na equação 9, teremos que

$$\sigma_S^2 = \frac{\int (\frac{f(x)}{g(x)} - \gamma)^2 g(x)}{N} \tag{10}$$

de forma que, se $g(x)$ for próxima de $f(x)$, a variância será reduzida.

É necessário então encontrar uma densidade de probabilidade $g(x)$ que aproxime razoavelmente bem a função $f(x)$. Dado o formato da curva do gráfico de $f(x)$, optou-se por aproximar a função utilizando uma distribuição exponencial $exp(\lambda)$.

Como a distribuição exponencial possui apenas um parâmetro, foi realizado o ajuste do gráfico testando alguns valores para λ. Dado que a distribuição exponencial é definida para o intervalo $[0, +\infty]$, precisamos truncar a função, de forma que a área sobre a curva entre [0,1] sejá igual a 1. Isto foi realizado considerando-se uma constante c, calculada como

$$c = \frac{1}{\int_0^1 \lambda e^{-\lambda x}\, dx} = \frac{1}{1 - e^{-\lambda}} \tag{11}$$

Para $\lambda = 0,5$, a constante calculada foi de $c = 2,5415$, enquanto que para $\lambda = 1$, a constante foi de $c = 1,5819$. Exibindo os ajustes no gráfico, considerando que a função é truncada (insto é, $g(x) = (c)(\lambda e^{-\lambda x})$), temos que

Figura 3: Comparação com $\lambda = 0,5$

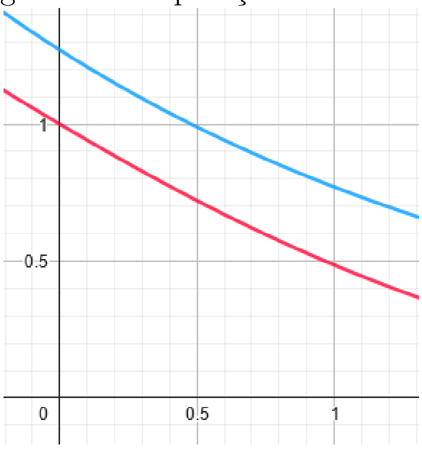
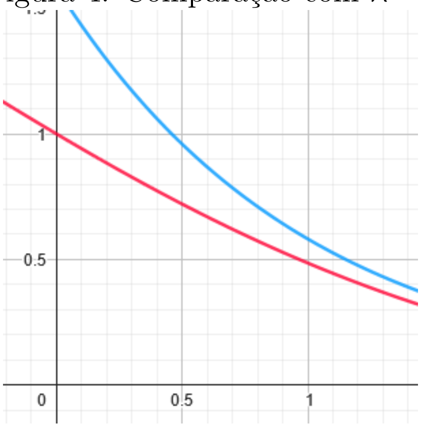


Figura 4: Comparação com $\lambda = 1$



É possível notar que, para ambos os valores de λ, o ajuste segue a mesma tendência de $f(x)$. Entretanto, optou-se por utilizar o valor de $\lambda = 0,5$, observando que a distância entre as curvas é a mesma ao longo de todo o intervalo $[0,1]$. A estimativa para o valor real da integral é calculada como sendo

$$\hat{\gamma} = \frac{1}{N} \sum_{i=1}^{N} \frac{f(x_i)}{g(x_i)} \tag{12}$$

2.5 Control Variates

O método de Váriavel Controle consiste na utilização de uma função $\varphi(x)$ correlacionada com $f(x)$ e que seja integrável analíticamente, de forma que seja calculado

$$\gamma = \int f(x) - \varphi(x) + \varphi(x)\, dx = [\int f(x) - \varphi(x)\, dx] + \omega \tag{13}$$

onde

$$\omega = \int \varphi(x)\, dx \tag{14}$$

Para a integral considerada na equação 13, teremos que

$$\sigma^2_{CV} = \frac{\sigma^2(f(x)) + \sigma^2(\varphi(x)) - 2\rho(f(x), \varphi(x))\sigma(f(x))\sigma(\varphi(x))}{N} \tag{15}$$

de forma que, quanto mais correlacionadas $f(x)$ e $\varphi(x)$, menor a variância encontrada.

Para a função estudada, dado que a curva é levemente linear no intervalo $[0, 1]$ (ver figura 1), optou-se por aproximar via série de Taylor, utilizando apenas a primeira derivada. Dessa forma, temos que

$$\varphi(x) = 1 - f^{(1)}(0) = 1 - Ax \tag{16}$$

onde A = 0.592216676

$$f^{(1)}(x) = -e^{-Ax}(Acos(Bx) + Bsen(Bx)) \tag{17}$$

A estimativa para o valor real da integral é calculada como sendo

$$\hat{\gamma} = \frac{\sum_{i=1}^{N} f(x_i) - \varphi(x_i) + \omega}{N} \tag{18}$$

onde ω é a integral da função auxiliar

$$\omega = \int_0^1 \varphi(x)\, dx\omega = \int_0^1 1 - Axdx = 0.703891662 \tag{19}$$

3 Resultados e discussões

Para calcular o tamanho amostral (N) necessário para cada método, foram considerados os valores de $Z = 1.96$ (95% em intervalo de confiânça Bi-caudal) e um tamanho de amostra piloto de 1000. Foram encontrados

Método	Variância	Média	N
Crude	0.0235	0.7341	10699142
Hit or Miss	0.1926	0.74	86470557
Importance Sampling	0.002	0.7260	948576
Control Variate	0.0005	0.7289	248163

Calculando as estimativas para o valor real da integral, com os valores estimados para o tamanho amostral, temos

Método	Estimativa
Crude	0.72832
Hit or Miss	0.72834
Importance Sampling	0.72829
Control Variate	0.72819

4 Conclusão

Em geral, os algoritmos crude e hit or miss possuem os maiores tamanhos amostrais, por serem algoritmos que não carregam nenhuma informação do problema a priori de sua execução, enquanto que os métodos de Importance Sampling e Control Variate tiveram os melhores resultados, visto que buscam reduzir a variância do método de estimação e, assim, diminuem o tamanho da amostra consumindo menos desempenho computacional.

Referências

[1] Whitlock P. A. Kalos, M. H. *Monte Carlo Methods: Basics*. 1986.

[2] Wilton MORETIN, Pedro; BUSSAB. *Estatística Básica*, volume 1. Saraiva, 2012.