Exercício Programa 01 - Cálculo do Valor de π

Alexandre Barsam Junqueira

21 de março de 2024

Resumo

Esse relatório busca desenvolver estimativas para o valor de π por meio de Simulações de Monte Carlo, isto é: (i) gerar pontos uniformemente distribuídos $x_i \in [1,1]^2$, $i \in \{1,\ldots,n\}$; (ii) definir a proporção de pontos cuja distância do centro é menor que 1; (iii) estimar π. Além disso, busca-se entender um valor de n ideal para um $erro \leq 0.05\%$.

1 Introdução ao Método de Monte Carlo

1.1 Algoritmo e Utilização

Criado por Stanislaw Ulam e John Von Neumann, o Método de Monte Carlo ou Simulação de Monte Carlo é um tipo de algoritmo computacional que utiliza amostragens aleatórias repetidas vezes para obter a probabilidade de um grupo de resultados ocorrer. O método é utilizado, principalmente, como forma de obter aproximações numéricas de funções complexas em que não é viável, ou é mesmo impossível, obter uma solução analítica ou determinística. Dessa forma, consegue-se atuar sobre problemas de otimização e integração numérica, por exemplo.

Na maioria dos casos, as simulações tendem a seguir um padrão:

1. Definir um domínio de possíveis entradas;
2. Gerar entradas randomicamente por meio de uma distribuição probabilística sobre o domínio;
3. Obter resultados para cada entrada;
4. Juntar os resultados.

Desse modo, o Método de Monte Carlo é uma ferramenta extremamente poderosa para o entendimento de situações complexas e para aproximações. No entanto, traz consigo problemas relacionados a custo computacional, precisão e dimensionalidade.

1.2 Intervalo de Confiança (IC) e Margem de Erro (e) para Proporções Populacionais

Por se tratar de um modelo probabilístico, não basta somente conduzir uma simulação e obter o resultado, é preciso entender a relevância estatística daquele resultado. Por isso, cabe observar o intervalo de confiança e a margem de erro do modelo.

Quando se deseja entender a proporção p de uma população de tamanho m que possui determinada característica x, $p = \frac{n_x}{m}$. No entanto, muitas vezes é imprático abordar toda a população, sendo preferível observar amostras aleatórias de tamanho n. Dessa forma, pode-se utilizar a proporção amostral p_a para estimar p.

Com o n grande, obtem-se uma aproximação de uma curva normal: $p_a \approx N(p, \frac{p1-p}{n})$. Assim, é possível obter um Intervalo de Confiança com coeficiente de confiança γ:

$$IC_p = \left[p_a - z_\gamma \cdot \sqrt{\frac{p_a(1-p_a)}{n}}, p_a + z_\gamma \cdot \sqrt{\frac{p_a(1-p_a)}{n}}\right]$$

Definindo-se a margem de erro como metade do tamanho do intervalo de confiança, tem-se que:

$$e = z_\gamma \cdot \sqrt{\frac{p_a(1-p_a)}{n}}$$

$$n = \left(\frac{z_\gamma}{e}\right)^2 p_a \cdot (1-p_a)$$

Com isso, consegue-se determinar um tamanho da amostra n para o qual é possível obter uma estimativa para p com erro e e confiança de γ.

Como $\pi = 4p$, sendo $p = \frac{n_{dentro}}{n_{total}}$, tem-se que $\pi_a = 4p$ e $\sigma_{\pi_m} = 4\sigma_{p_m}$. Assim:

$$IC_\pi = \left[\pi_a - 4z \cdot \sqrt{\frac{p_a \cdot (1-p_a)}{n}}, \pi_a + 4z \cdot \sqrt{\frac{p_a \cdot (1-p_a)}{n}}\right]$$

Portanto, analogamente ao processo anterior, tem-se que:

$$n_\pi = 16 \cdot \left(\frac{z}{e}\right)^2 \cdot p_a \cdot (1-p_a)$$

Similarmente, no pior dos casos:

$$n_\pi = 4 \cdot \left(\frac{z}{e}\right)^2$$

2 Cálculo do Valor de π com o Método de Monte Carlo

2.1 Algoritmo para a Estimativa

A Simulação de Monte Carlo consegue obter uma aproximação do valor de π. Mais precisamente, isso pode ser alcançando pelas seguintes etapas - exemplificado de maneira visual na Figura 1:

1. Determinar um quadrado no plano cartesiano por $x \in [-1, 1]$ e por $y \in [-1, 1]$;
2. Dispersar uniformemente um número de n de pontos dentro do quadrado;
3. Verificar a quantidade n_d de pontos cuja distância da origem é menor que 1, isto é, que estão dentro do círculo unitário;
4. A razão $\frac{n_d}{n}$ aproxima-se da razão da área da circunferência pela área do quadrado. Logo, $\frac{n_d}{n} \approx \frac{\pi}{4}$;
5. Com isso, estima-se $\pi \approx \frac{4 \cdot n_d}{n}$.

2.2 Intervalo de Confiança e Margem de Erro

Essa Simulação de Monte Carlo para o cálculo do valor de π é um problema de proporção populacional: de todos os pontos contidos dentro do quadrado, qual fração deles está dentro do círculo unitário. Sendo impossível computar todos os pontos nessa área, pode-se pegar uma amostra aleatória para estimar p por meio de p_a.

Dessa maneira, para definir o tamanho da amostra, precisa-se estabelecer o erro, o grau de confiança e uma proporção amostral. Assumindo $e = 0,05\%$ e $\gamma = 95\%$ (e, logo, $z_\gamma \approx 1,96$, falta definir p_a).

Conhecendo o valor de π, sabe-se que $p = \frac{\pi}{4}$. Não obstante, tomando esse valor como desconhecido, pode-se obter um valor de p_a satisfatório conduzindo diversas amostragens aleatórias e utilizando a média dos valores de π encontrados.

A Figura 2 demonstra a distribuição do valor de π encontrado em 10^6 amostragens aleatórias com $n = 10^3$. A média desses valores foi de $\mu \approx 3.1416$ com erro padrão $\sigma_m \approx 0,00013$.

Com isso, utilizando $p_a = \frac{\mu}{4}$, obtém-se que

$$n_\pi = 41.458.544$$

Outra opção quando se desconhece uma estimativa para p é simplesmente supor o caso em que n seja o maior possível, isto é, $p = 0,5$, resultando em:

$$n_\pi = 61.465.600$$

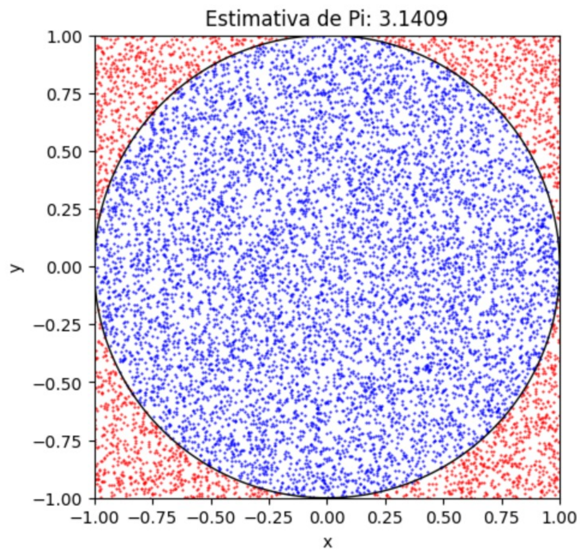

Figura 1: Exemplificação visual do Método de Monte Carlo para o cálculo de π.

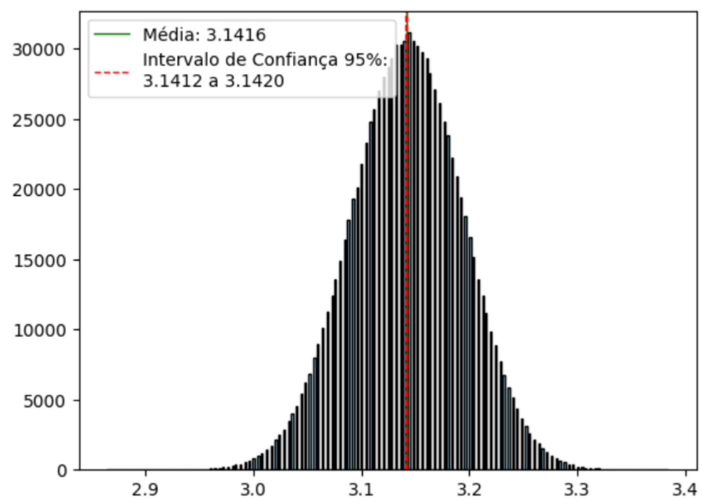

Figura 2: Distribuição das estimativas para π nas amostragens.

2.3 Análise de Convergência

Por fim, uma última maneira de compreender um valor propício para o tamanho da amostra n é observar graficamente a convergência das estimativas para o valor real. Como pode ser visto nas imagens a seguir, as estimativas de π apresentam um erro consistentemente abaixo de $0,05\%$ a partir de $3 \cdot 10^6$. No entanto, isso foi feito com poucas repetições, tornando o resultado obtido por intervalo de confiança e margem de erro uma melhor recomendação para n.

.

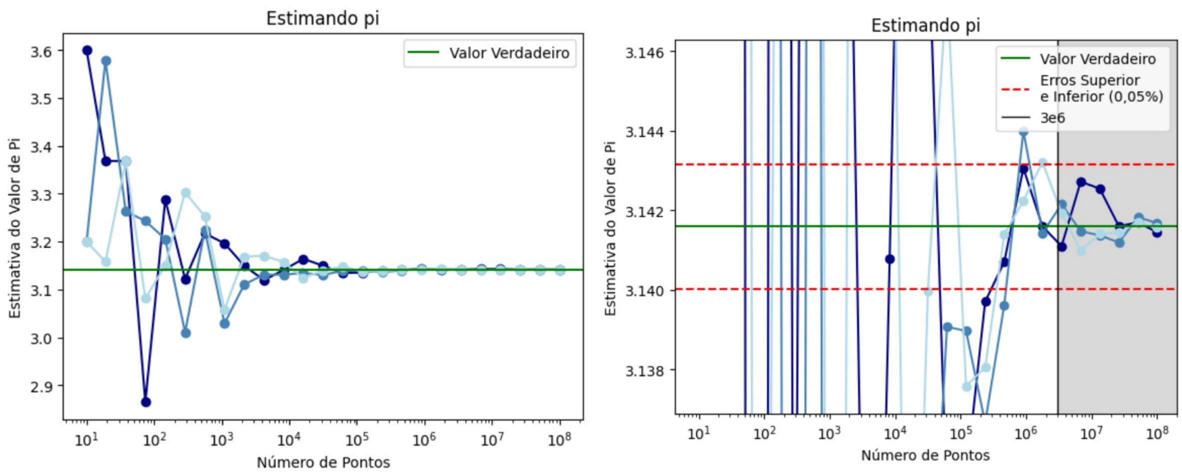

Figura 3: Simulações de Convergência com o aumento de n.

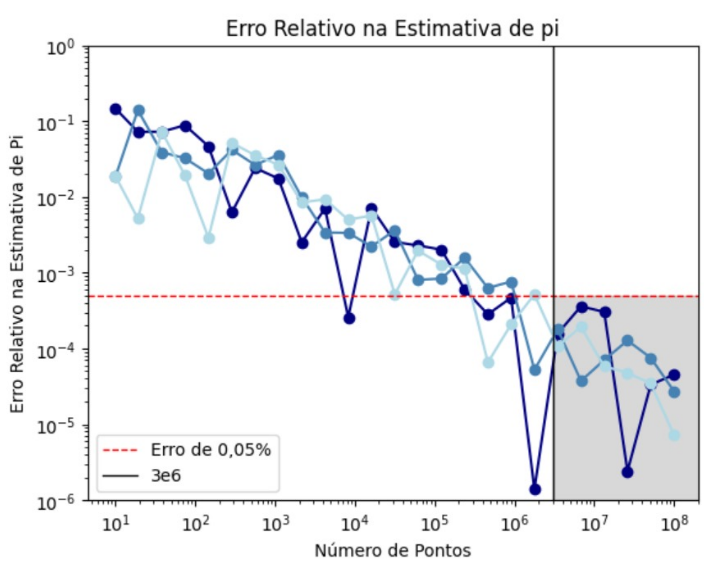

Figura 4: Análise do erro relativo das simulações conduzidas e sua evolução.