

Estimação de π através do algoritmo de Monte Carlo

Rennisson Davi D. Alves - 13687175

Abril de 2023

1 Introdução

O objetivo deste trabalho é tentar estimar a área do círculo unitário através do método de Monte Carlo. Para tratar deste problema e chegar no valor de π (número para o qual a área do círculo unitário converge) através do método computacional de Monte Carlo vamos primeiro considerar que o sorteio de pontos dentro do intervalo $(-1, 1)$ para todo x e y segue uma distribuição binomial. E após isso, vamos aproximar a binomial através da distribuição normal, a fim de facilitar alguns cálculos.

Como ferramenta, foi utilizada a linguagem de programação python com sua biblioteca numpy para desenvolvimento do algoritmo de Monte Carlo e realização dos experimentos. Para o seed do random, foi utilizado o NUSP 13687175.

2 Aproximação da binomial pela distribuição normal

Como dissemos anteriormente, o sorteio de pontos segue uma distribuição binomial. A equação da circunferência é $x^2 + y^2 \leq 1$ e a utilizaremos para analisar se determinado ponto está dentro ou não da circunferência unitária.

Não é difícil perceber que para nos aproximarmos o máximo possível do resultado esperado precisaremos de uma imensa quantidade de pontos, nesse caso, queremos uma precisão de 99.5%. E calcular probabilidade para tamanha quantidade de eventos não é tarefa fácil quando usamos a distribuição binomial.

Para simplificar o problema, podemos aproximar a binomial através da distribuição normal. O maior trabalho aqui é a manipulação algébrica para realizar essa transformação, porque depois de feita só precisamos substituir valores e encontrar valor-z correspondente na tabela normal.

De início, nossa variável X (pontos dentro da circunferência unitária) segue uma distribuição binomial com parâmetros n , números de eventos, e p , probabilidade do ponto escolhido estar na região $x^2 + y^2 \leq 1$. Tal distribuição é denotada por $X \sim \text{bin}(n, p)$.

Em geral, a normal tem como parâmetros a esperança (ou média) $E(X)$ e sua variância $\text{Var}(X)$, mais comumente representados por μ e σ^2 , respectivamente. A média da binomial é $E(X) = np$ e sua variância é $\text{Var}(X) = np(1-p)$. Para que ocorra a aproximação entre as duas distribuições basta que passemos a esperança e a variância da binomial como os respectivos parâmetros da normal, ficando então $\mu = np$ e $\sigma = np(1-p)$. Para não haver confusão, agora a nossa variável normalizada é $Y \sim N(\mu, \sigma)$.

3 Estimação dos parâmetros

Não temos ainda os valores exatos de μ e σ , então precisamos estimá-los.

Para o estimador \hat{p} , podemos observar que a circunferência está contida no quadrado de lado $2r$ (como $r=1$, o quadrado tem lado 2). Desse modo, deduzimos que a probabilidade de um ponto estar dentro da circunferência é $\hat{p} = \frac{\text{Área do círculo}}{\text{Área do quadrado}} = \frac{\pi r^2}{(2r)^2} = \frac{\pi}{4}$.

Com esse estimador já podemos trabalhar um pouco mais com a nossa distribuição para descobrirmos o erro ε que podemos cometer e conseguirmos dimensionar o número n mínimo de pontos necessários para uma boa aproximação.

$$P(|\hat{p} - p| < \varepsilon) \geq \gamma$$

$$P\left(\frac{-\varepsilon}{\frac{\sigma}{\sqrt{n}}} < \frac{\hat{p} - p}{\frac{\sigma}{\sqrt{n}}} < \frac{\varepsilon}{\frac{\sigma}{\sqrt{n}}}\right) = P\left(\frac{-\sqrt{n}\varepsilon}{\sigma} < Z < \frac{\sqrt{n}\varepsilon}{\sigma}\right) = P\left(-z_\gamma < Z < z_\gamma\right) \geq \gamma$$

Note que $z_\gamma = \frac{\sqrt{n}\varepsilon}{\sigma}$. Como foi exigida uma precisão $\gamma = 99.95\%$ para o nosso intervalo de confiança, temos que $z_\gamma = 3.27$, pela tabela normal padronizada. Dessa mesma relação, vamos estimar n :

$$z_\gamma = \frac{-\sqrt{n}\varepsilon}{\sigma} \implies n = \frac{z_\gamma^2 \sigma^2}{\varepsilon^2}$$

Já temos z_γ e σ . Vamos procurar ε agora:

$$\frac{|\hat{\pi} - \pi|}{\pi} \leq 0.0005$$

Lembre que $\hat{\pi} = 4 \cdot T(x_i) = 4 \cdot 1(x^2 + y^2 \leq 1)$. Então

$$\frac{|4 \cdot T(x_i) - \pi|}{\pi} = 4 \cdot \frac{|T(x_i) - \pi|}{\pi} \leq 0.0005$$

Realizamos alguns teste de antemão e vamos utilizar a média dos π encontrados nos experimentos realizados, apenas para termos um ponto de partida.

n	Area (π^{\wedge})	$\pi^{\wedge} - \pi$	Precisão (%)
1000	3.02	0.1215926535897931	96.1296
2000	3.112	0.029592653589793017	99.058
3000	3.092	0.049592653589793034	98.4214
4000	3.132	0.009592653589792999	99.6947
5000	3.0624	0.07919265358979333	97.4792
6000	3.11333	0.02825932025645983	99.1005
7000	3.09029	0.051306939304078814	98.3668
8000	3.089	0.05259265358979315	98.3259
9000	3.116	0.025592653589793013	99.1854
10000	3.1048	0.03679265358979311	98.8289

Primeiros experimentos

$$\frac{|3.093182 - \pi|}{\pi} \leq 0.000125$$

$$|3.093182 - \pi| \leq 0.000125\pi \implies \varepsilon = 0.000125\hat{\pi} = 0.00038665$$

Com o ε definido, basta substituir os valores e encontramos o n procurado:

$$n = \frac{z_{\gamma}^2 \sigma^2}{\varepsilon^2} = \frac{z_{\gamma}^2 \cdot (\hat{p}(1 - \hat{p}))}{\varepsilon^2} = \frac{3.27^2 (\frac{\pi}{4}(1 - \frac{\pi}{4}))}{0.00038665^2} = 12055432.3989$$

Logo, se utilizarmos um $n \geq 12055433$, vamos chegar a um valor próximo o suficiente do π real.

4 O algoritmo de Monte Carlo, experimentos e resultados

O algoritmo de Monte Carlo é uma técnica computacional desenvolvida para testes e aproximação de resultados de certos problemas. É uma ferramenta de inferência estatística utilizada geralmente em problemas muito complexos que só conseguem ser verificados através da ajuda dos computadores.

É justamente isso que precisamos neste caso. O problema em si não parece complexo, mas a verificação de seus resultados pode ser bastante trabalhosa. Fica mais clara ainda a sua necessidade se prestarmos atenção na última seção: para atingir um resultado satisfatório precisamos de um n maior que 7.000.000. Inviável sortear manualmente tal numero de pontos e verificar se cada um deles está na circunferência unitária. Por isso deixamos esse trabalho nas mãos do algoritmo de Monte Carlo. Deixamos que ele sorteie todos os pontos e verifique se cada ponto está ou não na circunferência unitária. Ao fim de cada experimento ele retorna a frequência de pontos dentro da circunferência, que se aproxima da área do próprio círculo a medida que vamos aumentando a nossa amostra de pontos.

Feita a análise dos parâmetros, vamos verificar os resultados obtidos através do algoritmo de Monte Carlo.

n	Area (π^{\wedge})	$\pi^{\wedge} - \pi$	Precisão (%)
12055433	3.14176	0.00016525755284257926	99.9947
12055433	3.1416	3.0070554943861794e-06	99.9999
12055433	3.1415	9.387872206323422e-05	99.997
12055433	3.14182	0.00023161767445767367	99.9926
12055433	3.1419	0.00030262300458572255	99.9904
12055433	3.14156	3.481821382589345e-05	99.9989
12055433	3.14165	5.3108947313695865e-05	99.9983
12055433	3.14143	0.00016521585279916096	99.9947
12055433	3.14181	0.00021801384952668812	99.9931
12055433	3.14238	0.0007913653002789012	99.9748
12055433	3.14192	0.0003278398507990943	99.9896
12055433	3.14149	0.00010482814212986469	99.9967
12055433	3.1414	0.00019441430630973144	99.9938
12055433	3.14168	8.463000508074359e-05	99.9973
12055433	3.14161	1.627907981749388e-05	99.9995
12055433	3.14163	3.253730961283452e-05	99.999
12055433	3.14157	1.7896382814353018e-05	99.9994
12055433	3.1418	0.00020341462277118083	99.9935
12055433	3.14207	0.0004764865232163018	99.9848
12055433	3.14179	0.00019545140817767148	99.9938

Experimentos com $n=12055433$

n	Area (π^{\wedge})	$\pi^{\wedge} - \pi$	Precisão (%)
15000000	3.14137	0.00022518692312623045	99.9928
15000000	3.14058	0.0010126535897931888	99.9678
15000000	3.14201	0.00041641307687356743	99.9867
15000000	3.14226	0.0006676130768736854	99.9787
15000000	3.14154	5.052025645957414e-05	99.9984
15000000	3.1417	0.00010627974354004976	99.9966
15000000	3.14191	0.0003145464102067841	99.99
15000000	3.14145	0.0001417202564599762	99.9955
15000000	3.14154	5.2386923126590546e-05	99.9983
15000000	3.14258	0.000990546410206683	99.9685
15000000	3.14224	0.0006454797435400117	99.9795
15000000	3.14227	0.0006796130768735864	99.9784
15000000	3.14153	6.518692312651453e-05	99.9979
15000000	3.1416	1.107974354042085e-05	99.9996
15000000	3.14146	0.00012918692312657853	99.9959
15000000	3.14116	0.0004374535897930798	99.9861
15000000	3.14144	0.00015078692312631148	99.9952
15000000	3.14139	0.0002041202564599942	99.9935
15000000	3.14198	0.00038814641020668006	99.9876
15000000	3.1414	0.00019425358979319185	99.9938

Experimentos com $n=15000000$

Observe que os resultados dos experimentos com n maior ou igual ao n estimado pelos processos algébricos das seções anteriores são excelentes. À exceção de um ou outro teste, quase todos os experimentos atingiram a precisão desejada. Tal desempenho dos experimentos realizados reforçam o caráter preciso e útil do método de Monte Carlo.