Estimando pi através do método de Monte Carlo

Guilherme Henrique Sampaio Marcelino

April 2022

1 Introduction

2 Desenvolvimento

2.1 Algoritmo

(1º)O Algoritmo inicia o total como 0, (2º) sorteia uniformemente x e y de um ponto $z = (x, y)$, (3º) utiliza da fórmula $T(z) = Ind(||z||_2 1)$, somando $T(z)$ ao total, e depois de refazer n vezes (2º) e (3º) o algoritmo irá estimar pi através do estimador $\hat{\pi} = 4\frac{\sum_{i=1}^{n} T_i}{n}$, onde $T_i = T(z_i)$ e $Total = \sum_{i=1}^{n} T_i$ e por fim retorna o valor estimado de pi.

2.2 Viés do estimador

Quando falamos de estimadores temos sempre que nos preocupar com se o estimador é ou não enviesado, isto é, se a média amostral do estimador é igual ao valor que queremos estimar(retirado do livro [CB]).
Definindo duas distribuições aleatórias uniformes $X \sim Unif(-1, 1)$ e $Y \sim Unif(-1, 1)$ temos que:
$f_T(x, y) = f_X(x)f_Y(y)I_{(0,1)}(x^2 + y^2)$
Isso vale pois, $T(x, y) = 1$, se $-1 < x < 1$, $-1 < y < 1$ e $x^2 + y^2 < 1$, como $-1 < x < 1$ está definido em X e $-1 < y < 1$ está definido em Y, pela definição de X e Y temos que $f_X(x) = \frac{1}{2}I_{(0,1)}(x)$ e $f_Y(y) = \frac{1}{2}I_{(0,1)}(y)$, substituindo temos:
$f_T(x, y) = \frac{1}{2}I_{(0,1)}(x)\frac{1}{2}I_{(0,1)}(y)I_{(0,1)}(x^2 + y^2)$
$f_T(x, y) = \frac{1}{4}I_{(0,1)}(x)I_{(0,1)}(y)I_{(-\sqrt[2]{1-x^2}, \sqrt[2]{1-x^2})}(y)$
$f_T(x, y) = \frac{1}{4}I_{(0,1)}(x)I_{(-\sqrt[2]{1-x^2}, \sqrt[2]{1-x^2})}(y)$
Como $f_T(x, y)$ é contínua, para calcular a probabilidade de T(x,y)=1 temos que calcular $f_T(x, y)$ em $-1 < x < 1$ e $-\sqrt[2]{1-x^2} < y < \sqrt[2]{1-x^2}$ através de integral dupla, logo:
$P(T(x, y) = 1) = \int_{-1}^{1}\int_{-\sqrt[2]{1-x^2}}^{\sqrt[2]{1-x^2}} \frac{1}{4}dydx$
$P(T(x, y) = 1) = \int_{-1}^{1} \sqrt[2]{1-x^2}\frac{1}{2}dx = \frac{\pi}{4}$
Como só temos $T(x, y) = 1$ ou $T(x, y) = 0$ temos que $T(x, y) \sim bernoulli(p)$,

e como $P(T(x,y)=1)=\frac{\pi}{4}$ temos que $T(x,y)$ $bernoulli(\frac{\pi}{4})$ logo a média de $T(x,y)$ é $(\frac{\pi}{4})$ (retirado de [CD]), agora podemos passar para o estimador utilizado para estimar π, $\hat{\pi}=4\frac{\sum_{i=1}^{n}T_i}{n}$, onde $T_i=T(x_i,y_i)$ agora calculando a média amostral temos:
$E(\hat{\pi})=E(4\frac{\sum_{i=1}^{n}T_i}{n})$
Como estamos gerando a amostra de forma mais aleatória possível podemos considerar que a amostra seja i.i.d(independentes e identicamente distribuídas), logo:
$E(\hat{\pi})=4\frac{\sum_{i=1}^{n}E(T_i)}{n}$
$E(\hat{\pi})=4E(T_i)=4\frac{\pi}{4}=\pi$
Logo o estimador é não-viesado.

2.3 Calculando n

Precisamos que o erro relativo de nosso estimador seja de no máximo 0.05%, isto é(retirado de [BF]):
$\frac{|\hat{\pi}-\pi|}{\pi}<0.05\%\rightarrow(1)|\hat{\pi}-\pi|<0.0005\pi$
Como sabemos que T(z) é uma bernoulli também podemos calcular sua média e variância facilmente, utilizando o TLC(Teorema do Limite Central, pego de [CD]) temos:
$P(-z_\gamma<\frac{\sum_{i=1}^{n}T_i-n\mu}{\sigma\sqrt{n}}<z_\gamma)=\gamma$
Como T $Bernoulli(\frac{\pi}{4})$ temos que $\mu=\frac{\pi}{4}$ e $\sigma=\sqrt{\frac{4\pi-\pi^2}{16}}$ (retirado de [CD]) logo:
$P(-z_\gamma<\frac{\sum_{i=1}^{n}T_i-n\frac{\pi}{4}}{\sqrt{\frac{4\pi-\pi^2}{16}}\sqrt{n}}<z_\gamma)=\gamma$
$P(-\sqrt{\frac{4\pi-\pi^2}{16}}z_\gamma<\sqrt{n}\frac{\sum_{i=1}^{n}T_i}{n}-\sqrt{n}\frac{\pi}{4}<\sqrt{\frac{4\pi-\pi^2}{16}}z_\gamma)=\gamma$
$P(-4\frac{1}{\sqrt{n}}\sqrt{\frac{4\pi-\pi^2}{16}}z_\gamma<\hat{\pi}-\pi<-4\frac{1}{\sqrt{n}}\sqrt{\frac{4\pi-\pi^2}{16}}z_\gamma)=\gamma$
$(2)P(|\hat{\pi}-\pi|<\frac{1}{\sqrt{n}}\sqrt{4\pi-\pi^2}z_\gamma)=\gamma$
Definindo $\gamma=0.9995$ temos $z_\gamma=3.5$ garantimos, com um coeficiente de coeficiança de 99.95%, que o nosso estimador não errará pi pela quantidade do erro relativo que queremos, o que falta agora é definir n para garantir o erro relativo de 0.05%:
$(2)|\hat{\pi}-\pi|<3,5\frac{1}{\sqrt{n}}\sqrt{4\pi-\pi^2}$
$(1)|\hat{\pi}-\pi|<0.0005\pi$
Como queremos que o erro relativo seja no máximo 0.0005, não temos problema se (2) for menor ou igual a (1), logo:
$3,5\frac{1}{\sqrt{n}}\sqrt{4\pi-\pi^2}\leq 0.0005\pi$
$(3,5)^2(\frac{4}{\pi}-1)\leq(0.0005)^2n$
Como é trivial que $\pi>2$ temos:
$(3,5)^2(\frac{4}{\pi}-1)<(3,5)^2(\frac{4}{2}-1)=(0.0005)^2n$
$\frac{(3,5)^2}{(0.0005)^2}\leq n\rightarrow n\geq 49000000$
Logo tomando $n=49000000$ conseguimos garantir, com um coeficiente de con-

fiança de 99.95%, que nosso estimador não errará por mais que 0.05%.

3 Resultado e Discussão

Temos que para $n = 49000000$ não é trivial nem para um computador rodar o algoritmo, já que o algoritmo tem complexidade $O(n)$ e n depende da função normal que, isso para apenas 0.05% de erro relativo, para erros menores esse algortimo não irá rodar para computadores que não são de ponta, isto é, irá ultrapassar os 100000000 operações por segundo comumente usados, mostrando que não é o melhor estimador para pi quando olhamos a complexidade do algoritmo, outros como utilizar o método de Newton(explicado no livro) em alguma função com pi como raíz(exemplo $sen(x)$) tem uma complexidade menor e é mais rápido.
Na estatística muitos dos problemas nascem no mundo real e se resolvem através de ferramentas estatística, uma dessas ferramentas é a inferência estatística, apartir da amostra tentamos estimar informações da distribuição, porém não precisamos usar estimadores apenas para isso, e esse algoritmo é um ótimo exemplo de que ao estimar parte da distribuições podemos estimar qualquer coisa que tenha ligação com a distribuição, como nesse caso em que T tinha relação com π.
Outro ponto do metódo de monte carlo é que para estimar regiões, constantes, e outros valores que são impossíveis por um motivo qualquer, ele é muito bom, pois mesmo sendo computacionalmente complexo são muito raros os problemas em que ele não funcione.

4 Conclusão

Mesmo conseguindo estimar pi como queríamos, ainda não é o melhor estimador quando se trata de precisão e complexidade computacional, mesmo para aplicações que não precisem de tanta precisão é melhor escrever pi que já foi estimado por outro método do que utilizar o método de monte carlo hoje em dia, mas a ideia por traz do método é um ótimo exemplo de utilização dos estimadores e para outros problemas o método de Monte Carlo funciona .

5 Bibliografia

[CB]G. Casella, R. L. Berger, Statistical Inference, 2nd ed., Pacific Grove: Duxbury/Thomson Learning, 2002.
[CD]C. A. B. Dantas, Probabilidade: um Curso Introdutório, São Paulo: Edusp, 1997.
[BF]Richard L. Burden, Douglas J. Faires, Annette M. Burden, Numerical Analysis, 10th ed,Cengage Learning,2015.