Aula 14: Integração de Monte Carlo

Prof. Dr. Eder Angelo Milani e Érika Soares Machado

21/06/2023

Integração de Monte Carlo

A integração de Monte Carlo é um método estatístico baseado em amostragem aleatória. Os métodos de Monte Carlo foram desenvolvidos no final dos anos 1940 após a Segunda Guerra Mundial.

Seja $g(x)$ uma função e suponha que queremos calcular

$$\int_a^b g(x)dx,$$

assumindo que a integral exista.

Observação Se X é uma v.a. com densidade $f(x)$, então a esperança da v.a $y = g(x)$ é

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

Além disso, um estimador não viesado para $E(g(X))$ é a média amostral.

Considere o problema de estimar

$$\theta = \int_0^1 g(x)dx$$

Se $x_1, ..., x_m$ é uma amostra aleatória da distribuição Uniforme(0,1), então

$$\hat{\theta} = \overline{g_m(x)} = \frac{1}{m}\sum_{i=1}^{m} g(x_i)$$

converge para $E(g(X)) = \theta$ com probabilidade 1, pela lei dos grandes números. Portanto, o simples estimador de $\int_0^1 g(x)dx$ é $\overline{g_m(x)}$.

Exemplo 1

Calcule

$$\theta = \int_0^1 e^{-x}dx$$

a partir de um estimador de Monte Carlo e compare com o verdadeiro valor.

Solução:

Método exato

$$\theta = \int_0^1 e^{-x} dx = -e^{-x}|_0^1 = -e^{-1} - (-e^0) = 1 - e^{-1} = 0.632120$$

Método de Monte Carlo Podemos escrever

$$\theta = \int_0^1 e^{-x} f(x) dx,$$

com $f(x)$ sendo a função densidade da distribuição Uniforme$(0, 1)$.

Disto, se $x_1, \ldots, x_m$ for uma amostra aleatória da distribuição Uniforme$(0, 1)$, então

$$\theta \approx \frac{1}{m} \sum_{i=1}^{m} g(x_i) = \frac{1}{m} \sum_{i=1}^{m} e^{-x_i}.$$

A partir destes resultados, a rotina a seguir calcula o valor de θ para diferentes tamanhos amostrais.

```
set.seed(2021)

theta_MC = function(m){
  a = runif(m)
  return(mean(exp(-a)))
}

theta_MC(100)      # aproximação com 100 valores
```

```
## [1] 0.6035229
```

```
theta_MC(1000)     # aproximação com 1000 valores
```

```
## [1] 0.6419337
```

```
theta_MC(10000)    # aproximação com 10000 valores
```

```
## [1] 0.628936
```

```
theta_MC(100000)   # aproximação com 100000 valores
```

```
## [1] 0.632366
```

```
theta_MC(1000000)   # aproximação com 1000000 valores
```

```
## [1] 0.6319968
```

Exemplo 2

Calcule

$$\theta = \int_0^1 x^2 dx$$

Solução

Exata

$$\theta = \int_0^1 x^2 dx = \frac{x^3}{3}|_0^1 = \frac{1^3}{3} - 0 = 1/3 = 0.3333$$

Monte Carlo Considerando que $x_1, \ldots, x_m$ é uma amostra aleatória da distribuição Uniforme$(0,1)$, então

$$\theta \approx \frac{1}{m}\sum_{i=1}^{m} g(x_i) = \frac{1}{m}\sum_{i=1}^{m} x_i^2.$$

```
set.seed(2021)

theta_MC = function(m){
  a = runif(m)
  return(mean(a^2))
}

theta_MC(100)      # aproximação com 100 valores
```

```
## [1] 0.3969113
```

```
theta_MC(1000)     # aproximação com 1000 valores
```

```
## [1] 0.321886
```

```
theta_MC(10000)    # aproximação com 10000 valores
```

```
## [1] 0.3387756
```

```
theta_MC(100000)   # aproximação com 100000 valores
```

```
## [1] 0.3331281
```

```
theta_MC(1000000)   # aproximação com 1000000 valores
```

```
## [1] 0.3335139
```

Podemos substituir a densidade da Uniforme(0,1) por qualquer outra densidade com suporte no intervalo entre os limites de integração. Por exemplo,

$$\int_a^b g(t)dt = (b-a)\int_a^b g(t)\frac{1}{b-a}dt$$

sendo Y a v.a. com distribuição no intervalo (a,b). O valor da integral desejada é portanto $(b-a)$ vezes o valor médio de $g(.)$ sobre (a,b).

Exemplo 3

Calcule o estimador de Monte Carlo para

$$\theta = \int_2^4 e^{-x} dx$$

e compare com o verdadeiro valor.

Solução

Método exato

$$\theta = \int_2^4 e^{-x} dx = -e^{-x}\Big|_2^4 = e^{-2} - e^{-4} = 0,117019644$$

Método de Monte Carlo A partir do método de Monte Carlo temos

$$\theta = \int_2^4 e^{-x} dx = \int_2^4 e^{-x} \frac{(4-2)}{(4-2)} dx = (4-2) \int_2^4 e^{-x} \frac{1}{(4-2)} dx,$$

logo, conseguimos introduzir no integrando a função densidade da distribuição Uniforme$(2, 4)$. Consequentemente, a geração dos valores aleatórios deverão apresentar tal distribuição.

```
set.seed(2021)
a = 2
b = 4
theta_MC2 = function(m, a, b){
  d = runif(m, a, b)
  e = exp(-d)
  (b-a)*mean(e)
}

theta_MC2(100,a, b)
```

```
## [1] 0.1081467
```

```
theta_MC2(1000,a, b)
```

```
## [1] 0.1207949
```

```
theta_MC2(10000,a, b)
```

```
## [1] 0.1159025
```

```
theta_MC2(100000,a, b)
```

```
## [1] 0.1171084
```

Exemplo 4

Calcule

$$\theta = \int_0^3 x^2 dx$$

Solução

Exata

$$\theta = \int_0^3 x^2 dx = \frac{x^3}{3}\Big|_0^3 = \frac{3^3}{3} - 0 = 9$$

Monte Carlo A partir do método de Monte Carlo temos

$$\theta = \int_0^3 x^2 dx = \int_0^3 x^2 \frac{(3-0)}{(3-0)} dx = (3-0)\int_0^3 x^2 \frac{1}{(3-0)} dx$$

logo, conseguimos introduzir no integrando a função densidade da distribuição Uniforme$(0,3)$. Consequentemente, a geração dos valores aleatórios deverão apresentar tal distribuição.

```
set.seed(2021)
a = 0
b = 3
theta_MC2 = function(m, a, b){
  d = runif(m, a, b)
  e = d^2
  (b-a)*mean(e)
}

theta_MC2(100,a, b)
```

```
## [1] 10.71661
```

```
theta_MC2(1000,a, b)
```

```
## [1] 8.690923
```

```
theta_MC2(10000,a, b)
```

```
## [1] 9.146941
```

```
theta_MC2(100000,a, b)
```

```
## [1] 8.994459
```

```
theta_MC2(10000000,a, b)
```

```
## [1] 9.00269
```

Alternativamente, para calcular:

$$\int_a^b g(t)dt$$

podemos fazer uma mudança nos limites da integral, alterando para 0 a 1. A transformação linear

$$y = \frac{t-a}{b-a} \Rightarrow y(b-a) + a = t$$

e

$$dy = \frac{1}{(b-a)}dt$$

disto,

$$\begin{aligned}\int_a^b g(t)dt &= \int_0^1 g(y(b-a)+a)(b-a)dy \\ &= (b-a)\int_0^1 g((b-a)y+a)dy\end{aligned}$$

Exemplo 5

Calcule um estimador de Monte Carlo para

$$\theta = \int_2^4 e^{-x}dx$$

Solução

Do método exato, temos que

$$\theta = \int_2^4 e^{-x}dx = 0,117019644$$

Usando a técnica da transformação linear, temos

$$y = \frac{x-2}{4-2} = \frac{x}{2} - 1,$$

logo,

$$dy = \frac{1}{2}dx$$

e

$$y = \frac{x}{2} - 1 \Rightarrow 2y + 2 = x,$$

disto,

$$\theta = \int_0^1 e^{-(2y+2)}(4-2)dy = 2\int_0^1 e^{-2y-2}dy$$

```r
set.seed(2021)
a = 2
b = 4
theta_MC3=function(m,a,b){
  y = runif(m)
  d = exp(-((b-a)*y+a))
  return((b-a)*mean(d))
}

theta_MC3(100,a,b)
```

```
## [1] 0.1081467
```

```
theta_MC3(1000,a,b)
```

```
## [1] 0.1207949
```

```
theta_MC3(10000,a,b)
```

```
## [1] 0.1159025
```

```
theta_MC3(100000,a,b)
```

```
## [1] 0.1171084
```

Integração de Monte Carlo para intervalos ilimitados

Para introduzir a integração de Monte Carlo para intervalos ilimitados utilizamos um exemplo, veja a seguir.

Exemplo 1

Use a abordagem de Monte Carlo para estimar a função distribuição acumulada da normal padrão. Ou seja,

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\Big(-\frac{t^2}{2}\Big) dt.$$

Primeiro, note que não podemos aplicar o algoritmo diretamente devido os limites de integração ser um intervalo ilimitado. Contudo, podemos dividir este problema em dois casos: $x \geq 0$ e $x < 0$, e usar a simetria da densidade da normal para lidar com o caso $x < 0$. Então o problema é estimar

$$\theta = \int_0^x e^{-\frac{t^2}{2}} dt, x > 0.$$

Podemos resolver gerando números da variável Uniforme$(0, x)$, mas seria necessário mudar o parâmetro da distribuição uniforme para cada diferente valor da função distribuição acumulada. Suponha que preferimos um algoritmo que sempre amostra da Uniforme$(0, 1)$. Note que

$$\theta = \int_0^x e^{-\frac{t^2}{2}} dt,$$

pode ser escrito mudando os intervalos de integração para 0 a 1, logo

$$y = \frac{t-a}{b-a} = \frac{t}{x} \Rightarrow t = xy, \Rightarrow dt = xdy,$$

disto,

$$\theta = \int_0^1 \exp\Big(-\frac{(yx)^2}{2}\Big) xdy,$$

sendo que o x é dado. Portanto,

$$\theta = E_Y\left(xe^{\frac{-(yx)^2}{2}}\right),$$

sendo que Y é distribuído segundo a v.a. Uniforme$(0,1)$.

Gerando números aleatórios $y_1, ..., y_m$ da Uniforme$(0,1)$ e calculando

$$\hat{\theta} = \overline{g_m(\underline{y})} = \frac{1}{m}\sum_{i=1}^{m} xe^{-\frac{(y_i x)^2}{2}},$$

a média amostral $\hat{\theta}$ converge para $E(\hat{\theta}) = \theta$ quando $m \to \infty$

Note que

- Se $x > 0$, a estimativa de $\Phi(x)$ é $0,5 + \frac{\hat{\theta}}{\sqrt{2\pi}}$;
- Se $x < 0$, calcule $\Phi(x) = 1 - \Phi(-x)$;
- Se $x = 0$, $\Phi(x) = 0,5$.

```
set.seed(2021)

fda = function(x,m){
  x1 = x
  x = ifelse(x<0, -x, x)
  y = runif(m)
  g = x*exp(-(y*x)^2/2)
  Fy = mean(g)/sqrt(2*pi) + 0.5
  fy = ifelse(x1<0, 1 - Fy, Fy)
  fy
}

fda(0.6,1000)
```

```
## [1] 0.7258884
```

```
# comparando com o valor obtido por meio do comando do R
pnorm(0.6)
```

```
## [1] 0.7257469
```

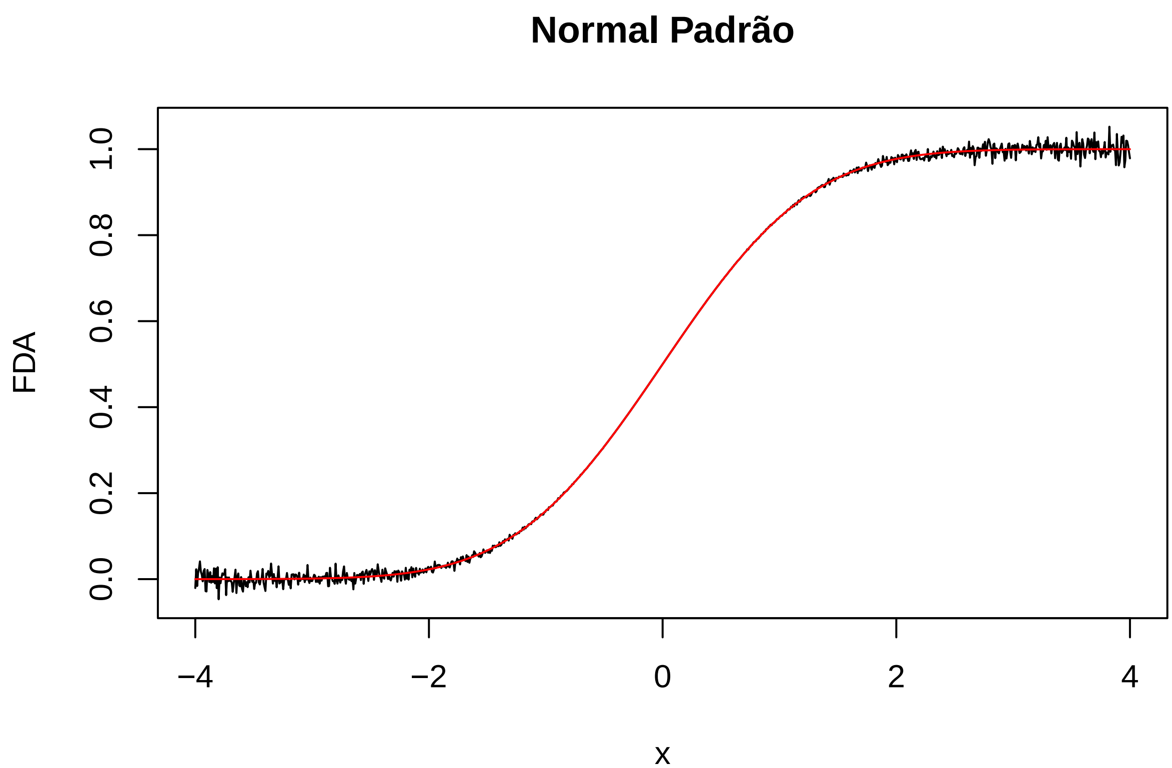
```
fda(-0.5,1000)
```

```
## [1] 0.3085872
```

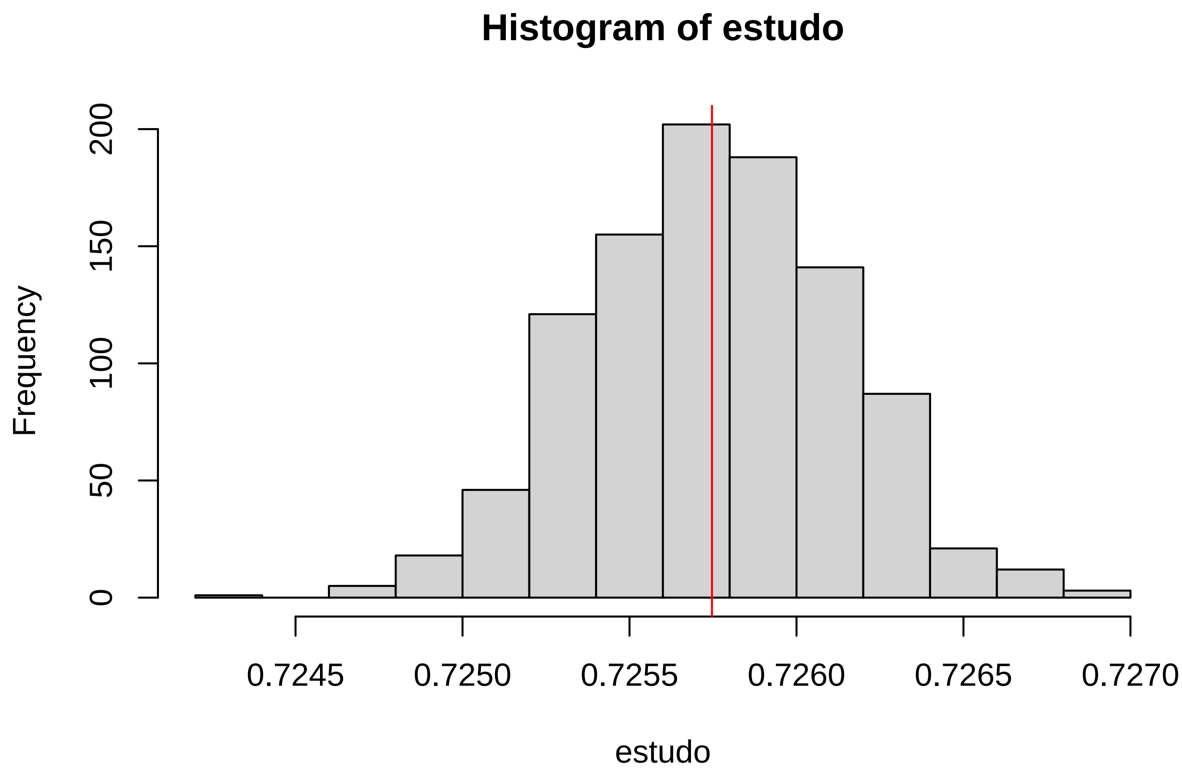
```
# comparando com o valor obtido por meio do comando do R
pnorm(-0.5)
```

```
## [1] 0.3085375
```

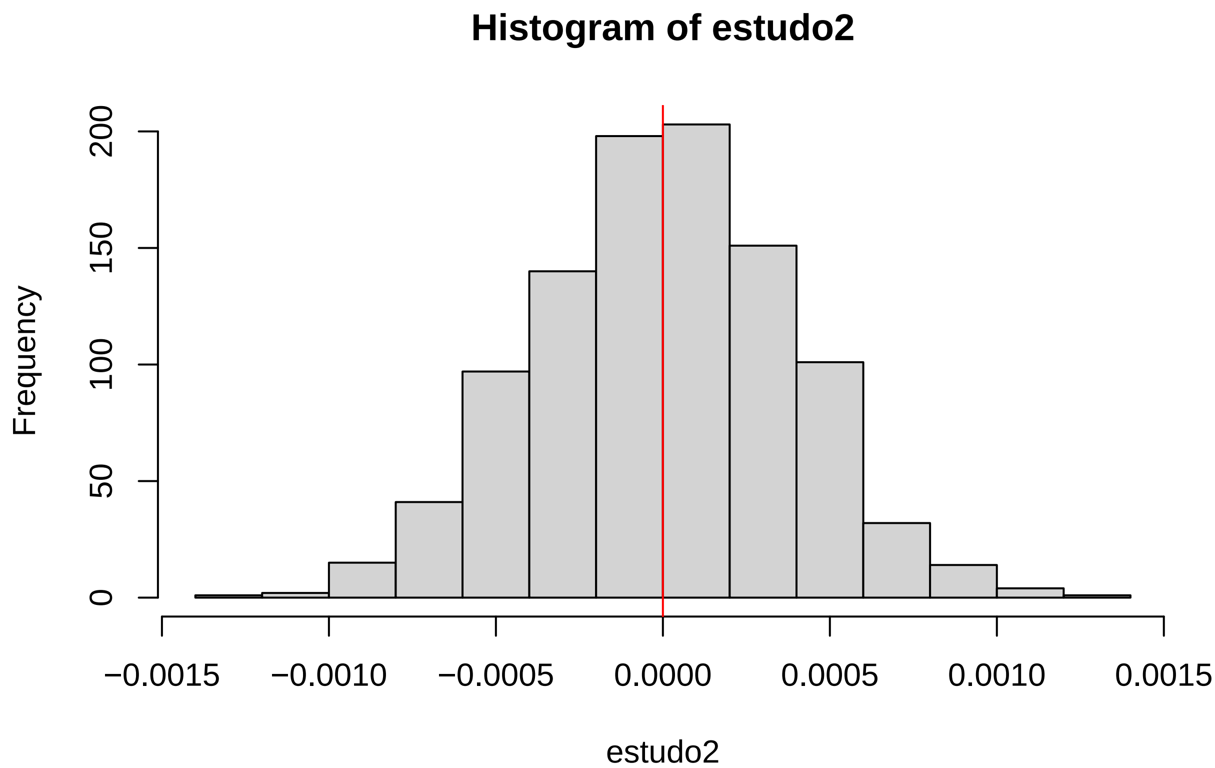
```
# construindo o gráfico da função distribuição acumulada

fda_np=matrix(NA,ncol=1000,nrow=3)

fda_np[1,]=seq(-4,4,length.out = 1000)
for( i in 1:1000){
fda_np[2,i]=fda(fda_np[1,i],1000)
fda_np[3,i]=pnorm(fda_np[1,i])
}

plot(fda_np[1,], fda_np[2,], type="l",ylab="FDA",xlab="x", main="Normal Padrão")
lines(fda_np[1,], fda_np[3,], col="red")
```




```
estudo=numeric()
estudo2=numeric()
for(i in 1:1000){
estudo[i]=fda(0.6,1000)
estudo2[i]=estudo[i]-pnorm(0.6)
}
hist(estudo)
abline(v=pnorm(0.6), col="red")
```




```
hist(estudo2)
abline(v=0, col="red")
```



Exercícios

1. Calcule o valor da integral $\int_0^1 \frac{1}{\sqrt{2\pi}} \exp(-0.5x^2)dx$
2. Calcule o valor da integral $\int_0^3 \frac{1}{\sqrt{2\pi}} \exp(-0.5x^2)dx$
3. Calcule o valor da integral $\int_{-3}^3 \frac{1}{\sqrt{2\pi}} \exp(-0.5x^2)dx$
4. Sabemos que a função densidade da distribuição t-Student é dada por

$$f(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})}\left(1 + \frac{x^2}{\nu}\right)^{-(\frac{\nu+1}{2})}$$

faça uma tabela com os valores da função distribuição acumulada, considerando $x \in (-4, 4)$, com saltos de $0,1$ e $\nu = 1$.