Exemplos_MC

Aishameriane Schmidt

10 de abril de 2017

Exemplos da aula do dia 10/04/17 - Integração por Monte Carlo

Exemplo 1: Integrando uma função determinística

Suponha que o objetivo é integrar $I = \int_1^2 \exp(\theta) d\theta$ sem usar integração analítica.

Reinterprete I com,o uma esperança em relação a $\theta \sim U(1,2)$ (escolhida de maneira conveniente dentro dos intervalos de integração). Sabemos que a densidade de uma $Uniforme[a,b]$ é $\frac{1}{b-a}$, de forma que $p_U(\theta) = 1/(2-1)$. Obtemos então:

$$I = \int\limits_1^2 \exp\{\theta\} d\theta = (2-1) \int\limits_1^2 \exp\{\theta\} \frac{1}{2-1} d\theta = (2-1)\mathbb{E}_{\mathcal{U}}[\exp\{\theta\}]$$

Para aproximar a integral, simule S observações de $\theta \sim U(1,2)$ e aproxime $\mathbb{E}[\exp(\theta)]$ através da média amostral:

$$\hat{I}_n = \frac{1}{S} \sum_{i=1}^{S} \exp(\theta_i)$$

Usando $S = 10.000$, obtemos 4.7080 (o valor está diferente das notas de aula possivelmente pela diferença na semente aleatória e no programa), que é uma aproximação razoável para o valor exato 4.6707743.

```
# Fixa uma semente aleatória
#(para poder reproduzir o exemplo depois obtendo os mesmos valores)
set.seed(1234)

# Faz 1000 retiradas de uma Uniforme(1,2) e armazena em um vetor
theta <- runif(1000,min = 1, max = 2)
head(theta)
```

```
## [1] 1.113703 1.622299 1.609275 1.623379 1.860915 1.640311
```

```
# Calcula a exponencial do vetor theta
exp_theta <- exp(theta)
head(exp_theta)
```

```
## [1] 3.045617 5.064723 4.999184 5.070196 6.429620 5.156771
```

```
# Faz a soma dos valores e divide por S
I_n <- (sum(exp_theta))/length(theta)
I_n
```

```
## [1] 4.708025
```

Exemplo 2: FDA da Normal

A F.D.A. Normal padrão, dada por

$$\Phi(\theta) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{\theta^2/2}\, d\theta$$

não possui fórmula fechada, então pode ser interessante considerar integração MC. Se amostrarmos $\theta^i \sim N(0,1)$, então

$$\Phi(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{\theta^2/2}\, d\theta \approx \widehat{\Phi}_S(t) = \frac{1}{S}\sum_{i=1}^{S} \mathbb{I}(\theta^i \leq t)$$

Note que $\widehat{\Phi}_S(t)$ é uma variável aleatória Bernoulli, logo sua variância é $\Phi(t)(1-\Phi(t))/S$. t próximo de 0 implica que a variância de $\widehat{\Phi}_S(t)$ é $1/4S$, logo precisamos de 200.000 observações, em média, para conseguirmos precisão de 4 dígitos.

```r
# Seta a semente
set.seed(1235)

# Cria um vetor de tamanho 200.000
theta<-rep(0,1)

# Fixa um t
t <- 0

# Cria um vetor para as indicadoras
indicadora <- rep(0,200000)

# Gera um valor aleatório da normal padrão e compara com o valor de t
for (i in 1:length(indicadora)){
  theta<-rnorm(1, mean = 0, sd = 1)
  ifelse(theta <= t, indicadora[i] <- 1, indicadora[i]<-0)
}

head(indicadora)
```

```
## [1] 1 1 0 0 0 0
```

```r
sum(indicadora)/length(indicadora)
```

```
## [1] 0.498735
```

```r
pnorm(t,mean=0,sd=1)
```

```
## [1] 0.5
```

Comparando o valor tabelado de 0.5 com o valor aproximado de 0.498735 obtemos um erro de 0.001265.

Para t menor que -4.5 precisaremos de muito mais observações ainda para conseguir uma estimativa acurada para esta probabilidade.

```r
# Seta a semente
set.seed(1234)

# Cria um vetor de tamanho 200.000
theta<-rep(0,1)

# Fixa um t
t <- -4.5
```

```r
# Cria um vetor para as indicadoras
indicadora <- rep(0,999000)

# Gera um valor aleatório da normal padrão e compara com o valor de t
for (i in 1:length(indicadora)){
  theta<-rnorm(1, mean = 0, sd = 1)
  ifelse(theta <= t, indicadora[i] <- 1, indicadora[i]<-0)
}

head(indicadora)
```

```
## [1] 0 0 0 0 0 0
```

```r
sum(indicadora)/length(indicadora)
```

```
## [1] 3.003003e-06
```

```r
pnorm(t,mean=0,sd=1)
```

```
## [1] 3.397673e-06
```

Mas então como calcular probabilidades de eventos raros utilizando métodos de Monte Carlo? Utilizando amostragem por eficiência.

Calcular probabilidade de eventos raros como $\Phi(-4.5)$ usando este método MC simples é difícil, pois muito raramente iremos amostrar θ^i tal que $\mathbb{I}(\theta^i \leq -4.5) = 1$, logo $\widehat{\Phi}_S(-4.5) = 0$ mesmo para um valor alto de S. Mas usando a regra de mudança de variáveis, podemos usar $v = \frac{1}{x}$:

$$\int_{-\infty}^{-4.5} \frac{1}{\sqrt{2\pi}} e^{\theta^2/2}\, d\theta = \int_{\frac{-1}{4.5}}^{0} \frac{\phi(1/v)}{v^2} dv = \frac{1}{4.5} \int_{\frac{-1}{4.5}}^{0} \frac{\phi(1/v)}{v^2} p_U(v) dv$$

Podemos amostrar $v_i \sim U(-1/4.5, 0)$, então:

$$\int_{-\infty}^{-4.5} \frac{1}{\sqrt{2\pi}} e^{\theta^2/2}\, d\theta \approx \widehat{\Phi}_S^U(-4.5) = \frac{1}{S} \sum_{i=1}^{S} \frac{\phi(1/v^i)}{4.5 v^{i^2}}$$

Note que a F.D.P. de v $p_U(v) = 4.5$ é usada no denominador para compensar o fato de que não amostramos da distribuição original, mas sim de uma distribuição alternativa.

```r
set.seed(1234)
vetor_v <- runif(n = 100000, min =(-1/4.5) , max = 0)
quociente<- pnorm(1/vetor_v, mean=0, sd=1)/(4.5*(vetor_v^2))
summary(quociente)
```

```
##      Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
## 0.000e+00 0.000e+00 0.000e+00 6.950e-07 7.908e-09 1.529e-05
```

```r
sum(quociente)
```

```
## [1] 0.06949705
```

```r
I_n <- (1/length(vetor_v))*sum(quociente)
I_n
```

```
## [1] 6.949705e-07
```

```r
pnorm(-4.5, mean=0, sd=1)
```

```
## [1] 3.397673e-06
```