Aula 4: Método da Aceitação-Rejeição

Prof. Dr. Eder Angelo Milani

03/05/2023

Método da Aceitação-Rejeitação

Suponha que X e Y são v.a.'s com densidade f e g, respectivamente, e exista uma constante c tal que

$$\frac{f(t)}{g(t)} \leq c,$$

para todo t, tal que $f(t) > 0$. Então, o método da aceitação-rejeição (ou método da rejeição) pode ser aplicado para gerar valores da v.a. X.

O método da aceitação-rejeição

1. Encontrar uma v.a. Y com densidade g que satisfaz $f(t)/g(t) \leq c$, para todo t, tal que $f(t) > 0$, e fornecer um método de geração da v.a. Y.
2. Para cada valor aleatório requerido faça:

(a) gerar uma valor aleatório y da distribuição com densidade g;

(b) gerar uma valor aleatório u da distribuição Uniforme$(0, 1)$;

(c) if $u < f(y)/(cg(y))$ aceitar y e fazer $x = y$, caso contrário rejeitar y e repetir o Passo 2.

Note que no Passo 2(c),

$$P(aceitar \mid Y) = P\Big(U < \frac{f(Y)}{cg(Y)} \mid Y\Big) = \frac{f(Y)}{cg(Y)}.$$

A última igualdade é simplesmente avaliar a função acumulada da distribuição Uniforme$(0, 1)$. A probabilidade de aceitação para qualquer iteração é portanto dada por

$$P(aceitar) = \sum_y P(aceitar \mid Y)P(Y = y) = \sum_y \frac{f(Y)}{cg(Y)} g(Y) = \frac{1}{c},$$

e o número de interações até a aceitação tem distribuição geométrica com média c. Portanto, em média, para cada valor amostral de X é necessário c iterações. Logo, para uma boa eficiência do método, Y deve ser de fácil simulação e c pequeno.

Para verificar que os valores aceitos tem a mesma distribuição que X, basta aplicar o Teorema de Bayes. No caso discreto, para cada k tal que $f(k) > 0$, temos

$$P(Y = k \mid aceito) = \frac{P(aceitar \cap k)}{P(aceitar)} = \frac{P(aceitar \mid k)g(k)}{P(aceitar)} = \frac{[f(k)/(cg(k))]g(k)}{1/c} = f(k)$$

Para o caso contínuo, a prova é similar.

Exemplo 1

Este exemplo ilustra o método da aceitação-rejeição para a distribuição beta. Na média, quantos números aleatórios devem ser simulados para gerar 1000 valores da distribuição Beta($\alpha = 2, \beta = 2$) utilizando este método?

Note que uma boa escolha para a distribuição g é fundamental para a eficiência do método. Observe que a densidade da distribuição Beta($\alpha = 2, \beta = 2$) é dada por

$$f(x) = 6x(1-x),\ 0 < x < 1,$$

adotando g como a distribuição Uniforme$(0,1)$, temos que

$$\frac{f(x)}{g(x)} = 6x(1-x),\ \text{para todo } 0 < x < 1,$$

portanto, podemos adotar $c = 6$, logo

$$\frac{f(x)}{g(x)} \leq 6,\ \text{para todo } 0 < x < 1.$$

Um valor aleatório y gerado de g é aceito se

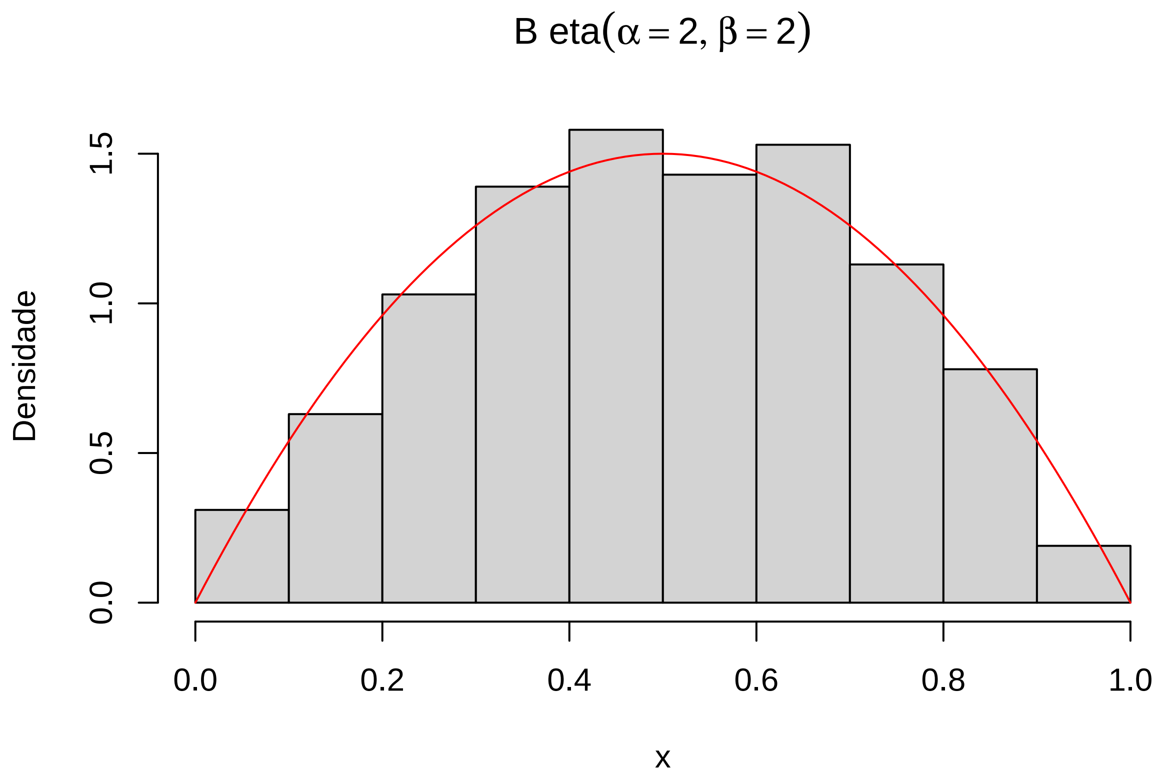
$$\frac{f(y)}{cg(y)} = \frac{6y(1-y)}{6(1)} = y(1-y) > u.$$

Na média, $c * n = 6 * 1000 = 6000$ iterações são necessárias para gerar uma amostra de tamanho 1000. A seguir é apresentado o código para a geração de 1000 valores da distribuição Beta($\alpha = 2, \beta = 2$).

```
set.seed(2023)
n <- 1000
x <- numeric(n) # amostra requerida
cont <- 0  # vai contar até atingir o tamanho amostral
j <- 0 # vai contar as iterações

while(cont<n){
  u <- runif(1)
  j <- j+1
  y <- runif(1) # gerando valores da densidade g
  if(y*(1-y)>u) { cont <- cont+1
                  x[cont] <- y
                  }
              }

cat("Quant. de iterações=",j, "\n")
```

```
## Quant. de iterações= 5947
```

```r
# histograma dos valores gerados

hist(x,prob=T, main=expression("B eta"(alpha==2, beta==2)), ylab="Densidade")
aux <- seq(0,1,0.01)
# curva da densidade beta(2,2)
lines(aux,dbeta(aux,2,2), col="red")
```




```r
# Para comparar a amostra obtida por meio dos decis - percentis

p <- seq(0.1,0.9,0.1)
Dhat <- quantile(x,p)
D <- qbeta(p,2,2)

round(rbind(Dhat, D),3)
```

```
##        10%   20%   30%   40%   50%   60%   70%   80%   90%
## Dhat 0.206 0.304 0.370 0.443 0.507 0.566 0.639 0.707 0.792
## D    0.196 0.287 0.363 0.433 0.500 0.567 0.637 0.713 0.804
```

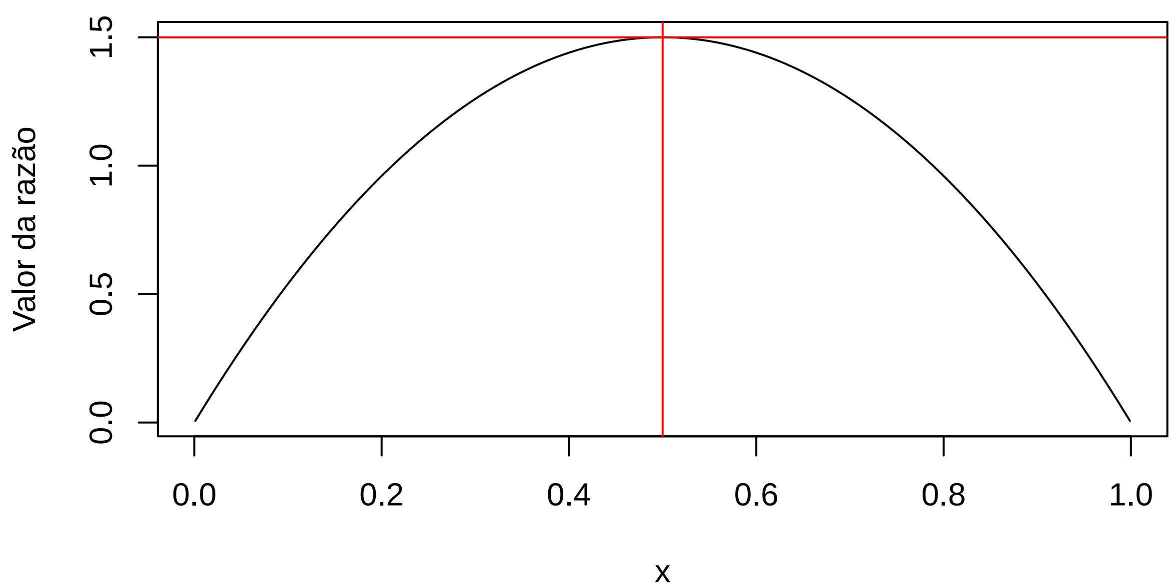
Exemplo 2

Refaça a análise do cálculo do calor c e obtenha uma nova geração para a distribuição Beta$(2, 2)$.

O gráfico a seguir calcula a razão

$$\frac{f(x)}{g(x)} = 6x(1-x), \text{ para todo } 0 < x < 1.$$

```
# função para calcular o valor da razão de f/g

razao <- function(x){
  6*x*(1-x)
}


x_aux <- seq(0.001,0.999,0.001)
razao_aux <- razao(x_aux)

plot(x_aux, razao_aux, type = 'l', ylab="Valor da razão", xlab="x")
abline(h=1.5,v=0.5, col='red')
```



Analisando o gráfico, vemos que a razão é menor ou igual a 1,5, para todo $0 < x < 1$. Portanto, um valor aleatório y gerado de g é aceito se

$$\frac{f(y)}{cg(y)} = \frac{6y(1-y)}{1,5(1)} = 4y(1-y) > u.$$

Na média, $c * n = 1,5 * 1000 = 1500$ iterações são necessárias para gerar uma amostra de tamanho 1000. A seguir é apresentado o código para a geração de 1000 valores da distribuição Beta($\alpha = 2, \beta = 2$).

Repetindo o código acima com as devidas alterações, temos

```r
set.seed(2023)
n <- 1000
x <- numeric(n) # amostra requerida
cont <- 0  # vai contar até atingir o tamanho amostral
j <- 0 # vai contar as iterações

while(cont<n){
  u <- runif(1)
  j <- j+1
  y <- runif(1) # gerando valores da densidade g
  if(4*y*(1-y)>u) { cont <- cont+1
                    x[cont] <- y
                    }
              }

cat("Quant. de iterações=",j, "\n")
```

```
## Quant. de iterações= 1503
```

```r
# Para comparar a amostra obtida por meio dos decis - percentis

p <- seq(0.1,0.9,0.1)
Dhat <- quantile(x,p)
D <- qbeta(p,2,2)

round(rbind(Dhat, D),3)
```

```
##        10%   20%   30%   40%   50%   60%   70%   80%   90%
## Dhat 0.197 0.280 0.351 0.420 0.481 0.546 0.626 0.696 0.781
## D    0.196 0.287 0.363 0.433 0.500 0.567 0.637 0.713 0.804
```

Nota: observe que a mudança do valor c exigiu um esforço computacional muito menor para a geração de uma amostra tão boa quanto. Portanto, a escolha do limitante influência muito na eficiência da geração.

Exercícios

1. Utilizando o método da aceitação-rejeição, obtenha uma amostra de tamanho 1.000 da distribuição Beta($\alpha = 3, \beta = 3$). Na média, quantos números aleatórios devem ser simulados para gerar 1.000 valores da distribuição, compare com o valor encontrado. Faça uma comparação dos decis amostrados com os teóricos.

2. Considere a distribuição triangular dada por

$$
\begin{aligned}
f(x) &= 4x, \text{se } 0 < x < 1/2 \\
&= 4 - 4x, \text{se } 1/2 \leq x < 1 \\
&= 0, \text{caso contrário}
\end{aligned}
$$

Utilizando o método da transformação inversa e apenas valores da distribuição uniforme, gerar 2.000 valores da distribuição triangular definida acima. Obter o histograma com a curva da distribuição.