Aula 2 - Método da transformação inversa - caso discreto

Prof. Dr. Eder Angelo Milani

24/04/2023

Método da Transformação Inversa - Caso Discreto

O método da transformação inversa também pode ser aplicado para distribuições discreta. Se X é uma v.a. discreta e os pontos de descontinuidade da função distribuição acumulada são

$$x_1 < \ldots < x_{i-1} < x_i < x_{i+1} < \ldots < x_k,$$

então a transformação inversa é

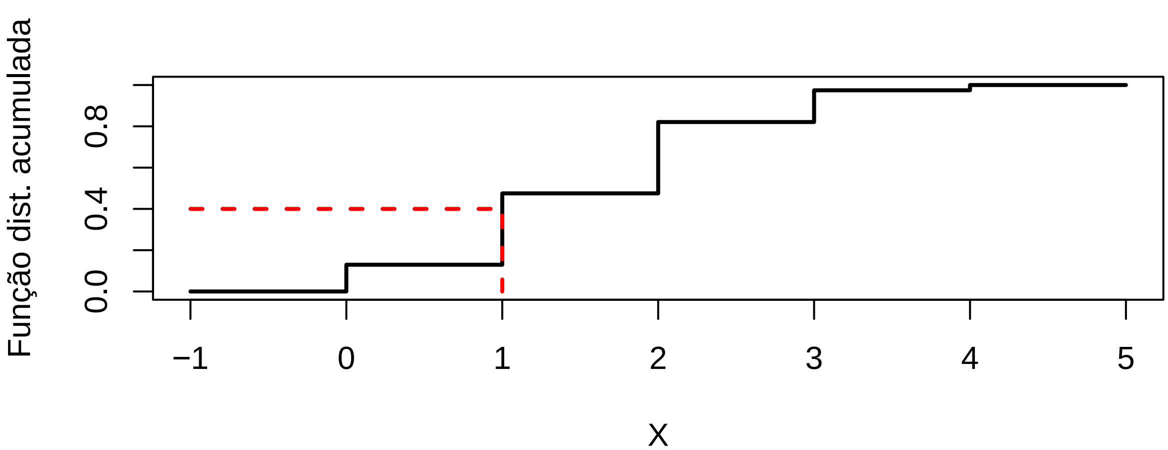
$$F_X^{-1}(u) = x_i, \text{ com } F_X(x_{i-1}) < u \leq F_X(x_i).$$

Para cada valor aleatória requerido é necessário seguir os passos

(a) gere uma valor aleatório u da distribuição Uniforme$(0, 1)$,

(b) calcule x_i satisfazendo a expressão $F_X(x_{i-1}) < u \leq F(x_i)$.

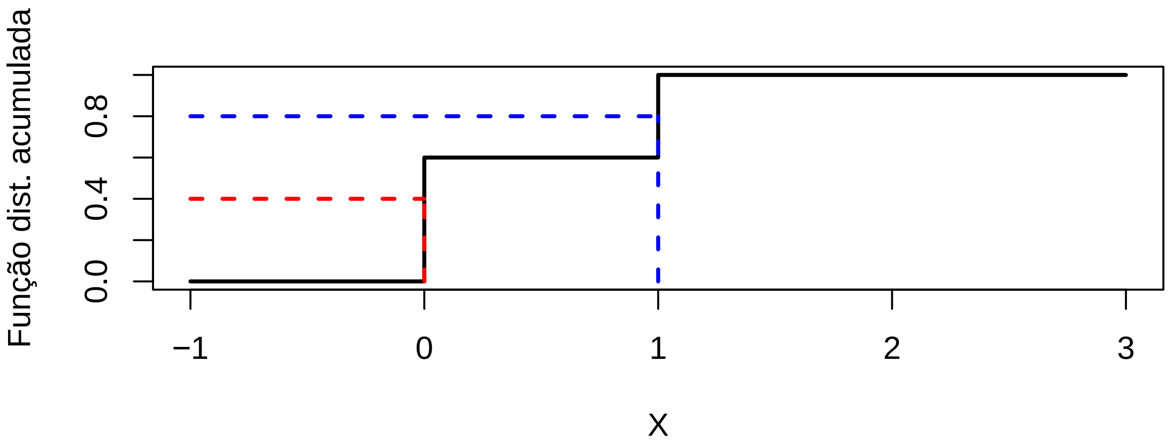
Ideia



Exemplo 1 - Adaptado de Rizzo (2007)

Utilize o método da transformação inversa para gerar 100 valores aleatórios da distribuição Bernoulli($p = 0,4$). Logo, $P(X = 0) = 0,6$ e $P(X = 1) = 0,4$.

Ideia



Note que $F_X(0) = P(X = 0) = 0,6$ e $F_X(1) = 1$. Logo, $F_X^{-1}(u) = 1$ se $u > 0,6$, e $F_X^{-1}(u) = 0$ se $u \leq 0,6$.

```r
set.seed(2023)
n <- 100
u <- runif(n)
head(u)
```

```
## [1] 0.46661394 0.33519095 0.16281756 0.39612002 0.03039173 0.12088487
```

```r
x <- ifelse(u > 0.6, 1, 0)
head(x)
```

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

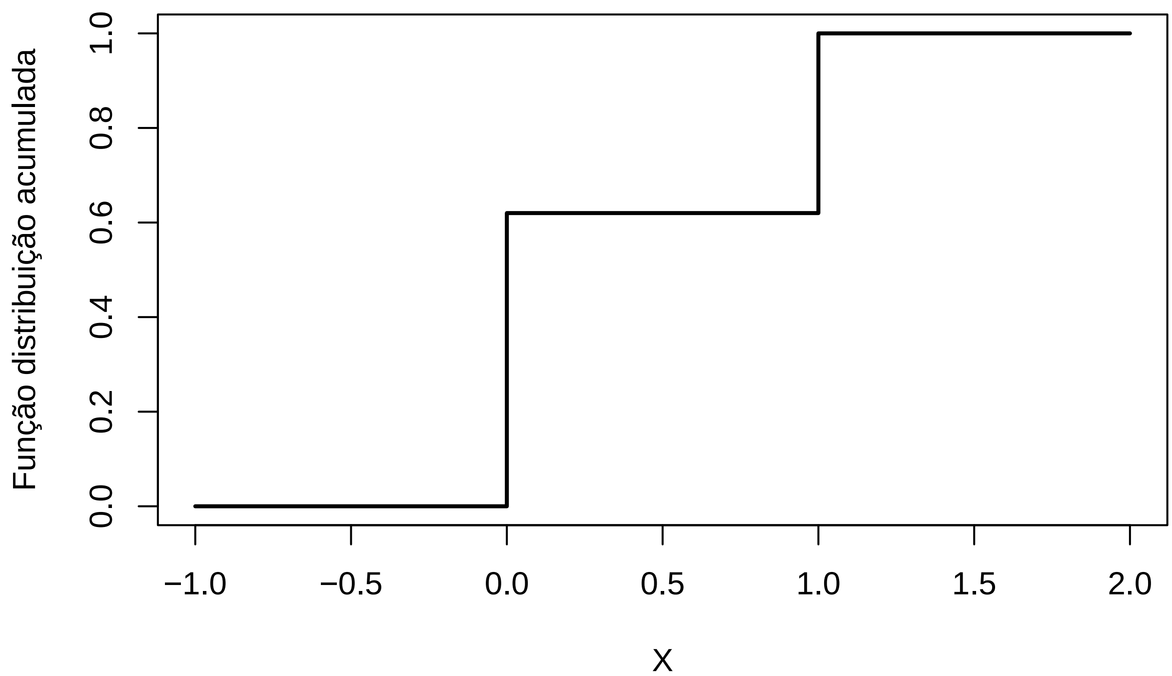
```r
# uma outra maneira é x <- as.integer(u > 0.6)

# resumo dos valores gerados
table(x)/n
```

```
## x
##    0    1
## 0.53 0.47
```

```r
cat("média=", mean(x))
```

```
## média= 0.47
```

```r
# Plot da função acumulada
plot(c(-1, 0, 1, 2), c(0, 0.62, 1, 1), lwd = 2, type = 's', ylab = "Função distribuição acumulada",
     xlab = "X", xlim = c(-1, 2))
```



Nota: veja a seguir outras formas de gerar valores da distribuição Bernoulli$(p = 0,4)$ **no R**

```r
set.seed(2023)
n <- 100

## usando a v.a. binomial

x1 <- rbinom(n, size = 1, prob = 0.4)
table(x1)
```

```
## x1
##  0  1
## 53 47
```

```r
cat("média=", mean(x1), "\n")
```

```
## média= 0.47
```

```r
set.seed(2023)
## usando a função sample
x2 <- sample(c(0, 1), size = n, replace = TRUE, prob = c(0.6, 0.4))
table(x2)
```

```
## x2
##  0  1
## 53 47
```

```r
cat("média=", mean(x2))
```

```
## média= 0.47
```

Exemplo 2 - Adaptado de Rizzo (2007)

Utilize o método da transformação inversa para gerar 1.000 valores aleatórios da distribuição geométrica com parâmetro $p = 1/4$.

Adote que a distribuição geométrica, com parâmetro p, é dada por

$$P(X = x) = pq^x,\ x = 0, 1, 2, \ldots,$$

sendo que $q = 1 - p$.

Note que os pontos de descontinuidade da distribuição acumulada são $x = 0, 1, 2, \ldots$ e a distribuição acumulada é dada por

$$F(x) = 1 - q^{x+1}, x \geq 0,$$

pois

1. se $x = 0$, $F(0) = 1 - q^{0+1} = 1 - q = 1 - (1 - p) = p$,
2. se $x = 1$, $F(1) = 1 - q^{1+1} = 1 - q^2 = 1 - (1 - p)^2 = 1 - (1 - 2p + p^2) = 2p - p^2 = p + p(1 - p) = p + pq$.

A ideia da demonstração deste resultado é feita por indução matemática.

Logo, para cada valor aleatório requerido precisamos gerar um valor aleatório u e encontrar o valor de x da expressão

$$\begin{aligned} F_X(x - 1) < u &\leq F_X(x) \\ 1 - q^x < u &\leq 1 - q^{x+1}. \end{aligned}$$

Resolvendo $1 - q^x < u$, temos $x < log(1 - u)/log(q)$. Enquanto que $u \leq 1 - q^{x+1}$ pode ser escrito como $log(1 - u)/log(q) \leq x + 1$, logo

$$x < log(1 - u)/log(q) \leq x + 1,$$

sendo que o valor de x é o maior valor inteiro que é menor ou igual a $log(1 - u)/log(q)$.

```r
set.seed(2023)
n <- 1000
u <- runif(n)
p <- 1/4
x <- floor(log(1 - u)/log( 1- p))

# resumo dos valores gerados

print("Resumos")
```

```
## [1] "Resumos"
```

```r
summary(x)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.000   2.000   2.877   4.000  44.000
```

```r
# Tabela dos valores gerados
print("Tabela dos valores gerados")
```

```
## [1] "Tabela dos valores gerados"
```

```r
table(x)/n
```

```
## x
##     0     1     2     3     4     5     6     7     8     9    10    11    12
## 0.264 0.190 0.141 0.108 0.081 0.045 0.053 0.038 0.020 0.016 0.006 0.005 0.008
##    13    14    15    16    17    18    19    20    23    25    44
## 0.005 0.002 0.005 0.005 0.001 0.001 0.002 0.001 0.001 0.001 0.001
```

```r
# A média da dist geo(p) é dada por (1-p)/p
# A variância da dist geo(p) é dada por (1-p)/p^2

cat("A média da dist geo(p) é ", (1-p)/p, "\n")
```

```
## A média da dist geo(p) é  3
```

```r
cat("A média cálculada a partir da amostra é ", mean(x), "\n")
```

```
## A média cálculada a partir da amostra é  2.877
```

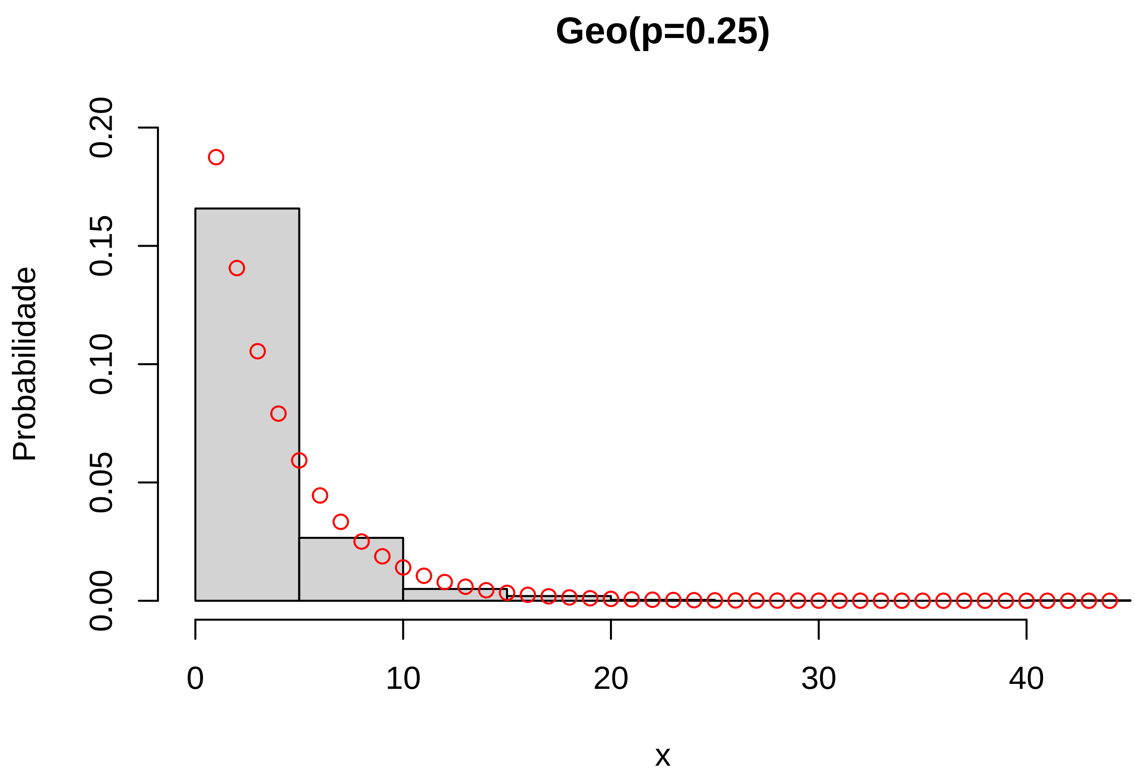
```r
cat("A variância da dist geo(p) é ", (1-p)/p^2, "\n")
```

```
## A variância da dist geo(p) é  12
```

```r
cat("A variância cálculada a partir da amostra é ", var(x), "\n")
```

```
## A variância cálculada a partir da amostra é  13.05292
```

```r
# densidade estimada pelo histograma
hist(x, prob=T, main="Geo(p=0.25)", ylab="Probabilidade", ylim=c(0, 0.2))
y <- seq(0, max(x))

# Calculo da probabilidade
points(y, p*(1-p)^y, col="red")
```



Exercícios

1. Utilize o método da transformação inversa para gerar 200 valores aleatórios da distribuição

x	0	1	2	3	4
P(X=x)	0,2	0,3	0,1	0,1	0,3

Faça comparação entre os valores gerados e os esperados.

2. Utilize o método da transformação inversa para gerar 1.000 valores aleatórios da distribuição binomial $(n = 3, p = 0.5)$. Faça comparação entre os valores gerados e os esperados.