Como gerar variáveis aleatórias discretas?

O Método da Transformação Inversa

ESTAT0090 – Estatística Computacional
Prof. Dr. Sadraque E. F. Lucena
sadraquelucena@academico.ufs.br

Cenário

Imagine que você está trabalhando com uma distribuição de probabilidade que foi definida recentemente. Por ser tão nova, ela ainda não tem funções de simulação em bibliotecas ou pacotes de programação. Para validar seus modelos, você precisa desenvolver o código do zero para gerar os dados aleatórios e realizar a sua simulação.

Objetivos da aula

Nesta aula, de hoje aprenderemos a gerar ocorrências de variáveis aleatórias discretas usando o método da transformação inversa.

Método da transformação inversa para variáveis discretas

O método da transformada inversa consiste em gerar um número aleatório entre 0 e 1 e encontrar o menor valor discreto cuja probabilidade acumulada é maior ou igual a esse número.

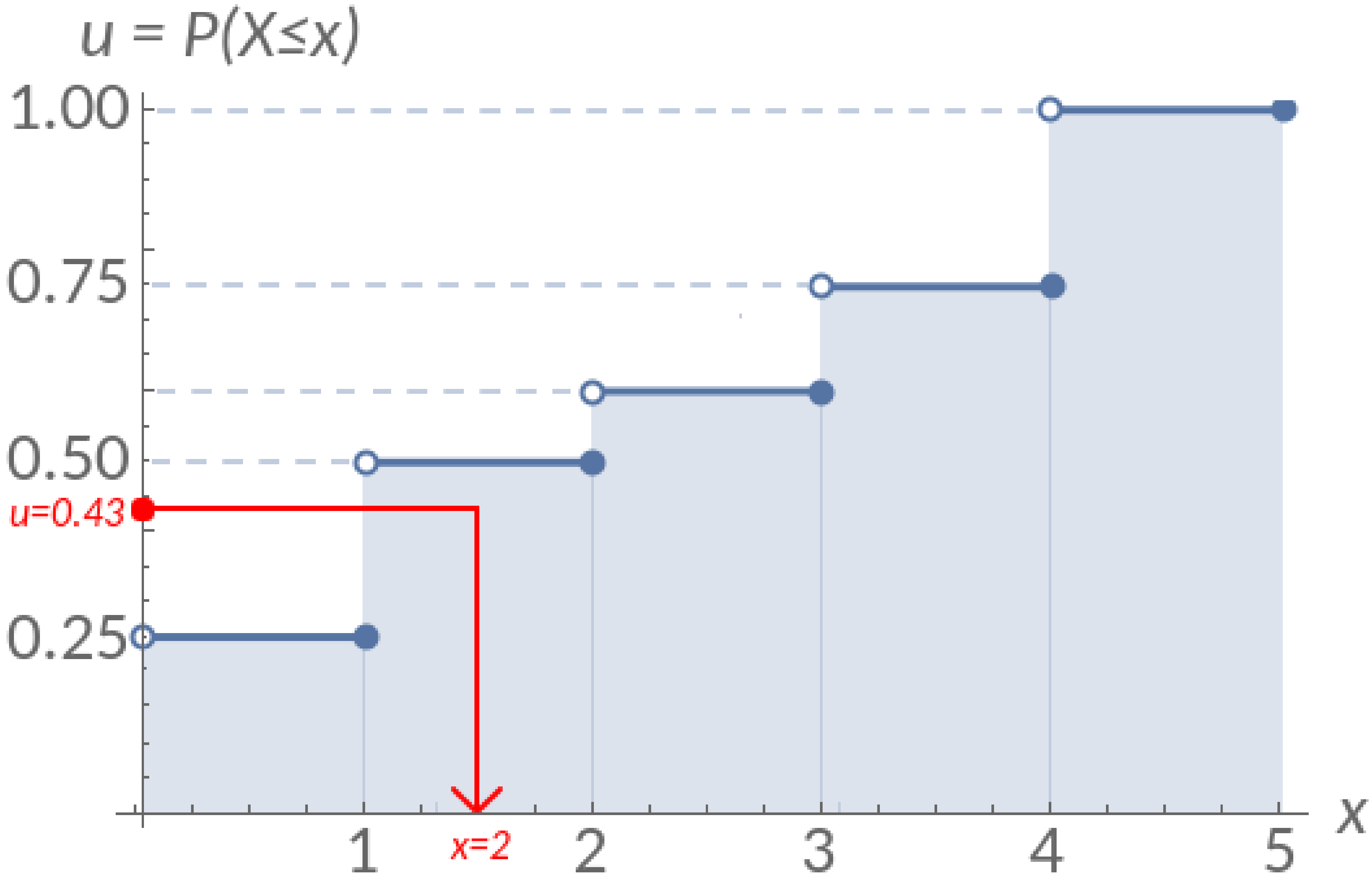
- Vejamos o método.

Método da transformação inversa para variáveis discretas

- Considere X uma variável aleatória tal que
  - $P(X = x_i) = p_i, \ i = 0, 1, \ldots$
  - $\sum_i p_i = 1.$
- Para gerar ocorrências de X, geramos um valor u de uma distribuição Uniforme$(0, 1)$ e obtemos um valor x fazendo

$$x = \begin{cases} x_0, & \text{se } u < p_0; \\ x_1, & \text{se } p_0 \leq u < p_0 + p_1; \\ x_2, & \text{se } p_0 + p_1 \leq u < p_0 + p_1 + p_2; \\ \vdots & \\ x_i, & \text{se } \sum_{j=0}^{i-1} p_j \leq u < \sum_{j=0}^{i} p_j; \\ \vdots & \end{cases}$$

Método da transformação inversa para variáveis discretas



Método da Transformada Inversa

Algoritmo

Passo 1: Gere $u \sim$ Uniforme$(0, 1)$;

Passo 2: Se $u < p_0$ faça $x = x_0$ e pare, caso contrário, vá para o próximo passo;

Passo 3: Se $u < p_0 + p_1$ faça $x = x_1$ e pare, caso contrário, vá para o próximo passo;

$\ldots$

Passo $k + 2$: Se $u < p_0 + p_1 + \cdots + p_k$ faça $x = x_k$ e pare, caso contrário, vá para o próximo passo.

Exemplo 10.1

Seja $X_i \sim \text{Bernoulli}(p)$.

a. Crie uma função no R que gera uma ocorrência de X. Teste para $p = 0{,}3$.
b. Crie uma função que gera o número de ocorrências da Bernoulli definida pelo usuário. Teste com $p = 0{,}3$ e $n = 1000$.

Exemplo 10.2

Seja X uma variável aleatória tal que

- $P(X = 1) = 0{,}2$
- $P(X = 2) = 0{,}15$
- $P(X = 3) = 0{,}25$
- $P(X = 4) = 0{,}4$.

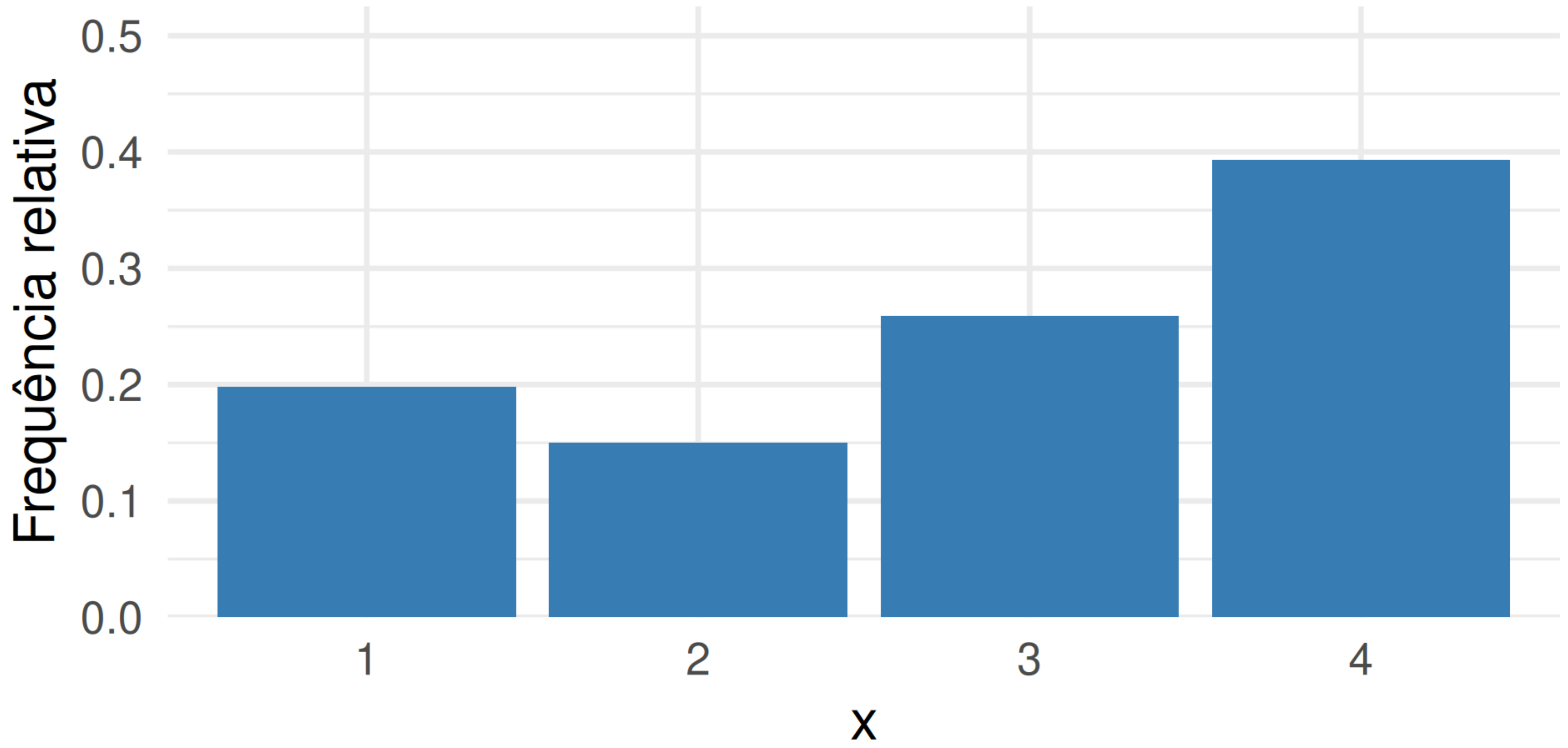
a. Crie uma função no R que gera uma ocorrência de X.

b. Crie uma função que gera o número de ocorrências definida pelo usuário. Teste com 1000 ocorrências e faça um gráfico de barras.

Um exemplo de código em R que implementa a função que gera os números aleatórios do Exemplo 10.2 é

```
gernum <- function(n, x, prob){
  amostra <- vector() # vetor que receberá valores gerados
  prob_acum <- cumsum(prob) # vetor de probabilidades acumuladas
  for (i in 1:n) { # início do laço
    u <- runif(1)  # valor de u
    for (j in 1:length(x)) {
      if (u <= prob_acum[j]) {
        amostra[i] <- x[j]
        break
      }
    }
  }
  return(amostra)
}

valores_x <- 1:4
probabilidades <- c(.2, .15, .25, .4)

set.seed(123) # fixando a semente para reprodutibilidade
amostra <- gernum(1000, valores_x, probabilidades)
```

```r
library(ggplot2)

# Calcular frequência relativa
df_freq <- as.data.frame(prop.table(table(amostra)))

# Renomear colunas para usar no ggplot
colnames(df_freq) <- c("x", "freq")
df_freq$x <- as.factor(df_freq$x)  # transformar em fator para manter
                                   # ordem no eixo x


# Gráfico de barras
ggplot(df_freq, aes(x = x, y = freq, fill = x)) +
  geom_col(fill = "#377EB8") +
  scale_y_continuous(limits = c(0, 0.5),
                     expand = expansion(mult = c(0, 0.05))) +
  labs(x = "x", y = "Frequência relativa") +
  theme_minimal(base_size = 25) +
  theme(legend.position = "none")
```


0.5
0.4
0.3
0.2
0.1
0.0
Frequência relativa
1
2
3
4
x

Exemplo 10.3

Sejam $X_1, \ldots, X_m$ v.a. tais que $X_i \sim$ Bernoulli(p).

Então $Y = \sum_{i=1}^{m} X_i \sim$ Binomial(m, p).

a. Crie uma função no R que gera uma ocorrência de Y. Teste com $m = 10$ e $p = 0{,}3$.

b. Crie uma função que gera o número de ocorrências definida pelo usuário. Teste com 1000 ocorrências e faça um gráfico.

Exemplo 10.4

Seja $X \sim$ Geométrica(p) tal que $P(X = i) = p(1 - p)^{i-1}$, para $i = 1, 2, \ldots$

a. Crie uma função no R que gera uma ocorrência de X. Teste com $p = 0{,}4$.
b. Crie uma função que gera o número de ocorrências definida pelo usuário. Teste com 1000 ocorrências e faça um gráfico.

Ganho da aula

- Compreensão da geração de números aleatórios discretos usando o método da transformação inversa.

Fim

Esta aula foi baseada no livro *Simulation (Sixth Edition)*, de Sheldon M. Ross, 2023.