

# Laborator 08

## Petculescu Mihai-Silviu

### Laborator 08

Petculescu Mihai-Silviu

Generarea unei variabile aleatoare discrete

Exemplu 1

Exemplu 2

Aplicație

Funcția de repartiție pentru variabile aleatoare

Aplicație

## Generarea unei variabile aleatoare discrete

Definiți o funcție care să genereze un eșantion de talie  $n$  dintr-o distribuție discretă definită pe mulțimea  $\{x_1, \dots, x_N\}$  cu probabilitățile  $\{p_1, \dots, p_N\}$ .

```
> GenerateDiscrete = function(n = 1, x, p, err = 1e-15) {  
  # x alfabetul ; # p probabilitatile  
  lp = length(p)  
  lx = length(x)  
  # verify if x and p have the same size  
  if(abs(sum(p) - 1) > err | sum(p >= 0) != lp)  
    stop("suma probabilitatilor nu este 1 || Probabilitatile sunt mai mici decat  
0")  
  else if(lx != lp)  
    stop("x si p ar trebui sa aiba aceeasi marime")  
  else {  
    out = rep(0, n)  
    indOrderProb = order(p, decreasing = TRUE) # index  
    pordered = p[indOrderProb] # rearrange the values of the probabilities  
    xordered = x[indOrderProb] # rearrange the values of x  
    porderedCS = cumsum(pordered)  
    for (i in 1:n) {  
      u = runif(1) # runif(n) - generate n uniforms  
      k = min(which(u <= porderedCS))  
      out[i] = xordered[k]  
    }  
  }  
  return(out)  
}
```

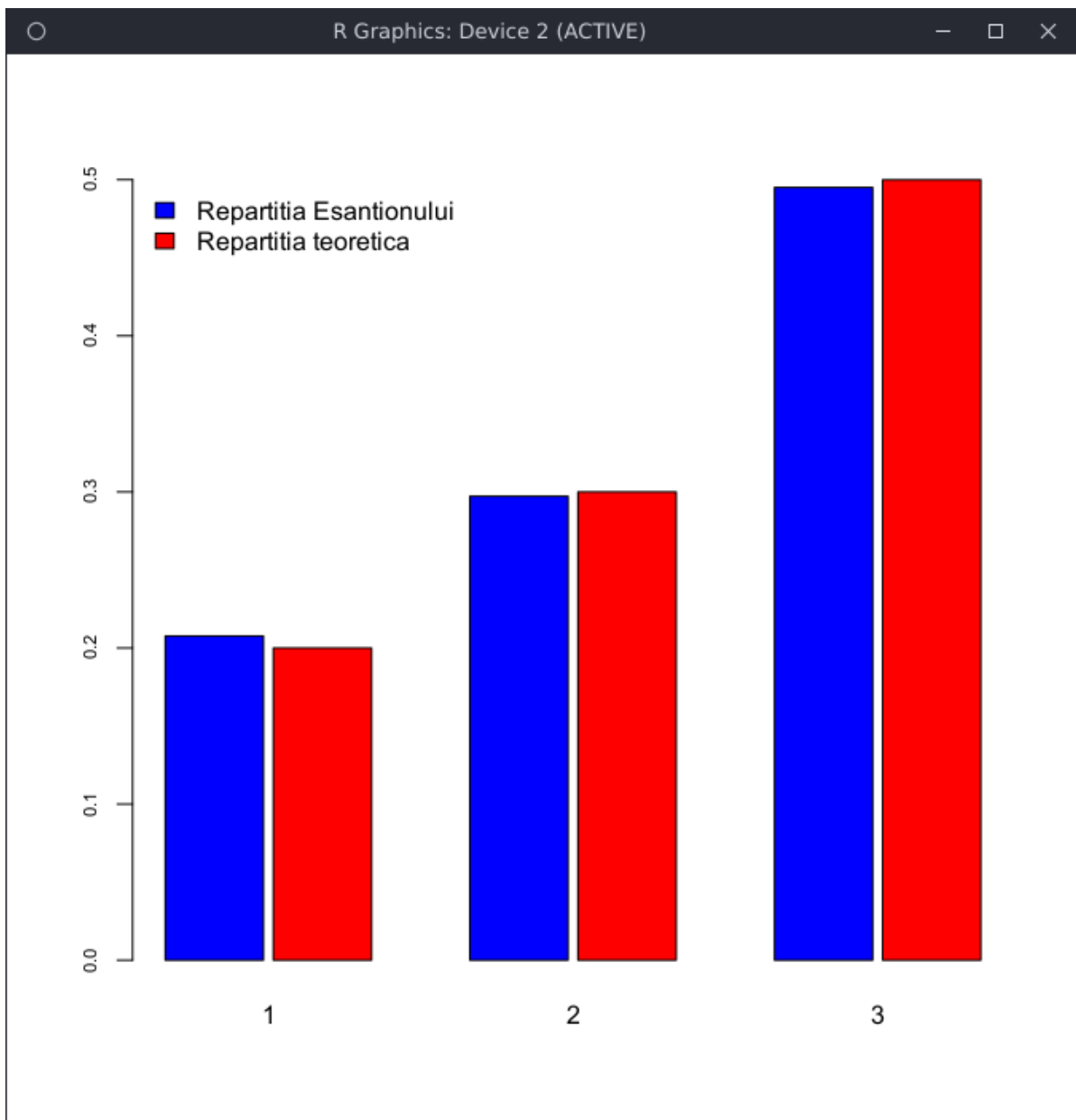
### Exemplu 1

Ne propunem să generăm observații din  $X \sim \begin{pmatrix} 1 & 2 & 3 \\ 0.2 & 0.3 & 0.5 \end{pmatrix}$ , în acest caz:  $x = [1, 2, 3]$  și  $p = [0.2, 0.3, 0.5]$ . Începem prin generarea a  $n = 10$  observații din repartiția lui  $X$ :

```
> GenerateDiscrete(10, c(1:3), c(0.2, 0.3, 0.5))  
[1] 3 2 3 3 3 1 2 2 3 3
```

Plecând de la un eșantion de  $n = 10000$  de observații vrem să comparăm, cu ajutorul diagramei cu bare verticale (`barplot`), repartiția eșantionului cu cea teoretică:

```
> n = 10000  
> x = GenerateDiscrete(n, c(1:3), c(0.2, 0.3, 0.5))  
> pX = table(x) / n  
> pT = c(0.2, 0.3, 0.5)  
> indx = c(1,2,3)  
> barplot(rbind(pX, pT), beside = T, space = c(0.1, 1), col = c("blue", "red"),  
names.arg = indx, cex.axis = 0.7, legend.text = c("Repartitia Esantionului",  
"Repartitia teoretica"), args.legend = list(x = "topleft", bty = "n"))
```



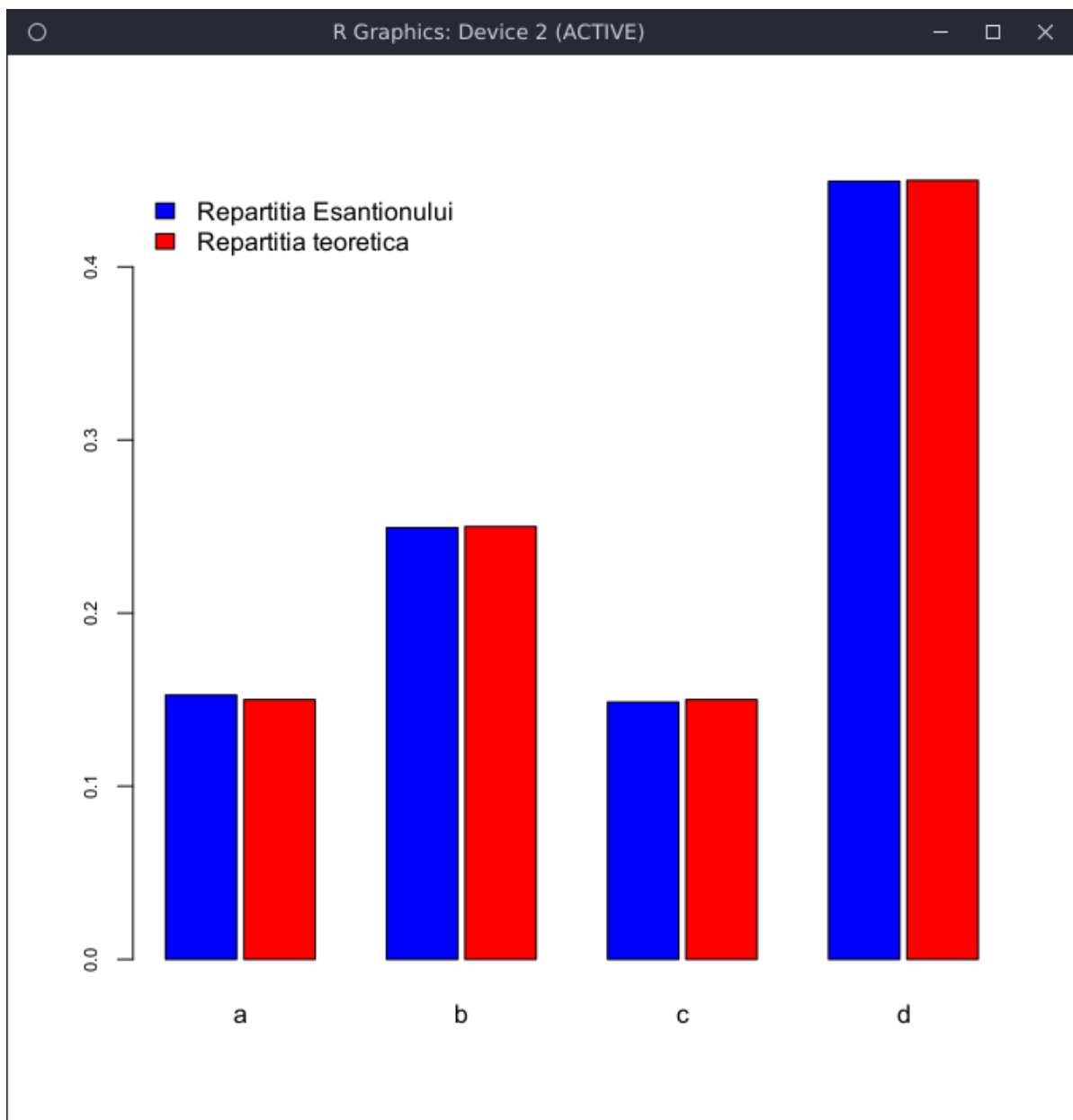
## Exemplu 2

În acest caz considerăm variabila aleatoare  $X \sim \begin{pmatrix} a & b & c & d \\ 0.15 & 0.25 & 0.15 & 0.45 \end{pmatrix}$ , deci  $x = [a, b, c, d]$  și  $p = [0.15, 0.25, 0.15, 0.45]$ . Mai jos generăm  $n = 15$  observații din repartiția variabilei aleatoare  $X$ :

```
> GenerateDiscrete(15, c("a", "b", "c", "d"), c(0.15, 0.25, 0.15, 0.45))  
[1] "a" "d" "a" "b" "c" "c" "d" "b" "d" "d" "b" "b" "d" "c" "d"
```

Ca și în cazul primului exemplu, vom compara repartiția teoretică cu cea a unui eșantion de  $n = 10000$  de observații:

```
> n = 10000  
> x = GenerateDiscrete(n, c("a", "b", "c", "d"), c(0.15, 0.25, 0.15, 0.45))  
> px = table(x) / n  
> pT = c(0.15, 0.25, 0.15, 0.45)  
> indx = c("a", "b", "c", "d")  
> barplot(rbind(px, pT), beside = T, space = c(0.1, 1), col = c("blue", "red"),  
names.arg = indx, cex.axis = 0.7, legend.text = c("Repartitia Esantionului",  
"Repartitia teoretica"), args.legend = list(x = "topleft", bty = "n"))
```



## Aplicație

Pentru variabila aleatoare  $X$  cu distribuția următoare  $X \sim \begin{pmatrix} \frac{1}{2} & 1 & 2 \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \end{pmatrix}$

a) Generați  $n = 12$  observații din repartiția variabilei aleatoare  $X$ .

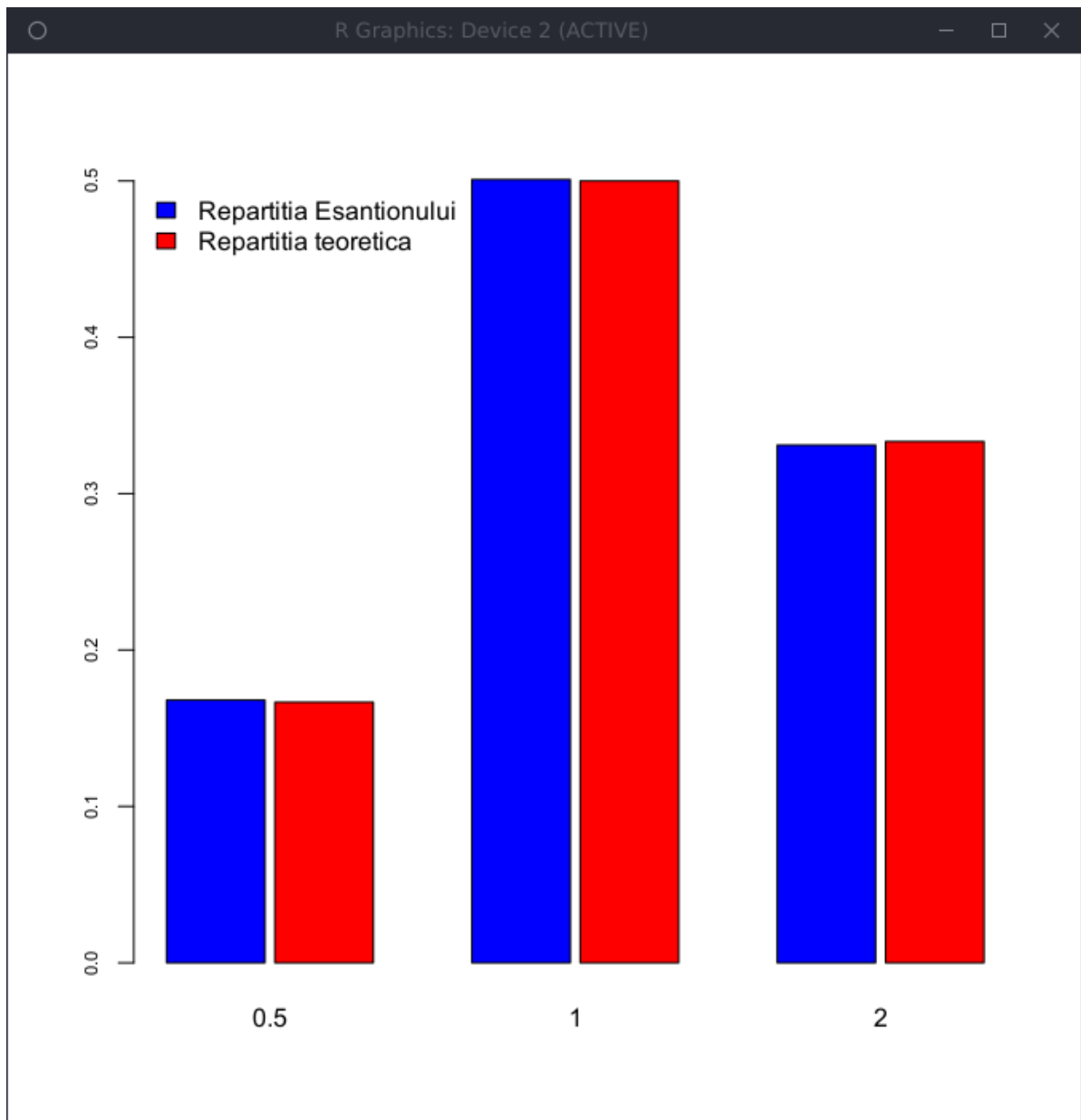
```
> GenerateDiscrete(12, c(1/2, 1, 2), c(1/6, 1/2, 1/3))  
[1] 0.5 0.5 2.0 0.5 0.5 2.0 1.0 2.0 1.0 2.0 1.0 2.0
```

b) Plecând de la un eșantion de  $n = 10000$  de observații comparați repartiția eșantionului cu cea teoretică.

```

> n = 10000
> x = GenerateDiscrete(n, c(1/2, 1, 2), c(1/6, 1/2, 1/3))
> pX = table(x) / n
> pT = c(1/6, 1/2, 1/3)
> indx = c(1/2, 1, 2)
> barplot(rbind(pX, pT), beside = T, space = c(0.1, 1), col = c("blue", "red"),
names.arg = indx, cex.axis = 0.7, legend.text = c("Repartitia Esantionului",
"Repartitia teoretica"), args.legend = list(x = "topleft", bty = "n"))

```



## Funcția de repartiție pentru variabile aleatoare

Scrieți o funcție în R care să traseze graficul funcției de repartiție a unei distribuții date. Verificați și documentația funcției `ecdf`.

```

> cdfPlot = function(dist, title, err = 1e-10) {
  # dist - repartitia discreta (sau discretizata)
  lp = length(dist)
  # dist[1,] in loc de dist
  if (abs(sum(dist[1,]) - 1) > err | sum(dist >= 0) != lp)
    stop("Eroare: vectorul de probabilitati nu formeaza o repartitie")
  else {

```

```

x = 0:(lp-1) # ia valori in 1:lp
cp = cumsum(dist)
plot(x, cp, type = "s", lty = 3, xlab = "x", ylab = "F",
     main = paste("Functia de repartitie:", title),
     ylim = c(0,1), col = "grey", bty = "n"
)
abline(h = 0, lty = 2, col = "grey")
abline(h = 1, lty = 2, col = "grey")
for( i in 1:(lp-1) ) {
  lines( c(x[i], x[i+1]), c(cp[i], cp[i]), col = "blue", lwd = 2 )
}
points(x, cp, col = "red", pch = 20, cex = 0.85)
}
}

```

## Aplicație

Fie variabila aleatoare discretă simplă  $X$  cu distribuția următoare

$X \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 0.1 & 0.2 & 0.3 & 0.1 & 0.2 & 0.1 \end{pmatrix}$ . Determinați funcția de repartitie a variabilei

aleatoare  $X$  și reprezentați grafic variabila și funcția ei de repartitie.

```

> x = GenerateDiscrete(n, c(1:6), c(0.1, 0.2, 0.3, 0.1, 0.2, 0.1))
> pX = table(x) / n
> pT = c(0.1, 0.2, 0.3, 0.1, 0.2, 0.1)
> cdfPlot(rbind(pX, pT), "")

```

