Laboratorul 5

1. Realizaţi un program care generează N numere pseudo-aleatoare pentru variabila aleatoare discretă X, care are distribuţia de probabilitate

$$X \sim \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix},$$

folosind indicațiile de pe pagina 2 și

[ ]:
```
from scipy.stats import uniform
```

Aplicaţie: Conform statisticilor medicale, 46% din oameni au grupa sanguină **0**, 40% au grupa sanguină **A**, 10% au grupa sanguină **B** şi 4% au grupa sanguină **AB**.

Simulaţi de N ori observarea grupei sanguine a unei persoane alese aleator şi afişaţi frecvenţa relativă de apariţie a fiecărei grupe sanguine. Afişaţi histograma datelor obţinute și barele corespunzătoare probabilităților teoretice.

[ ]:
```
from matplotlib.pyplot import bar, show, hist, grid, legend, xticks , yticks
```

2. Realizaţi un program care generează N numere pseudo-aleatoare pentru variabila aleatoare $X \sim Exp(\alpha)$, unde $\alpha > 0$, folosind indicațiile de pe pagina 2 și

[ ]:
```
from scipy.stats import uniform
from math import log
```

Aplicaţie: Timpul T (în secunde) necesar ca o imprimantă să printeze un afiş are distribuţia exponenţială cu parametrul $\frac{1}{12}$. Simulaţi de N ori timpul de printare al unui afiş.
i) Afișați o histogramă cu 12 bare, pe intervalul $[0, 60]$, pentru datele obținute, apoi desenați, în aceeași figură, graficul funcției de densitate, completând codul următor:

[ ]:
```
from matplotlib.pyplot import show, hist, grid, legend, xticks, plot
from scipy.stats import expon

alpha = ?

data = ?
hist(?, bins=12, density = True,range=(?,?))

x = range(60)
plot(x,expon.pdf(x,loc=?,scale=?),'r-')

xticks(range(0,60,5))
grid()
show()
```

ii) Estimați probabilitatea evenimentului E: "timpul de printare al afișului este cel puțin 5 secunde". Afișați valoarea teoretică pentru $P(E)$.

I. Generarea de numere pseudo-aleatoare pentru o distribuţie discretă (metoda inversei)

• Input: valorile $[x_1, x_2, \ldots, x_n]$, probabilităţile corespunzătoare $[p_1, p_2, \ldots, p_n]$ şi numărul N. Fie $p_0 = 0$.

• Se generează N numere aleatoare pentru distribuţia uniformă $Unif[0,1]$: $U[i]$, $i = \overline{1,N}$.

• Pentru fiecare $i = \overline{1,N}$: $X[i] = x_k$ dacă şi numai dacă

$$p_0 + p_1 + \cdots + p_{k-1} < U[i] \leq p_0 + p_1 + \cdots + p_k, \text{ unde } \ k \in \{1, \ldots, n\}.$$

• Output: $X[i], i = \overline{1,N}$.

Verificarea procedeului: Pentru fiecare $i = \overline{1,N}$ şi $k = \overline{1,n}$:
$P(X[i] = x_k) = P(\text{“se generează } \ x_k\text{”}) = P(p_0 + p_1 + \cdots + p_{k-1} < U[i] \leq p_0 + p_1 + \cdots + p_k) = p_k$.

II. Generarea de numere pseudo-aleatoare pentru o distribuţie continuă (metoda inversei)

Fie X o variabilă aleatoare continuă care are funcţia de repartiţie F astfel încât există F^{-1} pe $(0,1)$: pentru orice $u \in (0,1)$ există un unic $x \in \mathbb{R}$ astfel încât $F(x) = u$ ($\Leftrightarrow F^{-1}(u) = x$).

• Input: funcţia F^{-1} şi numărul N.

• Se generează N numere aleatoare pentru distribuţia uniformă $Unif[0,1]$: $U[i]$, $i = \overline{1,N}$.

• Pentru fiecare $i = \overline{1,N}$: $X[i] = F^{-1}(U[i])$.

• Output: $X[i], i = \overline{1,N}$.

Verificarea procedeului pentru $X \sim Exp(\alpha)$, unde $\alpha > 0$:
Avem funcţia de densitate pentru X: $f(t) = \begin{cases} \alpha e^{-\alpha t}, & t > 0 \\ 0, & t \leq 0 \end{cases} \implies$ funcţia de repartiţie a lui X este $F(x) = \displaystyle\int_{-\infty}^{x} f(t)\,dt = \begin{cases} \displaystyle\int_0^x \alpha e^{-\alpha t}\,dt, & x > 0 \\ 0, & x \leq 0 \end{cases} = \begin{cases} 1 - e^{-\alpha x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$.
Observăm că pentru orice $u \in (0,1)$:
$F(x) = u \Leftrightarrow 1 - e^{-\alpha x} = u \Leftrightarrow 1 - u = e^{-\alpha x} \Leftrightarrow \ln(1-u) = -\alpha x \Leftrightarrow -\frac{1}{\alpha}\log(1-u) = x$. Deci,

$$F^{-1}(u) = -\frac{1}{\alpha}\log(1-u), u \in (0,1).$$

Pentru $i = \overline{1,N}$, arătăm că $X[i]$ are funcţia de repartiţie a lui X: $\forall x \in \mathbb{R}$, $P(X[i] \leq x) = P(F^{-1}(U[i]) \leq x) = P(U[i] \leq F(x)) = F(x)$, deci $X[i]$ are aceeaşi distribuţie ca X.