Seminarul 4

1. Fie S mulțimea tuturor numerelor naturale cel mult egale cu 50, cu exact două cifre de parități diferite. Un număr este ales aleator din S. Fie X suma cifrelor numărului ales. Scrieți distribuția lui X, apoi calculați valoarea sa medie $E(X)$.

2. Considerăm următoarea problemă de *clasificare naivă Bayes* a unor restaurante (**R**), în
• *clasele*: **r**ecomandat sau **n**erecomandat,
în funcţie de următoarele *atribute* cu valorile lor posibile:
• cost (C): ieftin, mediu, scump;
• timp de aşteptare (T): puţin, mediu, îndelungat;
• mâncare (M): fadă, acceptabilă, bună, delicioasă.

R, C, T, M sunt variabilele aleatoare (categoriale) şi **r**, **n**, i, m, s, p, m, $\hat{\imath}$, f, a, b, d valorile de mai sus, în ordinarea în care sunt menţionate.

Considerăm următorul *tabel de date* furnizat de clienţii unor restaurante:

	Cost	*Timp de aşteptare*	*Mâncare*	**Restaurant**
1	mediu	îndelungat	acceptabilă	**nerecomandat**
2	scump	puţin	bună	**recomandat**
3	ieftin	îndelungat	delicioasă	**recomandat**
4	mediu	puţin	bună	**recomandat**
5	ieftin	mediu	acceptabilă	**nerecomandat**
6	ieftin	puţin	fadă	**nerecomandat**
7	mediu	puţin	acceptabilă	**nerecomandat**
8	mediu	mediu	delicioasă	**recomandat**
9	scump	puţin	delicioasă	**recomandat**
10	ieftin	îndelungat	bună	**nerecomandat**
11	scump	puţin	acceptabilă	**nerecomandat**
12	mediu	mediu	bună	**recomandat**
13	mediu	îndelungat	fadă	**nerecomandat**
14	scump	mediu	delicioasă	**recomandat**
15	ieftin	mediu	fadă	**nerecomandat**
16	mediu	puţin	delicioasă	**recomandat**
17	ieftin	puţin	acceptabilă	**recomandat**
18	scump	îndelungat	bună	**nerecomandat**
19	ieftin	puţin	fadă	**recomandat**
20	scump	îndelungat	delicioasă	**nerecomandat**

i) Folosind datele din tabel, determinaţi probabilităţile claselor şi probabilităţile condiţionate ale atributelor, ştiind clasa.
ii) Considerăm evenimentul dat de *vectorul de atribute*: $E = (C = s) \cap (T = m) \cap (M = b)$. Alegeţi o clasă pentru E, stabilind care din următoarele probabilităţi este mai mare: $P(\mathbf{R} = \mathbf{r}|E)$ sau $P(\mathbf{R} = \mathbf{n}|E)$.
iii) Determinaţi $P(E)$.

3. Ce valoare teoretică estimează programul următor? Calculați valoarea teoretică corespunzătoare.

[ ]:
```
import numpy as np

N=2000
S = np.concatenate((np.zeros(50),np.ones(70),2*np.ones(80)))
X=[]
for _ in range(N):
    k=0
    i= np.random.randint(len(S))
    while S[i] != 0:
        i= np.random.randint(len(S))
        k=k+1
    X.append(k)

print(" . . . . . :",np.mean(X))
```

4. Fie $a, b \in \mathbb{Z}$, $a \leq b$, și $c \in (0,1)$. Spunem că variabila aleatore X are o distribuție uniform discretă dacă

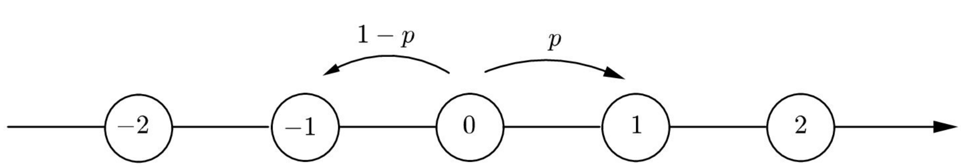
$$X \sim \begin{pmatrix} a & a+1 & \ldots & b \\ c & c & \ldots & c \end{pmatrix}.$$

a) Determinați valoarea lui c.
b) Pentru $a = 3$ și $b = 21$, calculați

$$P\Big(\Big\{X \leq \frac{a+b}{2}\Big\} \cup \Big\{\frac{a+b}{6} \leq X\Big\}\Big) \text{ și } P\Big(\Big\{X \leq \frac{a+b}{2}\Big\} \cap \Big\{\frac{a+b}{6} \leq X\Big\}\Big).$$

c) Determinați a și b, știind că $P(X = a) = \frac{1}{3}$ și $E(X) = 1$.

5. Un punct material se deplasează pe axa reală dintr-un nod spre un nod vecin, la fiecare pas, cu probabilitatea $p \in (0,1)$ la dreapta şi cu probabilitea $1 - p$ la stânga. Nodurile sunt centrate în numerele întregi:



Fie X variabila aleatoare care indică poziţia finală a punctului material după $n \in \mathbb{N}$ paşi ai unei deplasări ce porneşte din nodul 0. Determinaţi distribuţia şi valoarea medie lui X.

6. Considerăm vectorul aleatoar discret (X, Y) cu distribuţia dată sub formă tabelară:

$X \backslash Y$	-2	1	2
1	0,2	0,1	0,2
2	0,1	0,1	0,3

a) Să se determine distribuţiile de probabilitate ale variabilelor aleatoare X şi Y.
b) Calculaţi probabilitatea ca $|X - Y| = 1$, ştiind că $Y > 0$.
c) Sunt evenimentele $X = 2$ şi $Y = 1$ independente?
d) Sunt variabilele aleatoare X şi Y independente?
e) Sunt evenimentele $X = 1$ şi $Y = 1$ condiţional independente, cunoscând $X + Y = 2$?
f) Este variabila aleatoare X condiţional independentă de Y, cunoscând $X + Y$?
g) Calculaţi valoarea medie a variabilei aleatoare $2X + Y^2$.

7. O monedă este aruncată de 10 ori. Fie X variabila aleatoare care indică diferenţa dintre numărul de capete şi numărul de pajuri obţinute. Determinaţi:
i) distribuţia de probabilitate a lui X;
ii) valoarea medie a lui X.