Seminarul 6

1. Un jucător de darts ocheşte discul roşu (denumit “bullseye”) cu centrul în centrul ţintei şi diametru 1 cm. La o aruncare, distanţa dintre centrul ţintei şi punctul nimerit de săgeata jucătorului urmează distribuţia uniformă pe intervalul $[a, b]$, unde $0 \leq a < b$, cu valoarea medie $\frac{3}{2}$ cm şi deviaţia standard $\frac{\sqrt{3}}{2}$ cm. Aruncările jucătorului sunt independente. Determinaţi:
a) probabilitatea ca jucătorul să nimerească discul roşu;
b) probabilitatea ca jucătorul să nimerească de 2 ori discul roşu din 10 aruncări.

Funcţia de densitate pentru distribuţia uniformă $Unif[a, b]$ este $f : \mathbb{R} \to \mathbb{R}$ definită prin $f(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b] \\ 0, & x \notin [a, b] \end{cases}$.

R: a) X=distanţa de la săgeată la centru $\implies f_X = f$ este funcţie de densitate pentru $X \implies E(X) = \int_a^b \frac{x}{b-a}dx = \frac{a+b}{2} = \frac{3}{2}, V(X) = E(X^2) - E(X)^2 = \int_a^b \frac{x^2}{b-a}dx - \frac{(a+b)^2}{4} = \frac{a^2+b^2-2ab}{12} = \frac{3}{4}$. Avem: $\begin{cases} a + b = 3 \\ ab = 0 \end{cases} \implies$ $a = 0, b = 3$. p=probabilitatea de a nimeri discul roşu $\implies p = P(X \leq \frac{1}{2}) = \int_0^{\frac{1}{2}} \frac{1}{b-a}dx = \frac{1}{6}$.
b) Z=numărul de reuşite din 10 aruncări $\implies Z \sim Bino(10, p) \implies P(Z = 2) = C_{10}^2 p^2(1 - p)^8 = 45 \cdot \frac{5^8}{6^{10}} \approx 29\%$.

2. a) Fie datele statistice $(x_i)_{i=\overline{1,10}}$: 2, 1, 3, 1, 5, 2 ,3, 5, 1, 1. Să se calculeze expresia funcţiei de repartiţie empirice corespunzătoare acestor date

$$\mathcal{F}_{10} : \mathbb{R} \to [0, 1] \text{ definită prin } \mathcal{F}_{10}(x) = \frac{\# \{i \in \{1, \ldots, 10\} : x_i \leq x\}}{10}.$$

R.: $\mathcal{F}_{10} : \mathbb{R} \to [0, 1]$ $\qquad \mathcal{F}_{10}(x) = \begin{cases} 0, & \text{dacă } x < 1 \\ 0{,}4 & \text{dacă } 1 \leq x < 2 \\ 0{,}4 + 0{,}2 & \text{dacă } 2 \leq x < 3 \\ 0{,}4 + 0{,}2 + 0{,}2, & \text{dacă } 3 \leq x < 5 \\ 1, & \text{dacă } 5 \leq x. \end{cases}$

b) Fie $(X_n)_n$ un şir de variabile aleatoare independente care au aceeaşi distribuţie. Notăm cu F funcţia de repartiţie comună.

b$_1$) Fie $x \in \mathbb{R}$ fixat şi se consideră pentru $n \in \mathbb{N}^*$ v.a. $Y_n(\omega) = \begin{cases} 1, & \text{dacă } X_n(\omega) \leq x \\ 0, & \text{dacă } X_n(\omega) > x. \end{cases}$

Ce distribuţie au Y_n, respectiv $Y_1 + ... + Y_n$?
R.: $Y_n \sim Bernoulli(F(x))$, $Y_1 + ... + Y_n \sim Bino(n, F(x))$.
b$_2$) Spre ce valoare converge a.s. şirul $\left(\frac{1}{n}(Y_1 + ... + Y_n)\right)_n$?

R.: $\frac{1}{n}(Y_1 + ... + Y_n) \xrightarrow{a.s.} F(x)$, folosind LTNM, pentru şirul $(Y_n)_n$, care este un şir de v.a. independente şi $E(Y_n) = P(X_n \leq x) = F(x), \forall n \in \mathbb{N}^*$.
b$_3$) Pentru $n \in \mathbb{N}^*$ fie

$$\mathcal{F}_n : \mathbb{R} \times \Omega \to [0, 1] \qquad \mathcal{F}_n(x, \omega) = \frac{\# \{i \in \{1, \ldots, n\} : X_i(\omega) \leq x\}}{n},$$

funcţia de repartiţie empirică calculată în punctul $x \in \mathbb{R}$.

Ce relaţie există între cele două v.a. $\frac{1}{n}(Y_1 + ... + Y_n)$ şi $\mathcal{F}_n(x, \cdot)$?

R.: $\frac{1}{n}(Y_1(\omega) + ... + Y_n(\omega)) = \mathcal{F}_n(x, \omega)\ \forall \omega \in \Omega$.

b$_4$) Este $\mathcal{F}_n(x, \cdot)$ un *estimator nedeplasat* şi *consistent* pentru $F(x)$?

R.: Da, pentru că $E(\mathcal{F}_n(x, \cdot)) = E\left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(Y_i) = F(x)$ şi b$_2$) implică $\mathcal{F}_n(x, \cdot) = \frac{1}{n}\sum_{i=1}^{n} Y_i \xrightarrow{a.s.} F(x)$.

3. Durata (în minute) a unei plăţi pentru o factură la un ghişeu într-o bancă urmează distribuţia continuă $Unif[1, 3]$. Ştiind că duratele oricăror plăţi sunt independente, demonstraţi că:

i) media aritmetică a duratelor plăţilor a n facturi converge a.s. la 2 minute, când $n \to \infty$.
ii) media geometrică a duratelor plăţilor a n facturi converge a.s. la $\frac{3\sqrt{3}}{e}$ minute, când $n \to \infty$.
iii) media armonică a duratelor plăţilor a n facturi converge a.s. la $\frac{2}{\ln 3}$ minute, când $n \to \infty$.
R: Fie X_n durata plăţii celei de a n-a facturi. $(X_n)_n$ este un şir de variabile aleatoare independente care urmează disitribuţia $Unif[1,3]$.
i) LTNM implică $\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{a.s.} E(X_1) = \int_1^3 \frac{x}{2}\,dx = 2.$
ii) LTNM implică $\sqrt[n]{\prod_{i=1}^{n} X_i} = e^{\frac{1}{n}\sum_{i=1}^{n}\ln X_i} \xrightarrow{a.s.} e^{E(\ln X_1)} = e^{\int_1^3 \frac{\ln x}{2}dx} = \frac{3\sqrt{3}}{e} \approx 1,91.$
iii) LTNM implică $\frac{n}{\sum_{i=1}^{n}\frac{1}{X_i}} = \frac{1}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i}} \xrightarrow{a.s.} \frac{1}{E(\frac{1}{X_1})} = \frac{1}{\int_1^3 \frac{1}{2x}dx} = \frac{2}{\ln 3} \approx 1{,}82.$

4. Un computer este conectat la două imprimante: I_1 and I_2. Computerul trimite printarea unui document lui I_1 cu probabilitatea 0,4, respectiv lui I_2 cu probabilitatea 0,6. Ştiind că a fost aleasă imprimanta I_1, un poster $A2$ este printat în T_1 secunde, unde T_1 are distribuţia $Exp(\frac{1}{5})$. Ştiind că a fost aleasă imprimanta I_2, un poster $A2$ este printat în T_2 secunde, unde T_2 are distribuţia uniformă $Unif[4,6]$. Un inginer solicită printarea unui poster A2 de pe computer. Calculaţi valoarea medie şi deviaţia standard pentru timpul (în secunde) de printare a posterului.
R.: T=timpul de printare a posterului; F_T=funcţia de repartiţie a lui T; fie evenimentele A: computerul este conectat la imprimanta I_1; $\bar{A}$: computerul este conectat la imprimanta I_2. Formula probabilităţii totale $\Longrightarrow$

$$\begin{aligned} F_T(t) &= P(T \le t) = P(T \le t|A)P(A) + P(T \le t|\bar{A})P(\bar{A}) \\ &= 0{,}4 \cdot P(T_1 \le t) + 0{,}6 \cdot P(T_2 \le t) = 0{,}4\int_{-\infty}^{t} f_{T_1}(\tau)d\tau + 0{,}6\int_{-\infty}^{t} f_{T_2}(\tau)d\tau,\ t \in \mathbb{R}, \end{aligned}$$

unde f_{T_i} este funcţie de densitate pentru T_i, $i = 1,2 \Longrightarrow f_T(t) = F'(t) = 0{,}4 f_{T_1}(t) + 0{,}6 f_{T_2}(t), t \in \mathbb{R}$, este funcţie de densitate pentru $T \Longrightarrow E(T) = \int_{-\infty}^{\infty} t f_T(t)dt = 0{,}4\int_0^{\infty}\frac{1}{5}te^{-\frac{t}{5}}dt + 0{,}6\int_4^6 \frac{t}{2}dt = 0{,}4\cdot 5 + 0{,}6 \cdot 5 = 5$ (secunde).

5. Fie v.a. $U \sim Unif[1,3]$. Să se justifice de ce U^2 *nu* urmează distribuţia $Unif[1,9]$!
Indicaţie: Se pot folosi proprietăţi ale valorii medii!
R.: Fie $Y \sim Unif[1,9]$. Folosind calculele de la Problema 1-Seminar 6 şi faptul că $U \sim Unif[1,3]$, avem $E(U) = \frac{1+3}{2} = 2, E(Y) = \frac{1+9}{2} = 5$. Dar,

$$E(U^2) = \int_1^3 t^2 \frac{1}{3-1}dt = \frac{1}{2}\cdot \frac{t^3}{3}\Big|_1^3 = \frac{13}{3} \Longrightarrow E(U^2) \neq E(Y)$$

$\Longrightarrow U^2$ şi Y nu pot avea aceeaşi distribuţie.

6. Fie v.a. independente $U_1, U_2 \sim Unif[0,3]$. Să se justifice de ce $U_1 + U_2$ *nu* urmează distribuţia $Unif[0,6]$!
Indicaţie: Se pot folosi proprietăţi ale varianţei!
R.: Fie $Z \sim Unif[0,6]$. Folosind calculele de la Problema 1-Seminar 6 şi faptul că $U_1, U_2 \sim Unif[0,3]$, avem $V(U_1) = V(U_2) = \frac{9}{12} = \frac{3}{4}, V(Z) = \frac{36}{12} = 3$. Dar,

$$V(U_1+U_2) = V(U_1) + V(U_2)\ (U_1, U_2 \text{ sunt independente}) \Longrightarrow V(U_1+U_2) \neq V(Z)$$

$\Longrightarrow U_1 + U_2$ şi Z nu pot avea aceeaşi distribuţie.