Tema 4

Solutii

Exerciţiul 1 (Metoda Box-Muller)

Fie $R = \sqrt{-2\log(U_2)}$ şi $\Theta = 2\pi U_1$, atunci $(X_1, X_2) = g(R, \Theta)$ cu $g(r, \theta) = (r\cos(\theta), r\sin(\theta))$, $g : [0, \infty) \times [0, 2\pi] \to \mathbb{R}^2$. Cum U_1 şi U_2 sunt independente obţinem că şi R şi Θ sunt independente (ca funcţii de v.a. independente). Mai mult, cum $U_1 \sim \mathcal{U}([0,1])$ avem că $\Theta \sim \mathcal{U}([0, 2\pi])$ iar din $R = h(U_2)$ cu $h(u) = \sqrt{1 - 2\log(u)}$ rezultă

$$f_R(r) = f_{U_2}(h^{-1}(r)) \left|\frac{d}{dr}h^{-1}(r)\right| = |r|e^{-\frac{r^2}{2}}$$

Obţinem astfel că

$$\begin{aligned}
f_{(X_1,X_2)}(x_1, x_2) &= f_{(R,\Theta)}\left(g^{-1}(x_1, x2)\right)\left|\det J_{g^{-1}}\right| = f_{(R,\Theta)}\left(\sqrt{x_1^2 + x_2^2}, \arctan\frac{x_2}{x_1}\right)\frac{1}{\sqrt{x_1^2 + x_2^2}} \\
&= f_R\left(\sqrt{x_1^2 + x_2^2}\right) f_\Theta\left(\arctan\frac{x_2}{x_1}\right)\frac{1}{\sqrt{x_1^2 + x_2^2}} = \sqrt{x_1^2 + x_2^2}e^{-\frac{x_1^1 + x_2^2}{2}}\frac{1}{2\pi}\frac{1}{\sqrt{x_1^2 + x_2^2}} \\
&= \left(\frac{1}{\sqrt{2\pi}}e^{-\frac{x_1^2}{2}}\right)\left(\frac{1}{\sqrt{2\pi}}e^{-\frac{x_2^2}{2}}\right)
\end{aligned}$$

unde am folosit faptul că determinantul Jacobian-ului este $\det J_{g^{-1}} = \frac{1}{\sqrt{x_1^2 + x_2^2}}$. Astfel densitatea cuplului (X_1, X_2) se poate scrie ca un produs de densităţi care depind de x_1 şi respectiv x_2 ceea ce conduce la concluzia problemei (densităţile din factorizare sunt tocmai densităţile normalei standard).

Exerciţiul 2

Cum variabilele aleatoare $X_1, \ldots, X_n$ sunt independente avem că $\mathbb{E}[Y_n] = \mathbb{E}[X_1] \cdot \mathbb{E}[X_2] \cdots \mathbb{E}[X_n] = c^n$, $c \in (0, 1)$. Aplicand inegalitatea lui Markov obţinem, pentru $\varepsilon > 0$,

$$\mathbb{P}(|Y_n| > \varepsilon) \leq \frac{\mathbb{E}[|Y_n|]}{\varepsilon} = \frac{c^n}{\varepsilon} \overset{n\to\infty}{\to} 0,$$

unde am ţinut seama de faptul că v.a. sunt pozitive, deci $|Y_n| = Y_n$. Cum $\varepsilon > 0$ a fost ales arbitrar rezultă că $Y_n \overset{\mathbb{P}}{\to} 0$.

Exerciţiul 3

a) Pentru $M_n = \max(X_1, X_2, \ldots, X_n)$ observăm că pentru $x \in (0, 1)$

$$F_{M_n}(x) = \mathbb{P}\left(M_n \leq x\right) = \mathbb{P}(X_1 \leq x,\, X_2 \leq x, \ldots,\, X_n \leq x) \overset{indep.}{=} \prod_{i=1}^{n}\mathbb{P}(X_i \leq x) = F(x)^n = x^n.$$

Dacă $x < 0$ atunci $F_{M_n}(x) = 0$ iar dacă $x \geq 1$ avem $F_{M_n}(x) = 1$. In mod similar pentru $m_n = \min(X_1, X_2, \ldots, X_n)$ şi $x \in (0, 1)$ rezultă că

$$
\begin{aligned}
F_{m_n}(x) &= \mathbb{P}\left(m_n \leq x\right) = 1 - \mathbb{P}(m_n > x) = 1 - \mathbb{P}(X_1 > x,\, X_2 > x, \dots,\, X_n > x)\\
&\overset{indep.}{=} 1 - \prod_{i=1}^{n} \mathbb{P}(X_i > x) = 1 - [1 - F(x)]^n = 1 - (1-x)^n.
\end{aligned}
$$

Pentru a calcula densitatea v.a. m_n şi M_n este suficient să derivăm expresiile de mai sus şi obţinem $f_{m_n}(x) = n(1-x)^{n-1}$ şi $f_{M_n}(x) = nx^{n-1}$ pentru $x \in [0,1]$ şi 0 in rest.

b) Fie $Z_n = n(1 - M_n)$. Pentru calculul funcţiei de repartiţie avem

$$
F_{Z_n}(z) = \mathbb{P}(Z_n \leq z) = \mathbb{P}\left(M_n \geq 1 - \frac{z}{n}\right) = 1 - \left(1 - \frac{z}{n}\right)^n,\ z > 0.
$$

Cum $\left(1 - \frac{z}{n}\right)^n \to e^{-z}$ pentru $n \to \infty$ rezultă că

$$
\lim_{n\to\infty} F_{Z_n}(z) = 1 - e^{-z}
$$

această limită reprezentand funcţia de repartiţie a unei v.a. repartizată exponenţial de parametru 1.

Exerciţiul 4

Observăm că relaţia din cerinţa problemei se poate rescrie sub forma

$$
\begin{aligned}
\mathbb{P}\left(I_\alpha(X_1, X_2, \dots, X_n) \text{ conţine } \theta\right) &= \mathbb{P}\left(\bar{X} - z^\star n^{-\frac{1}{2}} \leq \theta \leq \bar{X} + z^\star n^{-\frac{1}{2}}\right)\\
&= \mathbb{P}\left(-z^\star \leq \sqrt{n}(\bar{X} - \theta) \leq z^\star\right)
\end{aligned}
$$

unde am notat $z^\star = z_{1-\frac{\alpha}{2}}$. Fie $Z = \sqrt{n}(\bar{X} - \theta)$; cum $\bar{X} \sim \mathcal{N}\left(\theta, \frac{1}{n}\right)$ obţinem că $Z \sim \mathcal{N}(0,1)$ deci

$$
\mathbb{P}\left(-z^\star \leq \sqrt{n}(\bar{X} - \theta) \leq z^\star\right) = \mathbb{P}\left(-z^\star \leq Z \leq z^\star\right) = \Phi(z^\star) - \Phi(-z^\star)
$$

unde $\Phi(x)$ este funcţia de repartiţie a normalei standard. Din definiţia lui $z^\star = z_{1-\frac{\alpha}{2}}$ deducem că $\Phi(z^\star) = 1 - \frac{\alpha}{2}$ şi prin simetrie $\Phi(-z^\star) = \Phi\left(z_{\frac{\alpha}{2}}\right) = \frac{\alpha}{2}$. Prin urmare

$$
\mathbb{P}\left(I_\alpha(X_1, X_2, \dots, X_n) \text{ conţine } \theta\right) = \Phi(z^\star) - \Phi(-z^\star) = 1 - \frac{\alpha}{2} - \frac{\alpha}{2} = 1 - \alpha.
$$

Exerciţiul 5

Pentru $\hat{\pi}_1$: observăm că v.a. X_i sunt v.a. de tip Bernoulli cu

$$
\begin{aligned}
\mathbb{P}(X_i = 1) &= \mathbb{P}\left(U_{i1}^2 + U_{i2}^2 < 1\right) = \iint_{\{u^2+u^2<1\}\cap[0,1]^2} f_{(U_{i1},U_{i2})}(u,v)\, dudv\\
&\overset{indep.}{=} \iint_{\{u^2+u^2<1\}\cap[0,1]^2} f_{U_{i1}}(u) f_{U_{i2}}(v)\, dudv = \int_0^1 \int_0^{\sqrt{1-u^2}} 1\, dvdu = \int_0^1 \sqrt{1-u^2}\, du\\
&\overset{u=\sin\alpha}{=} \int_0^{\frac{\pi}{2}} \cos^2\alpha\, d\alpha = \int_0^{\frac{\pi}{2}} \frac{\cos 2\alpha + 1}{2}\, d\alpha = \frac{\pi}{4} + \frac{1}{2}\int_0^{\frac{\pi}{2}} \cos 2\alpha\, d\alpha = \frac{\pi}{4}
\end{aligned}
$$

O altă variantă de calcul pentru $\mathbb{P}(X_i = 1)$ era să observam că această probabilitate se exprima şi ca raportul dintre aria mulţimii $\{(u,v) \in [0,1]^2\,|, u^2 + u^2 < 1\}$ şi cea a pătratului $[0,1]^2$, deci tot $\frac{\pi}{4}$.

Dacă $T = \sum_{i=1}^{n} X_i$ atunci $T \sim \mathcal{B}\left(n, \frac{\pi}{4}\right)$ de unde avem că media este $\mathbb{E}[T] = \frac{n\pi}{4}$ iar varianţa

$$\mathbb{V}[T] = n\frac{\pi}{4}\left(1 - \frac{\pi}{4}\right).$$

Cum $\hat{\pi}_1 = \frac{4}{n}T$ deducem că $\mathbb{V}[\hat{\pi}_1] = \frac{4\pi - \pi^2}{n}$. Din *Legea Numerelor Mari* obţinem că $\hat{\pi}_1 = \frac{4}{n}\sum_{i=1}^{n} X_i \overset{a.s.}{\to} 4\mathbb{E}[X_1] = 4\mathbb{P}(X_1 = 1) = \pi$.

Pentru $\hat{\pi}_2$, să observăm pentru inceput că media lui Y_1 este

$$\mathbb{E}[Y_1] = \int_0^1 \sqrt{1-u^2}\,du = \frac{\pi}{4}$$

iar varianţa lui Y_1 este

$$\mathbb{V}[Y_1] = \mathbb{E}[Y_1^2] - \mathbb{E}^2[Y_1] = \int_0^1 1 - u^2\,du - \frac{\pi^2}{16} = \frac{2}{3} - \frac{\pi^2}{16}.$$

Prin aplicarea *Legii Numerelor Mari* rezultă că

$$\hat{\pi}_2 = \frac{4}{n}\sum_{i=1}^{n} Y_i \overset{a.s.}{\to} 4\mathbb{E}[Y_1] = \pi$$

iar varianţa lui $\hat{\pi}_2$ este

$$\mathbb{V}[\hat{\pi}_2] = \frac{16}{n^2}\sum_{i=1}^{n}\mathbb{V}[Y_i] = \frac{16}{n}\left(\frac{2}{3} - \frac{\pi^2}{16}\right).$$

Pentru a vedea care dintre cei doi estimatori este mai eficient trebuie să verificăm care are varianţa mai mică. Cum $\frac{32}{3} < 12 < 4\pi$ rezultă că $\mathbb{V}[\hat{\pi}_2] < \mathbb{V}[\hat{\pi}_1]$ deci al doilea estimator este mai eficient.