Metody Probabilistyczne i Statystyka

LISTA 11

Informatyka algorytmiczna (I st.) WIT – 2023/2024

Zadanie 1. (*Normal approximation to binomial distribution*) Niech $X \sim \mathrm{Bin}(n, p)$.

(a) Korzystając z centralnego twierdzenia granicznego oszacuj prawdopodobieństwo tego, że X odchyla się od swojej wartości oczekiwanej o więcej niż $3\,\sigma_X$, gdzie $\sigma_X = \sqrt{\mathbf{var}\,(X)}$. Czy takie oszacowanie „ma sens" dla dowolnych wartości n?

(b) Porównaj to oszacowanie z ograniczeniem uzyskanym przy pomocy nierówności Czebyszewa oraz dokładną wartością szacowanego prawdopodobieństwa dla $n = 1000$ oraz $p = 1/2$.

HINT: Przeczytaj rozdział 4.3 z podręcznika M. Baron, *Probability and Statistics for Computer Scientists*, 2nd Edition, Chapman & Hall/CRC Press, 2013.

Zadanie 2. (*Statistical sampling*) W wyborach o fotel prezydenta ubiega się dwóch kandydatów – B i T. Chcemy oszacować frakcję p wyborców popierających T. W tym celu wybieramy losową próbę n wyborców (zakładamy, że każdy wyborca jest wybierany do sondażu niezależnie z jednakowym prawdopodobieństwem, nie odmawia odpowiedzi i mówi prawdę[1]) i jako estymator wartości p przyjmujemy frakcję $\hat{p}$ osób z wybranej próby, które popierają T.

(a) Wyznacz $\mathbf{E}\,(\hat{p})$ oraz $\mathbf{var}\,(\hat{p})$.

(b) Jak, w oparciu o centralne twierdzenie graniczne, dobrać wielkość próby n, żeby z prawdopodobieństwem $\geqslant 1 - \alpha$ pomylić się o $\leqslant \varepsilon$ dla zadanych α i ε? Oszacuj minimalne wartości n dla $\alpha = 0{,}05$ i $\varepsilon = 0{,}05$ oraz $\varepsilon = 0{,}01$.

(c) Zaproponuj alternatywne podejście oparte o nierówność Czebyszewa, jeśli metoda z podpunktu (b) zwróci zbyt małą wartość n, żeby aproksymacja rozkładu dwumianowego przez rozkład normalny była wiarygodna.

HINT: Przeczytaj rozdział 8.3 z podręcznika G. Grimmet, D. Welsh, *Probability. An introduction*, 2nd Edition, Oxford University Press, 2014 oraz rozdział 5.3.1 z podręcznika M. Baron, *Probability and Statistics for Computer Scientists*, 2nd Edition, Chapman & Hall/CRC Press, 2013 (warto zajrzeć też do rozdz. 8.1, 9.2 (wprowadzenie), 9.3.2 (s. 256 i początek 257) i 9.3.3)

Zadanie 3. Ustalmy $0 < \lambda < 1$. Dla $n \geqslant 1$ niech $X_n \sim \mathrm{Geo}\,(p_n)$, gdzie $p_n = \dfrac{\lambda}{n}$ i niech $Y_n = \dfrac{X_n}{n}$. Pokaż, że ciąg $(Y_n)_{n \geqslant 1}$ jest zbieżny według rozkładu i wyznacz rozkład graniczny.

Zadanie 4. Podaj przykład ciągu zmiennych losowych zbieżnego według rozkładu (słabo zbieżnego), ale nie według prawdopodobieństwa.

[1] W praktyce wybór reprezentatywnej próby np. w badaniach opinii publicznej jest dosyć złożonym zadaniem.

★ **Zadanie 5.** (*Zbieżność względem momentów*) Niech $X, X_1, X_2, \ldots : \Omega \to \mathbb{R}$ będą zmiennymi losowymi o skończonych p-tych momentach dla pewnego $p \geqslant 1$, tj. $\mathbf{E}\left(|X|^p\right) < \infty$ oraz $\mathbf{E}\left(|X_n|^p\right) < \infty$ dla każdego $n \geqslant 1$. Mówimy, że ciąg $(X_n)_{n\geqslant 1}$ jest **zbieżny względem p-tych momentów**[2] do X, co oznaczamy $X_n \xrightarrow{L^p} X$, jeżeli $\lim\limits_{n\to\infty} \mathbf{E}\left(|X_n - X|^p\right) = 0$. Można pokazać, że jeśli $X_n \xrightarrow{L^p} X$ dla pewnego $p \geqslant 1$, to w szczególności $\lim\limits_{n\to\infty} \mathbf{E}\left(X_n\right) = \mathbf{E}\left(X\right)$.

(a) Niech $X_n \sim U\left(\left[0, \frac{1}{n}\right]\right)$, $n \geqslant 1$. Pokaż, że $X_n \xrightarrow{L^p} 0$ dla dowolnego $p \geqslant 1$.

(b) Pokaż, korzystając z nierówności Markowa, że jeśli $X_n \xrightarrow{L^p} X$ dla pewnego $p \geqslant 1$, to $X_n \xrightarrow{P} X$.

(c) Podaj przykład ciągu zmiennych losowych zbieżnych według prawdopodobieństwa, ale nie względem wartości oczekiwanych, tj. ciągu $(X_n)_{n\geqslant 1}$ takiego, że $X_n \xrightarrow{P} X$ dla pewnej zmiennej losowej X, ale $\lim\limits_{n\to\infty} \mathbf{E}\left(X_n\right) \neq \mathbf{E}\left(X\right)$.

★ **Zadanie 6.** Korzystając z nierówności Cauchy'ego-Schwarza pokaż, że jeśli $X_n \xrightarrow{L^p} X$ dla $p = 2$, to $\lim\limits_{n\to\infty} \mathbf{E}\left(X_n\right) = \mathbf{E}\left(X\right)$.

★ **Zadanie 7.** Ustalmy $\lambda > 0$. Niech $X_n \sim \mathrm{Bin}(n, p_n)$, gdzie $p_n = \frac{\lambda}{n}$ i niech $X \sim \mathrm{Po}\left(\lambda\right)$. Pokaż, korzystając z funkcji charakterystycznych, że $X_n \xrightarrow{d} X$.

★ **Zadanie 8.** Ustalmy przestrzeń probabilistyczną $([0,1], \mathcal{B}([0,1]), \lambda)$, gdzie λ jest miarą Lebesgue'a. Niech $X_n(\omega) = n \cdot \mathbb{1}\{\omega \in [0, \frac{1}{n}]\}$ i niech $X(\omega) = 0$ dla $\omega \in [0,1]$. Pokaż, że $X_n \xrightarrow{a.s.} X$.

[2] W literaturze zbieżność ta nazywana jest również zbieżnością w przestrzeni L^p i jest definiowana dla dowolnych $p > 0$.