Statystyka
Lista 2

Histogramem rozkładu μ dla próbki $X_1, \ldots, X_n \sim_{i.i.d} \mu$ na przedziale $[a, b]$ względem podziału $a = c_0 < c_1 < \cdots < c_k = b$ nazywamy funkcję (losową) daną wzorem

$$\hat{f}_n(x) = \frac{1}{n(c_j - c_{j-1})} \sum_{i=1}^{n} \mathbb{I}(c_{j-1} < X_i \leq c_j) \quad \text{gdzie } c_{j-1} < x \leq c_j \ .$$

Zad 1. Udowodnić następujące stwierdzenie z rachunku prawdopodobieństwa: Jeśli $\mathbb{E}(X_n) \to a$ i $\mathrm{Var}(X_n) \to 0$, to $X_n \xrightarrow{P} a$ (a jest liczbą).

Zad 2. Niech $X_1, \ldots, X_n, \cdots \sim_{i.i.d} \mu$, gdzie μ posiada gęstość f skupioną na przedziale $[a, b]$. Rozpatrzmy histogram $\hat{f}_n$ z próbki $X_1, \ldots, X_n$ mający przedziały jednakowej długości h_n. Załóżmy, że $h_n \to 0$ i $nh_n \to \infty$ przy $n \to \infty$. Pokazać, że dla każdego punktu x, w którym f jest ciągła

$$\hat{f}_n(x) \xrightarrow{P} f(x).$$

Zad 3. Pokazać, że statystyka pozycyjna $X_{k:n}$ ma dystrybuantę

$$P(X_{k:n} \leq x) = \sum_{i=k}^{n} \binom{n}{i} F(x)^i (1 - F(x))^{n-i}$$

Zad 4. Pokazać, że jeśli badany rozkład ma gęstość $f(x)$ to k-ta statystyka pozycyjna $X_{k:n}$ ma gęstość daną wzorem

$$\frac{d}{dx} P(X_{k:n} \leq x) = n \binom{n-1}{k-1} f(x) F(x)^{k-1} (1 - F(x))^{n-k}$$

Zad 5. Obliczyć $\mathbb{E}(U_{k:n})$, gdzie $U_{k:n}$ oznacza statystykę pozycyjną z rozkładu jednostajnego $\mathcal{U}(0, 1)$.

Zad 6. Zbadać zbieność według rozkładu ciągu zmiennych losowych $n(1 - U_{n:n})$, gdzie $U_{n:n}$ oznacza ostatnią statystykę pozycyjną (maksimum z próbki) z rozkładu jednostajnego $\mathcal{U}(0, 1)$.

Zad 7. Pokazać, że wektor statystyk pozycyjnych $(U_{1:n}, \ldots, U_{n:n})$ z rozkładu jednostajnego $\mathcal{U}(0, 1)$ ma łączny rozkład jednostajny na sympleksie $\{u : u_1 \leq u_2 \leq \ldots \leq u_n \leq 1\}$.

Zad 8. Pokazać, że $X_{k:n} = F^{-1}(U_{k:n})$[1], gdzie $X_{k:n}$ jest k-tą statystyką pozycyjną z rozkładu o dystrybuancie F, a $U_{k:n}$ jest k-tą statystyką pozycyjną z rozkładu jednostajnego $\mathcal{U}(0, 1)$.

Zad 9. Wylosowano 10 liczb z rozkładu jednostajnego $\mathcal{U}(0, 1)$: 0,88; 0,39; 0,76; 0,13; 0,29; 0,14; 0,45; 0,63; 0,84; 0,38.

a) Na jednym rysunku wykonaj wykres dystrybuanty rozkładu $\mathcal{U}(0, 1)$ oraz dystrybuanty empirycznej dla podanego zdarzenia.

b) Oblicz prawdopodobieństwo, że dystrybuanta empiryczna rozkładu $\mathcal{U}(0, 1)$ dla dziesięcioelementowej próbki w punkcie $x = 1/2$ będzie miała wartość $0, 2$.

c) Oblicz prawdopodobieństwo, że w rozkładzie $\mathcal{U}(0, 1)$ ósma statystyka pozycyjna (z dziesięciu) przyjmie wartość większą niż $1/2$.

e) Oblicz średnią, wariancję, medianę i kwartyle z podanej próby i porównaj je z odpowiednimi statystykami rozkładu $\mathcal{U}(0, 1)$.

---

[1] $F^{-1}(y) := \min\{x : F(x) \geq y\}$.