

Lista 0

Zadanie 1

Niech X_1, \dots, X_n oznacza losową próbę prostą z rozkładu normalnego $N(\mu, \sigma^2)$:

1. Wygeneruj $n \in \{100, 1000\}$ zmiennych losowych z tego rozkładu dla $\mu = 0, \sigma^2 = 4$. Wykreśl ich trajektorie.
2. Porównaj dystrybuantę teoretyczną rozkładu normalnego z dystrybuantą empiryczną. Na osobnych wykresach przedstaw wynik dla $n = 100$ oraz $n = 1000$. Parametry średniej i wariancji możesz wyestymować za pomocą funkcji wbudowanych.
3. Porównaj gęstość teoretyczną rozkładu normalnego z histogramem z danych. Na osobnych wykresach przedstaw wynik dla $n = 100$ oraz $n = 1000$. Parametry średniej i wariancji możesz wyestymować za pomocą funkcji wbudowanych.
4. Porównaj empiryczną oraz teoretyczną funkcję charakterystyczną dla rozkładu normalnego.

$$\gamma_Y(t) = e^{iYt} = \exp(\mu it - \frac{\sigma^2 t^2}{2}), \quad t \in \mathbb{R}, \quad (1)$$

$$\hat{\gamma}_Y(t) = \frac{1}{n} \sum_{j=1}^n \exp(ix_j t). \quad (2)$$

Na potrzeby zadania weźmy $t \in [-5, 5], n = 100$. Aby wykonać zadanie należy:

- Wygenerować prostą próbę losową x_1, \dots, x_n ,
- Dla każdego $t \in \{-5, \dots, 5\}$ obliczyć wartości funkcji (1) i (2).

Zadanie 2

Niech Y będzie zmienną losową zdefiniowaną jako $Y = |X - EX|$, gdzie $X \sim N(\mu, \sigma^2)$.

1. Wyznacza dystrybuantę zmiennej losowej Y .
2. Na podstawie dystrybuanty policz gęstość.
3. Korzystając z metody momentów oraz metody największej wiarygodności wyznacz estymator parametru σ .
4. Przetestuj estymatory wykonując testy symulacyjne (np. na wykresach pudełkowych). Ustal, że liczba prób Monte Carlo MC=200. Zbadaj przypadki, gdzie długość próbki wynosi $n = 50, n = 100$ oraz $n = 1000$.

Zadanie 3

Na podstawie danych rzeczywistych z załączonego pliku wykonaj następujące zadania:

1. Oblicz wartości podstawowych statystyk opisujących dane.
2. Narysuj wykres rozproszenia i odpowiedz na pytanie, czy w danych widoczna jest zależność funkcyjna.
3. Oblicz współczynnik korelacji Pearsona i współczynnik korelacji Spermana. O czym świadczą uzyskane wartości?
4. Narysuj trajektorię x i y w zależności od czasu.

Zadanie 4

Niech X_1, \dots, X_n będą prostą próbą losową z rozkładu określonego za pomocą następującej funkcji gęstości:

$$f(x) = \frac{x^2}{2\Theta^3} e^{-x/\Theta}, \quad x > 0. \quad (3)$$

Wiadomo, że $\forall_{i=1,2,\dots,n} \quad E[X_i] = 3\Theta, \quad Var[X_i] = 3\Theta^2$.

1. Wyznacz estymator parametru Θ metodą momentów, wykorzystując informacje podane w treści zadania.
2. Wyznacz estymator parametru Θ metodą największej wiarygodności.
3. Wykaż, że otrzymane estymatory są estymatorami nieobciążonymi.

Zadanie 5

Ceny otwarcia dla dwóch podobnych akcji X_1 oraz X_2 są niezależnymi zmiennymi losowymi z tego samego rozkładu danego za pomocą gęstości:

$$f(x) = 0.5e^{-0.5(x-3)}, \quad x \geq 3. \quad (4)$$

Pewien inwestor planuje kupić tańszą akcję.

1. Wyznacz funkcję gęstości oraz dystrybuantę statystyki $Y = \min(X_1, X_2)$, która reprezentuje cenę, jaką zapłaci kupujący.
2. Wyznacz wartość oczekiwaną statystyki Y .

Zadanie 6*

Załóżmy, że chcemy ze sobą porównać dwa estymatory parametru średniej określone następującymi wzorami:

$$\Theta_1 = \frac{\sum_{i=1}^n X_i}{n}, \quad (5)$$

$$\Theta_2 = \text{mediana}(X), \quad (6)$$

Zbadaj jak zmienia się błąd średniokwadratowy obu estymatorów wraz ze wzrostem długości danych. W tym celu wykonaj 500 prób Monte Carlo. W każdej próbie wygeneruj n zmiennych losowych z rozkładu normalnego $N(\mu = 2, \sigma^2 = 4)$ oraz oblicz wartość estymatorów. Dla każdego estymatora wyznacz błąd średniokwadratowy. Powtórz eksperyment dla różnych wartości $n \in \{5, 10, 20, 50, 80, 100, 300\}$. Wyniki przedstaw na wykresie.