

## Statystyka

### Lista 2

#### Zadanie 1.

Wygeneruj  $n$  obserwacji z rozkładu dwumianowego  $b(5, p)$ .

- (a)  $n = 50, p = 0.1$ ,
- (b)  $n = 50, p = 0.3$ ,
- (c)  $n = 50, p = 0.5$ ,
- (d)  $n = 50, p = 0.7$ ,
- (e)  $n = 50, p = 0.9$ .

Na tej podstawie wyznacz wartość estymatora największej wiarygodności wielkości  $P(X \geq 3)$ , gdzie  $X \sim b(5, p)$ . Doświadczenie powtórz 10 000 razy. Oszacuj wariancję, błąd średniokwadratowy oraz obciążenie analizowanego estymatora. Przedyskutuj uzyskane wyniki w zależności od wyboru parametru  $p$ .

#### Zadanie 2.

Wygeneruj  $n$  obserwacji z rozkładu Poissona z parametrem  $\lambda$ .

- (a)  $n = 50, \lambda = 0.5$ ,
- (b)  $n = 50, \lambda = 1$ ,
- (c)  $n = 50, \lambda = 2$ ,
- (d)  $n = 50, \lambda = 5$ .

Na tej podstawie wyznacz wartość estymatora największej wiarygodności wielkości  $P(X = x)$ ,  $x = 0, 1, \dots, 10$ , gdzie  $X \sim \pi(\lambda)$ . Doświadczenie powtórz 10 000 razy. Oszacuj wariancję, błąd średniokwadratowy oraz obciążenie analizowanego estymatora. Przedyskutuj uzyskane wyniki w zależności od wyboru parametru  $\lambda$ .

#### Zadanie 3.

Liczby losowe czy pseudolosowe? Przedyskutuj wybór jednego z określeń na podstawie rozdziału 8.2.1, Koronacki i Mielniczuk (2009), str. 427-429.

#### Zadanie 4.

Wygeneruj  $n$  obserwacji z rozkładu beta z parametrami  $\theta$  i 1.

- (a)  $n = 50, \theta = 0.5$ ,
- (b)  $n = 50, \theta = 1$ ,
- (c)  $n = 50, \theta = 2$ ,
- (d)  $n = 50, \theta = 5$ .

Doświadczenie powtórz 10 000 razy. Na tej podstawie wyznacz wartość estymatora  $\widehat{I(\theta)}$  informacji Fishera parametru  $\theta$ . Wynik zapamiętaj.

Wygeneruj, niezależnie,  $n$  obserwacji z rozkładu beta z parametrami  $\theta$  i 1. Wyznacz wartość estymatora największej wiarygodności parametru  $\theta$ . Zdefiniuj nową zmienną  $Y = \sqrt{n\widehat{I(\theta)}}(\hat{\theta} - \theta)$ . Oblicz jej wartość na podstawie zaobserwowanej próby oraz zapamiętanego wcześniej wyniku.

Doświadczenie powtórz 10 000 razy. Narysuj histogram oraz wykres kwantylowo-kwantylowy. Przedyskutuj wybór liczby klas w histogramie oraz sposób wyznaczania kwantyli teoretycznych na wykresie kwantylowo-kwantylowym. Czy rozkład zmiennej losowej  $Y$  jest normalny? Odpowiedź uzasadnij.

### Zadanie 5.

Wygeneruj  $n$  obserwacji z rozkładu Laplace’a z parametrem przesunięcia  $\theta$  i skali  $\sigma$ .

(a)  $n = 50, \theta = 1, \sigma = 1,$

(b)  $n = 50, \theta = 4, \sigma = 1,$

(c)  $n = 50, \theta = 1, \sigma = 2.$

Na tej podstawie oblicz wartość estymatora parametru  $\theta$  postaci

(i)  $\hat{\theta}_1 = \bar{X} = (1/n) \sum_{i=1}^n X_i,$

(ii)  $\hat{\theta}_2 = Me\{X_1, \dots, X_n\},$

(iii)  $\hat{\theta}_3 = \sum_{i=1}^n w_i X_i, \sum_{i=1}^n w_i = 1, 0 \leq w_i \leq 1, i = 1, \dots, n,$  z własnym wyborem wag,

(iv)  $\hat{\theta}_4 = \sum_{i=1}^n w_i X_{i:n},$  gdzie  $X_{1:n} \leq \dots \leq X_{n:n}$  są uporządkowanymi obserwacjami  $X_1, \dots, X_n,$

$$w_i = \varphi\left(\Phi^{-1}\left(\frac{i-1}{n}\right)\right) - \varphi\left(\Phi^{-1}\left(\frac{i}{n}\right)\right),$$

przy czym  $\varphi$  jest gęstością, a  $\Phi$  dystrybuantą standardowego rozkładu normalnego  $N(0, 1)$ .

Doświadczenie powtórz 10 000 razy. Na tej podstawie oszacuj wariancję, błąd średniokwadratowy oraz obciążenie każdego z estymatorów. Przedyskutuj uzyskane wyniki. Który estymator jest optymalny i dlaczego? Skonfrontuj aktualne wyniki z rezultatami uzyskanymi w zadaniu 1 z listy 1.

### Zadanie 6.

Powtórz eksperyment numeryczny z zadań 1, 2, 4, 5 dla  $n = 20$  i  $n = 100$ . Przedyskutuj uzyskane rezultaty w nawiązaniu do wcześniejszych wyników.

### Literatura

Koronacki, J., Mielniczuk, J. (2009). *Statystyka dla studentów kierunków technicznych i przyrodniczych*. Wydawnictwa Naukowo-Techniczne, Warszawa.