2. ESTYMACJA PUNKTOWA

ZADANIE 2.1 Posługując się pakietem R, wyznaczyć estymatory największej wiarygodności parametrów a i d rozkładu Cauchy'ego o gęstości

$$f(x) = \frac{d}{\pi(d^2 + (x-a)^2)}, \quad \text{gdzie} \quad d > 0, \quad a \in \mathbb{R}, \quad x \in \mathbb{R},$$

dla następującej próby z tego rozkładu:

-1.11, 2.90, -1.10, -3.80, -1.66, -3.07, -3.25, -3.01,
-1.27, -5.86, -14.99, -10.23, -3.91, 99.34, -3.56, -5.39.

ZADANIE 2.2 (a) Niech $X_1, X_2, \ldots, X_n$ będzie prostą próbą losową z rozkładu wykładniczego $Exp(\lambda)$, czyli rozkładu o gęstości

$$f_\lambda(x) = \begin{cases} \lambda \exp(-\lambda x) & \text{dla } x > 0 \\ 0 & \text{dla } x \leqslant 0 \end{cases},$$

gdzie $\lambda > 0$. Wyprowadzić wzór na estymator największej wiarygodności parametru λ.

(b) W celu oszacowania czasu działania pewnych bateryjek, dział kontroli jakości zmierzył czas pracy 8 losowo wybranych bateryjek i otrzymał następujące wyniki (w godz.):

483, 705, 2623, 347, 620, 2719, 1035, 421.

Wiadomo, że czas pracy tych bateryjek ma rozkład wykładniczy $Exp(\lambda)$ z nieznaną $\lambda > 0$. Dla danych zebranych przez dział kontroli jakości, podać wartość estymatora największej wiarygodności parametru λ. Porównać wynik otrzymany przy użyciu wzoru z pkt. (a) z wynikiem otrzymanym przy użyciu funkcji `fitdistr()`.

(c) Dla danych z pkt. (b) wyznaczyć estymator największej wiarygodności dla

(i) średniego czasu działania bateryjki,

(ii) prawdopodobieństwa, że bateryjka będzie działać krócej niż 1000 godz.

WSKAZÓWKA: Jeśli $X \sim Exp(\lambda)$, to $EX = \frac{1}{\lambda}$. Ponadto $P(X < a) = \int_{-\infty}^{a} f(x)dx$. Przydaje się tu także tw. 2.1 z wykładu.

ZADANIE 2.3 Niech $X_1, X_2, \ldots, X_n$ będzie prostą próbą losową z rozkładu dwumianowego $bin(m, \theta)$ z $m = 10$ i z nieznanym parametrem $\theta \in (0, 1)$. Wyznaczyć estymator największej wiarygodności

(a) parametru θ,

(b) wartości oczekiwanej X_1,

(c) wariancji X_1.

WSKAZÓWKI: Rozkład dwumianowy $bin(m, \theta)$, $m \in \{0, 1, \ldots\}$, $\theta \in (0, 1)$, to rozkład dyskretny o funkcji masy prawdopodobieństwa

$$p_\theta(x) = \binom{m}{x} \theta^x (1-\theta)^{m-x} \text{ dla } x = 0, 1, \ldots, m.$$

Jeśli $X \sim bin(m, \theta)$, to $EX = m\theta$ i $Var(X) = m\theta(1-\theta)$.

ZADANIE 2.4 Niech $X_1, X_2, \dots, X_n$ będzie próbą losową z rozkładu ujemnego dwumianowego z parametrami $r = 2$ i $p \in (0,1)$, tzn. dla $i = 1, 2, \dots, n$,

$$P(X_i = x) = (x+1)p^2(1-p)^x, \quad x = 0, 1, 2, \dots.$$

Wyznaczyć estymator największej wiarygodności

(a) parametru p,

(b) prawdopodobieństwa $P(X_1 = 0)$.

ZADANIE 2.5 Niech $Gamma(a, \beta)$ oznacza rozkład gamma z parametrem kształtu a i drugim parametrem β, tzn. rozkład o gęstości

$$f_{a,\beta}(x) = \begin{cases} \frac{\beta^a}{\Gamma(a)} x^{a-1} \exp(-\beta x) & \text{dla } x > 0 \\ 0 & \text{dla } x \leqslant 0 \end{cases}, \quad a > 0, \beta > 0,$$

gdzie $\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx$. Wygenerować $n = 100$ obserwacji z rozkładu $Gamma(3, 2)$.

```
> rgamma(n=100,shape=3,rate=2)
```

Następnie przyjąć, że zapomnieliśmy wartości parametrów rozkładu gamma, z którego wygenerowaliśmy dane i, używając R, oszacować te parametry stosując metodę największej wiarygodności.

Powtórzyć zadanie z $n = 1000$.