ANALIZA PRZEŻYCIA

Lista 3

1. Na podstawie danych opisanych w zadaniu 3 z listy 2, przyjmując, że są one realizacjami zmiennych z rozkładu wykładniczego,

   (a) wyznaczyć oszacowania największej wiarogodności średniego czasu do remisji choroby dla pacjentów leczonych lekiem A oraz pacjentów leczonych lekiem B,

   (b) wyznaczyć realizację przedziału ufności, na poziomie ufności $1-\alpha$, dla średniego czasu do remisji choroby dla pacjentów leczonych lekiem A oraz pacjentów leczonych lekiem B, przyjmując $\alpha = 0.05$ i $\alpha = 0.01$.

2. Przyjmując w zadaniu 1, że obserwacje czasu do remisji choroby były prowadzone do momentu, w którym u dziesięciu pacjentów zostanie ona zaobserwowana, wyznaczyć oszacowania z punktu (a) i (b) zadania 1.

3. Niech $X_1, \ldots, X_n$ będą niezależnymi zmiennymi losowymi o jednakowym rozkładzie wykładniczym $\mathcal{E}(\vartheta)$, tzn. o funkcji gęstości

$$f_\vartheta(x) = \vartheta \exp\{-\vartheta x\} \mathbf{1}_{(0,\infty)}(x). \tag{1}$$

Niech t_0 będzie ustalonym czasem obserwacji. Wówczas, na podstawie danych cenzurowanych I-go typu, możemy szacować nieznaną wartość parametru ϑ przy użyciu następujących estymatorów

$$\hat{\vartheta} = \frac{R}{T_1}, \tag{2}$$

gdzie

$$R = \sum_{i=1}^{n} \mathbf{1}(X_i \leq t_0),$$

$$T_1 = \sum_{i=1}^{R} X_{(i)} + t_0(n-R),$$

oraz

$$\tilde{\vartheta} = -\frac{\log(1 - \frac{R}{n})}{t_0}. \tag{3}$$

W celu porównania dokładności estymatorów określonych wzorami (2) i (3), przeprowadzić symulacje, których wynikiem będzie tabela z oszacowaniami obciążeń $Bias(\hat{\vartheta}, \vartheta) = E_\vartheta(\hat{\vartheta} - \vartheta)$, $Bias(\tilde{\vartheta}, \vartheta) = E_\vartheta(\tilde{\vartheta} - \vartheta)$ i błędów średniokwadratowych $MSE(\hat{\vartheta}, \vartheta) = E_\vartheta(\hat{\vartheta} - \vartheta)^2$, $MSE(\tilde{\vartheta}, \vartheta) = E_\vartheta(\tilde{\vartheta} - \vartheta)^2$ obu estymatorów, dla wartości $\vartheta = 1$, $n = 10, 30$ i $t_0 = 0.5, 1, 2$.

Zadania dodatkowe

1. Udowodnić, że statystykę

$$T_2 = \sum_{i=1}^{m} X_{(i)} + (n-m)X_{(m)} \tag{4}$$

można przedstawić w postaci

$$T_2 = \sum_{i=1}^{m} D_i, \tag{5}$$

gdzie

$$D_i = (n-i+1)(X_{(i)} - X_{(i-1)}) \tag{6}$$

są tzw. unormowanymi spacjami i przyjmujemy $X_{(0)} = 0$. W przypadku obserwacji z rozkładu wykładniczego można pokazać, że D_i, $i = 1, \ldots, m$, określone wzorem (6), są niezależne i mają jednakowy rozkład wykładniczy z parametrem ϑ.

2. Załóżmy, że $X_1, \ldots, X_n$ są niezależnymi zmiennymi z rozkładu Rayleigha $\mathcal{R}a(\sigma)$ o gęstości

$$f_\sigma(x) = \frac{x}{\sigma^2} \exp\Big(-\frac{x^2}{2\sigma^2}\Big)\mathbf{1}_{(0,\infty)}(x),$$

gdzie $\sigma > 0$ jest nieznanym parametrem. Podać postać estymatora NW, przedziału ufności dla σ w oparciu o dane cenzurowane I-go typu.

3. Wyznaczyć dokładny rozkład $\hat{\mu} = 1/\hat{\vartheta}$, pod warunkiem, że $R > 0$ (zobacz Bartholomew, 1963).

Literatura

[1] Bartholomew, D. J. (1963). *The sampling distribution of an estimate arising on life-testing.* Technometrics, 5:361–374.

Alicja Jokiel-Rokita

23 września 2025