Zadania i problemy do wykładu *Statystyka*

(Zestaw nr 5)

Zadania

Zadanie 1. Dane są dwa estymatory parametru θ, S i T, o których wiadono, że $Var(S) = 40$ i $Var(T) = 3$.

1. Załóżmy, że $E(S) = \theta$ i $E(T) = \theta + 3$. Który z estymatorów wybierzesz i dlaczego?

2. Załóżmy, że $E(S) = \theta$ i $E(T) = \theta + a$ dla pewnej dodatniej liczby a. Który z estymatorów wybierzesz i dlaczego?

Zadanie 2. Niech $X_1, ..., X_n$ będzie próba losową z rozkładu wykładniczego $Exp(\lambda)$.

1. Uzasadnij, że $T_1 = \frac{1}{n}(X_1 + X_2 + ... + X_n)$ jest nieobciążonym estymatorem $\frac{1}{\lambda}$.

2. Niech M_n oznacza minimum z próby, $M_n = \min\{X_1, X_2, ..., X_n\}$. Wykaż, że statystyka $T_2 = nM_n$ jest również nieobciążonym estymatorem $\frac{1}{\lambda}$.

3. Który, z estymatorów T_1 i T_2 wybierzesz i dlaczego?

Zadanie 3. Przy przyjętych w zadaniu 2 założeniach i oznaczeniach rozpatrzy estymatory parametru $\frac{1}{\lambda}$, postaci

$$W_c = c\,(X_1 + X_2 + ... + X_n)\,, \text{ gdzie } c \text{ jest liczbą rzeczywistą.}$$

1. Policz błąd średniokwadratowy estymatora W_c.

2. Dla jakiego c, otrzymamy najlepszy, w sensie błędu średniokwadratowego, estymator?

3. Jak jest jego względna efektywność w porównaniu do nieobciążonego estymatora $\bar{X}_n$, (średniej próbkowej)?

Zadanie 4. Niech $X_1, ..., X_n$ będzie próba losową z rozkładu 0-1, dla którego $P\{X_i = 1\} = p$. Rozpatrujemy estymatory postaci

$$T_1 = \frac{1}{n}(X_1 + X_2 + ... + X_n) \quad \text{i} \quad T_n = \min\{X_1, X_2, ..., X_n\}\,.$$

1. Czy estymatory T_1 i T_2 są nieobciążonymi estymatorami parametru p?

2. Pokaż, że błędy średniokwadratowe estymatorów T_1 i T_2 są równe odpowiednio, $\frac{1}{n}p(1-p)$ oraz $p^n - 2p^{n+1} + p^2$.

3. Który, z estymatorów T_1 i T_2 jest bardziej efektywny, gdy $n = 2$?

Problemy

Problem 1. *(Zagadnienie regresji liniowej)*
Dla ustalonych liczb rzeczywistych x_1, x_2, ..., x_n niech

$$Y_i = \beta\, x_i + U_i,$$

gdzie U_1, U_2, ... U_n są niezależnymi zmiennymi losowymi, dla których $E(U_i) = 0$ i $Var(U_i) = \sigma^2$. Rozpatrujemy trzy estymatory parametru β, a mianowicie

$$T_1 = \frac{\sum_{i=1}^n x_i Y_i}{\sum_{i=1}^n x_i^2}, \quad T_2 = \frac{1}{n}\sum_{i=1}^n \frac{Y_i}{x_i}, \quad T_1 = \frac{\sum_{i=1}^n Y_i}{\sum_{i=1}^n x_i}.$$

1. Wykaż, że wszystkie trzy estymatory są nieobciążonymi estymatorami parametru β.
2. Policz błąd średniokwadratowy każdego z nich.
3. Wykaż, że estymator T_1 jest zawsze bardziej efektywny niż estymator T_3, oraz że estymator T_3 jest zawsze bardziej efektywny niż estymator T_2.

   Wskazówka. Skorzystaj odpowiednio z nierówności Cauchy'ego-Schwartza oraz nierówności Jensena

Problem 1. Czasami mamy do dyspozycji dwa nieobciążone estymatory U i T o tej samej wariancji $Var(U) = Var(T)$. W tej sytuacji nie ma powodów aby przedkładać jeden z nich nad drugi. Możemy jednak rozpatrzyć trzeci estymator W, postaci

$$W = \frac{U+T}{2}.$$

Łatwo widać, że W jest nieobciążonym estymatorem. W celu ocenienia estymatora powinniśmy policzyć jego wariancję. Jednak ponieważ nie dysponujemy informacją o łącznym rozkładzie U i T nie jest to możliwe.
Jednak, w każdym przypadku powinniśmy wybrać estymator W.

Wskazówka. Policz względną efektywność estymatora T, czyli $\frac{Var(W)}{Var(U)}$.