

Wybór optymalnej kombinacji liniowej dwóch statystyk porządkowych przy estymacji kwantyli

Luiza Pańczyk, Mariusz Bieniek

Instytut Matematyki UMCS w Lublinie

Niech $X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n}$ oznaczają statystyki porządkowe z próby rozmiaru n . Rozważamy problem estymacji kwantyli rozkładu populacji $F^{-1}(p)$ ustalonego rzędu $p \in (0, 1)$ przy użyciu kombinacji liniowych dwóch sąsiednich statystyk porządkowych postaci $(1 - \alpha)X_{j:n} + \alpha X_{j+1:n}$. Dla zadanego rozmiaru próby i rzędu kwantyla szukamy takich wartości j oraz α , dla których estymator kwantyla powyższej postaci jest optymalny w klasie wszystkich rozkładów ze skończoną wariancją. Intuicyjnie jasne jest, że j powinno być bliskie np , więc tradycyjnie najczęściej jako j wybiera się $[np]$ lub $[np] + 1$. Ponadto, zwykle α jest wybierane jako liniowa funkcja np .

Zastosujemy nowe kryterium optymalności oparte na osiągalnych oszacowaniach obciążenia $EX_{j:n} - F^{-1}(p)$. Pokażemy, że optymalny wybór dla α nie pokrywa się z tradycyjnym wyborem oraz porównamy uzyskany estymator z klasycznymi estymatorami kwantyli (zob. Hyndman, Fan (1996)).

Literatura

- [1] R. J. Hyndman, Y. Fan (1996), *Sample Quantiles in Statistical Packages*, The American Statistician 50, 361-365