

Metody Probabilistyczne i Statystyka

LISTA 13 do wykładu 14 (statystyka)

Informatyka algorytmiczna (I st.) WliT – 2022/2023

Zadanie 1. Pragniesz zorientować się ile mniej więcej znasz słów w języku angielskim. Jak to sprawdzić? Zaproponuj konkretną i łatwą do realizacji procedurę, która nadawałaby się do zautomatyzowania.

Zakładamy, że masz pod ręką słownik angielsko-polski w wersji cyfrowej. Jest to baza danych, w której rekord składa się z następujących pól: numer kolejny, słowo angielskie, tłumaczenie na polski.

Zadanie 2. Estymator \hat{M} mediany dla wartości funkcji f w populacji P obliczamy w następujący sposób:

- wybieramy losowo (i niezależnie) $u_1, \dots, u_n \in P$,
- obliczamy wartości $f(u_1), \dots, f(u_n)$ i znajdujemy ich medianę m ,
- podstawiamy $\hat{M} := m$

pytanie 1: czy estymator \hat{M} jest nieobciążony?

pytanie 2: jeśli jest nieobciążony, to czy \hat{M} jest zgodny (consistent)?

Zadanie 3. Jakie są 1, 10, 25, 50, 75, 99 kwantyle dla rozkładu normalnego o wartości oczekiwanej oraz wariancji równych 1?

Zadanie 4. Wykonaj serię 5 eksperymentów dla zmiennej losowej X z rozkładem eksponencjalnym¹ z parametrem $\lambda = 1.1212$ dla dwóch sytuacji:

a) próbka składa się z wartości X_1, \dots, X_5

b) próbka składa się z wartości X_1, \dots, X_{50}

W każdym wykonaniu oblicz:

1. estymator wartości oczekiwanej X
2. estymator dla wariancji X
3. dla porównania sprawdź jak się zachowuje estymator wariancji, gdzie w mianowniku zamiast $n - 1$ (w naszym przypadku 4 oraz 49) wstawiamy n

Jakie wnioski mógłaby wyciągnąć osoba nie znająca λ na podstawie tych wyników?

Przedstaw pisemny raport z wyników. Wnioski do dyskusji na ćwiczeniach.

Zadanie 5. Zmienna losowa X przyjmuje 3 wartości: 0 z prawdopodobieństwem p , 1 z prawdopodobieństwem q oraz 2 z prawdopodobieństwem $1 - p - q$. Parametry p oraz q są nieznanne. Przedstaw procedure wyliczenia estymatorów dla p i q opartą na metodzie momentów.

O ile zadanie to rozwiązujesz po 2 wykładzie za statystyki, to pokaż jak wyliczyć estymatory p oraz q za pomocą metody *maximum likelihood*.

Dla sprawdzenia, zrób serię eksperymentów.

¹Uwaga: nie musisz – a nawet nie powinieneś – ręcznie generować X , możesz użyć np. MATLABa i funkcji `exprnd(λ)`