

POSUDOK OPONENTA BAKALÁRSKEJ PRÁCE

Názov: James-Steinův odhad
Autor: Vojtěch Novotný

ZHRNUTIE OBSAHU PRÁCE

Predložená bakalárska práca predstavuje klasický James-Steinov (J-S) odhad parametru polohy viacrozmerného normálneho rozdelenia. V kapitole 1 je predstavený problém a definovaný pojem prípustnosti odhadu. Kapitola 2 interpretuje J-S odhad ako empirický bayesovský odhad, a uvádza výraz pre jeho riziko. (Bayesovské) riziko J-S a ďalších podobných odhadov je odvodené v kapitole 3; ukazuje sa slávny výsledok, že v prípade jediného pozorovania J-S odhad dominuje najlepší nestranný odhad v dimenzii $k > 2$. Záverečná kapitola 4 obsahuje krátku simulačnú štúdiu.

CELKOVÉ HODNOTENIE PRÁCE

Téma práce. Téma práce je veľmi zaujímavá a vhodná.

Vlastný príspevok. Autor pomerne prehľadne spracoval niekoľko zdrojov z odbornej literatúry. Je prezentovaná drobná simulačná štúdia v kapitole 4 práce.

Matematická úroveň. Práca obsahuje niekoľko definícií a dokázaných tvrdení, najmä v hlavnej kapitole 3. Matematické vyjadrovanie sa zdá byť zväčša korektné, niektoré odvodenia by si ale zaslúžili úplnejší a detailnejší komentár.

Práca so zdrojmi. Matematické vety prezentované v kapitole 3 sú prevzaté z článku Efron a Morris (1973). Tento článok je riadne citovaný. Najmä dôkaz hlavnej vety 4 s odvodením rizika J-S odhadu sa však zdá byť doslovne prebratý z literatúry, a postup by si zaslúžil lepšie vysvetlenie. Odkazy do literatúry by mohli byť presnejšie, tj mali by vždy smerovať priamo na konkrétne tvrdenia z článkov.

Formálna úprava. Formálna úprava práce je priemerná. Text obsahuje isté množstvo formálnych chýb a problémov v značení, celkovo je ale text dobre čitateľný a zrozumiteľný.

ŠPECIFICKÉ PRIPOMIENKY

1. Vedel by autor uviesť nejaký intuitívne jasný argument vysvetľujúci James-Steinov paradox?
2. V sekcii 2.2 je argumentované, že štatistika T je postačujúca pre α . Nie je mi však jasné v akom modele, a podľa čoho vidíme, že táto štatistika je postačujúca. Rovnako mi nie je jasné, odkiaľ sa berie odhad $\hat{\alpha}(T) = (k - 2)/T$. Prečo by malo ísť o rozumný odhad parametru α ?
3. Aký je rozdiel medzi $\hat{A}(T)$ a $\hat{\alpha}(T)$ v znení lemy 2?
4. Je možné, že RSL definované na str. 11 môže byť aj záporné?
5. Odkiaľ sme získali výrazy pre negatívne momenty χ^2 -rozdelenia použité v dôkaze vety 3?
6. Ako presne sme odvodili nerovnosť v dôkaze dôsledku 1?
7. Mohol by autor vysvetliť, prečo je možné obe strany rovnice (3.5) napísať ako funkciu normy vektoru θ_X , ako je tvrdené v dôkaze vety 4? Čo znamená, že „rozdelenia $\|\theta_X\|^2$ sú úplne ako funkcie τ^2 “, a ako z toho plynie, že f a g sú identické?

8. V kapitolách 1–3 boli odhady uvažované iba pre prípad jediného pozorovania, tj $n = 1$. Čo teda znamená napr. J-S odhad použitý na rozsah $n > 1$ v simuláciách v kapitole 4? Je možné interpretovať, prečo sa chyby odhadov J-S a B v tabuľke 4.1 typicky zvyšujú s rastúcim n ?
9. Nesúhlasím s tvrdením zo záveru práce, kde je uvedené, že J-S odhad prezentuje „reálnu alternatívu“ k maximálne vierohodnému odhadu strednej hodnoty normálneho rozdelenia. Konštrukcia J-S odhadu poukazuje na limity štandardnej teórie odhadu. Nemyslím si ale, že J-S odhad je vhodný pre použitie v štatistickej praxi.

NIEKTORÉ FORMÁLNE CHYBY

1. Píše sa „Bayesova veta“, ale „bayesovský odhad“.
2. Odkazy na čísla kapitol a sekcií by nemali byť v zátvorkách, tj píše sa „v kapitole 1.5,“ a nie „v kapitole (1.5).“
3. Str. 3: Čo je to „náhodné rozdelení“ na str. 3?
4. Str. 3: „odhadování z jednoho parametru“ má byť „odhadování z jednoho pozorování.“
5. Str. 4: V definícii 2 by malo byť vysvetlené, čo znamená E_{θ} .
6. Vzorec (3.1) a inde: Pri niekoľkoriadkových vzorcoch sa riadky nezakončujú výrazmi =, < apod. Tieto sa uvádzajú iba na začiatku ďalšieho riadku.
7. Rozsahy ako „ $n=1$ “ v tabuľke 4.1 by mali byť „ $n = 1$ “ apod.

ZÁVER

Jedná sa o mierne podpriemernú až priemernú prácu, ktorú však je možné uznať ako bakalársku prácu na MFF UK.

Stanislav Nagy
KPMS MFF UK
28. júla 2023