Posudek diplomové práce:

Blanka Hamplová: Testování normality v lineárním modelu

Předložená práce se zabývá testováním normality v lineárním modelu. První kapitola obsahuje krátký a přehledný souhrn vlastností lineárního modelu a definice čtyř druhů reziduí. Druhá kapitola obsahuje stručný popis deseti často používaných testů normality. Ve třetí kapitole jsou popsána rozdělení náhodných chyb, použitá v simulační studii, která ve čtvrté kapitole zkoumá hladinu a sílu testů normality v závislosti na typu testu, typu lineárního modelu a zejména na typu reziduí. Závěry simulační studie jsou stručně a přehledně shrnuty v závěrečné páté kapitole.

Práce se zabývá zajímavým praktickým problémem. Potřebná teorie i celá simulační studie jsou vysvětleny velice přehledně a pečlivě a k celé práci mám pouze několik drobných připomínek:

str. 7, řádek 8: Je počet prvků tvořících jeden řádek regresní matice k nebo $k+1$? Předpokládá se v celé práci, že lineární model obsahuje absolutní člen?

kapitola 2: V seznamu testů normality postrádám oblíbený „Jarque-Bera“ test.

str. 17, první vzorec: Má být ve jmenovateli rozptyl nebo odmocnina rozptylu?

str. 24, poslední vzorec: Symbol μ zde není definován.

str. 26, sekce 3.6: Proč zde nejsou parametry směsi zvoleny tak, aby výsledné rozdělení mělo rozptyl 1?

str. 27, třetí odstavec: Opravdu by nebylo možné vytvářet matici modelu jednoduchého třídění postupně? Nešlo by poskládat matici modelu například z pětic složených vždy ze dvou jedniček následovaných třemi nulami tak, že pro uvažované rozsahy výběrů by byl poměr jedniček a nul vždy stejný?

Na druhou stranu, možná by bylo vždy lepší vytvářet výběry nezávisle pro každý rozsah výběru: pokud se při postupu použitém v předložené práci náhodou vygeneruje jediné „špatné“ pozorování, bude automaticky ovlivňovat i všechny vyšší rozsahy výběru a příslušný test by pak mohl kvůli jednomu nepravděpodobnému pozorování vypadat horší, než ve skutečnosti je (jinými slovy: výsledky simulací pro různé hodnoty n mohou být korelované a pokud test „náhodou“ překročí kritickou hodnotu odvozenou na straně 29 pro n_1, tak ji pak se zvýšenou pravděpodobností překročí i pro $n_2 > n_1$).

str. 33, obrázek 4.3: Proč síla Cramér-von Misesova testu klesá s rostoucím počtem pozorování?

Shrnutí: Uchazečka s přehledem řeší zajímavý praktický problém. Předložená práce splňuje všechny požadavky kladené na diplomovou práci a doporučuji ji uznat jako diplomovou.

RNDr. Zdeněk Hlávka, Ph.D.