

ZÁKLADY EKONOMETRIE (KOMBINOVANÁ FORMA) – SKUPINOVÝ ÚKOL 1

Pro účely zpracování úkolu si vytvořte pracovní skupinu o max. 5 lidech. Úkoly zpracujte a odevzdejte do stanoveného termínu. Veškeré podklady (např. relace gretlu) a komentáře, které mohou být i součástí relace gretlu, odevzdejte do příslušné odevzdárny.

Termín odevzdání: 16. 11. 2015 (včetně)

Zadání úkolu

1. (Hill et. al. (2008): př. 5.8) Soubor `cocaine.gdt` na záložce PoE v `gretlu` obsahuje 56 pozorování proměnných vztahujících se k prodeji kokainu v severovýchodní Kalifornii v období let 1984–1991. Data jsou podmnožinou dat použitých ve studii Culkins, J.P. a Padman, R. (1993): „Quantity Discounts and Quality Premia for Illicit Drugs,“ *Journal of the American Statistical Association*, 88, 748-757. Proměnné jsou

- *price* = cena za gram kokainu v rámci dané transakce;
- *quant* = počet gramů kokainu prodaných v dané transakci;
- *qual* = kvalita kokainu vyjádřená jako procento čistoty;
- *trend* = časová proměnná s hodnotami od 1984=1 až po 1991=8.

Předpokládejme regresní model

$$price = \beta_0 + \beta_1 quant + \beta_2 qual + \beta_3 trend + \epsilon.$$

- (a) Jaká znaménka koeficientů byste očekávali u parametrů β_1 , β_2 a β_3 ?
 - (b) Odhadněte daný model a interpretujte dosažené hodnoty odhadů koeficientů. Jsou znaménka parametrů v souladu s vaším očekáváním?
 - (c) Jaký podíl variability v ceně kokainu je vysvětlen společnou variabilitou prodaného množství kokainu, jeho kvalitou a časem?
 - (d) Říká se, že čím větší objem těchto nelegálních obchodů člověk provádí, tím větší riziko, že jej dostihne ruka zákona. Prodejci tak jsou ochotni akceptovat nižší cenu, pokud prodávají větší množství dané drogy. Pokuste se testovat tuto hypotézu. Stanovte tedy odpovídající nulovou a alternativní hypotézu a proveďte odpovídající test hypotézy.
 - (e) Ověřte hypotézu, že kvalita kokainu nemá vliv na jeho cenu, oproti alternativní hypotéze, že u kvalitního kokainu je placena odpovídající prémie.
 - (f) Jaká je průměrná (očekávaná) roční změna ceny kokainu? Zamyslete se nad tím, proč by se mohla cena takto měnit.
 - (g) Jakou modifikaci modelu byste museli provést, abyste byli schopni testovat hypotézu, že vliv kvality kokainu na prodejní cenu je tím větší, čím větší je množství prodaného kokainu? Model modifikujte a danou hypotézu ověřte.
 - (h) Dokázali byste model upravit tak, aby dokázal zachytit odlišné vlivy objemu transakce cenu do 50 dolarů za gram kokainu a pro cenu nad 50 dolarů za gram? Pokuste se o tuto modelovou modifikaci a tuto možnou heterogenitu vlivu objemu transakce pro takto zvolené dvě cenové kategorie kokainu otestujte.
2. (Hill et. al. (2008): př. 7.12) Lion Forest je velmi úspěšný profesionální hráč golfu. Ve věku 45 let však jeho hra přestala být tou, kterou bývala dříve. Svou profesionální kariéru začal, když mu bylo 20 a po dovršení svých 45. narozenin ho náhle začala zajímat analýza historického průběhu jeho výsledků v závislosti na tom, jak postupně stárnul. Soubor `golf.m` na záložce PoE v `gretlu` obsahuje údaje o jeho konečném skóre (skutečné skóre mínus tzv. „par“) ze 150 turnajů spolu s údaji o věku (v jednotkách odpovídajících deseti letům). V souboru jsou obsaženy výsledky z šesti hlavních turnajů (pro každý rok) v průběhu posledních 25 let. Označíme-li si výsledek

turnaje jako $SCORE$ a jeho věk jako AGE , odhadněte následující model a pokuste se i o modelovou predikci jeho výsledků v rámci vzorku:

$$SCORE = \beta_1 + \beta_2 AGE + \beta_3 AGE^2 + \beta_4 AGE^3 + \epsilon.$$

- (a) Ověřte (testujte), který z modelů je vhodnější, zda-li kvadratický nebo kubický.
- (b) Použijte modelové predikce (v rámci vzorku, tzn. pro věk od 20 do 45 let) pro zodpovězení následujících otázek:
- V jakém věku byl Lion na vrcholu své kariéry?
 - V jakém období jeho věku docházelo ke zlepšování jeho hry, a to rostoucím tempem růstu?
 - V jakém období docházelo ke zlepšování Lionovy hry, a to snižujícím se tempem růstu?
 - Ve kterém věku začal hrát Lion hůře než na začátku své kariéry (což bylo ve věku 20 let)?
 - Od kterého věku Lion už nebyl (či nebude) schopen hrát pod par? (v průměru)
- (c) Když bude Lionovi 70, bude (podle našeho modelu) schopen zahrát turnaj na 100 úderů? Předpokládáme, že par turnaje je 72.
3. (Hill et al. (2011): př. 5.16) Datový soubor `tuna.gdt` dostupný na záložce PoE 4th ed. v `gretlu` obsahuje data o týdenních prodeích jedné ze značek tuňáka v konzervě prodávaných v jednom řetězci supermarketů ve velkém městě Spojených států na středozápadě. Data jsou z roku z poloviny 90. let (pokud instalace `gretlu` datové soubory z nového vydání Hill et al. (2011) neobsahuje, je datový soubor součástí podkladových materiálů ke cvičení). K dispozici je 52 pozorování proměnných:

- $SAL1$, počet prodaných jednotek tuňáka v konzervě značky č. 1;
- $APR1$, cena za konzervu tuňáka značky č. 1;
- $APR2$, $APR3$, ceny za konzervu tuňáka značek č. 2 resp č. 3.

- (a) Ceny $APR1$, $APR2$ a $APR3$ jsou vyjádřeny v dolarech. Vynásobte každé z pozorování těchto proměnných stovkou, pro vyjádření cen v centech. Označte tyto nové proměnné $PR1$, $PR2$ a $PR3$. Odhadněte následující regresní model a komentujte dosažené výsledky:

$$SAL1 = \beta_1 + \beta_2 PR1 + \beta_3 PR2 + \beta_4 PR3 + \epsilon.$$

- (b) Interpretujte odhady b_2 , b_3 a b_4 . Jsou jejich znaménka v souladu s očekáváním?
- (c) S využitím vhodného jednostranného testu a 5% hladiny významnosti testujte, zda-li je každý jednotlivý koeficient b_2 , b_3 a b_4 statisticky významný od nuly.
- (d) S využitím 5% hladiny významnosti testujte následující hypotézy:
- Zvýšení ceny tuňáka značky č. 1 o jeden cent sníží jeho prodej o 300 konzerv.
 - Zvýšení ceny tuňáka značky č. 2 o jeden cent zvýší prodej tuňáka značky č. 1 o 300 konzerv.
 - Zvýšení ceny tuňáka značky č. 3 o jeden cent zvýší prodej tuňáka značky č. 1 o 300 konzerv.
 - Vliv zvýšení ceny tuňáka značky č. 2 na prodej značky č. 1 je stejný jako vliv zvýšení ceny tuňáka značky č. 3 na prodejnost značky č. 1. Jsou výstupy tohoto testu v rozporu se závěry testů předchozích dvou podotázek?
 - Pokud se ceny všech tří značek zvýší o 1 cent, nedojde ke změně v prodaných kusů tuňáka značky č. 1.
- (e) Nalezněte 95% intervalový odhad očekávaného prodaného množství tuňáka č. 1 pokud $PR1 = 90$, $PR2 = 75$ a $PR3 = 75$. Je něco na získaném intervalu špatně?

- (f) Odhadněte alternativní model

$$\ln(SAL1) = \alpha_1 + \alpha_2 PR1 + \alpha_3 PR2 + \alpha_4 PR3 + \epsilon.$$

Nalezněte 95% intervalový odhad očekávaného logaritmu prodejů, pokud $PR1 = 90$, $PR2 = 75$ a $PR3 = 75$. Převeďte tento interval do očekávaného počtu prodaných kusů a porovnejte výsledky s výsledky předchozí otázky.

- (g) Jak se změnila interpretace koeficientů v modelu s $\ln(SAL1)$ oproti modelu s vysvětlovanou proměnnou $SAL1$?