Příklad 2

Phillipsova křivka, tak jak ji popsali Samuelson a Sollow, se vyznačuje nepřímo negativním vztahem mezi mírou nezaměstnanosti (U) a mírou inflace (Pi). Na datech odhadněte koeficienty lineárního regresního modelu tvaru: $Pi = \beta_0 + \beta_1 \cdot U + \epsilon$, pokud máte k dispozici měsíční data za poslední tři roky. Data jsou uložena v souboru SC_LINREG_P2.

1. Jaké očekáváte znaménko u parametru β_1? Zamyslete se nad kontextem teorie dlouhého a krátkého období.
2. Odhadněte krátkodobé Phillipsovy křivky za každý rok zvlášť. Dá se z tohoto příkladu usuzovat, že teorie Phillipsovy křivky jsou platné pro každé z období? Je znaménko u parametru β_1 v souladu s vašim očekáváním?
3. Odhadněte dlouhodobou Phillipsovu křivku. Je znaménko u parametru β_1 v souladu s vašim očekáváním?
4. Otázky na zamyšlení:
   - Kolik řádků a sloupců mají regresní matice jednotlivých modelů?
   - Kolik lze provést maximálně t-testů?
   - Jak zní nulová hypotéza pro celkový F-test?

Příklad 3

Empirickým pozorováním bylo zjištěno, že byty, které mají větší započitatelnou plochu ZP v metrech čtverečních bývají v přepočtu na k/m^2 levnější. Tuto nelinearitu ověřte pomocí lineárního regresního modelu tvaru $Cena = \beta_0 + \beta_1 \cdot ZP + \beta_1 \cdot ZP^2 + \epsilon$. Data jsou k dispozici v souboru SC_LINREG_P3.

1. Jaké očekáváte znaménka u parametrů β_1 a β_2? Rovnici odhadněte. Jsou výsledky v souladu s Vaším očekáváním?
2. Proveďte celkový F-test a dílčí t-testy. Na základě těchto testů rozhodněte, zda je model vhodný a slovně interpretujte.
3. Otázky na zamyšlení:
   - Kolik řádků a sloupců má regresní matice modelu ze zadání?
   - Kolik lze provést maximálně t-testů?
   - Jak zní nulová hypotéza pro celkový F-test?
   - Co o modelu říká koeficient determinace?

Příklad 4

V souboru *SC_LINREG_P4* jsou uložena data tzv. Anscombeho kvartetu. Spočtěte základní popisné statistiky a odhadněte lineární regresní modely tvaru $Y_i = \beta_0 + \beta_1 \cdot X_i + \epsilon$. Výsledky interpretujte graficky pomocí scatter plotů. Zamyslete se nad „vhodností“ využití lineárního regresního modelu a nad tím, na co si musí dát výzkumník při zpracování dat pozor.

Řešení

Př. 2 a) V krátkém období by β_1 měla nabývat záporných hodnot, v dlouhém období by měla být nulová.

b) Znaménka jsou v souladu s očekáváním. Statistická nevýznamnost β_1 potvrzena pouze v roce 1. Dá se říci, že krátkodobá teorie PC je na datech pozorovatelná.

Výsledky regrese se závislou proměnnou : Inflace (SC_LINREG_P2)
R= ,30995110 R2= ,09606968 Upravené R2= ,00567665
F(1,10)=1,0628 p<,32687 Směrod. chyba odhadu : ,27827
Zhrnout podmínku: Rok=1

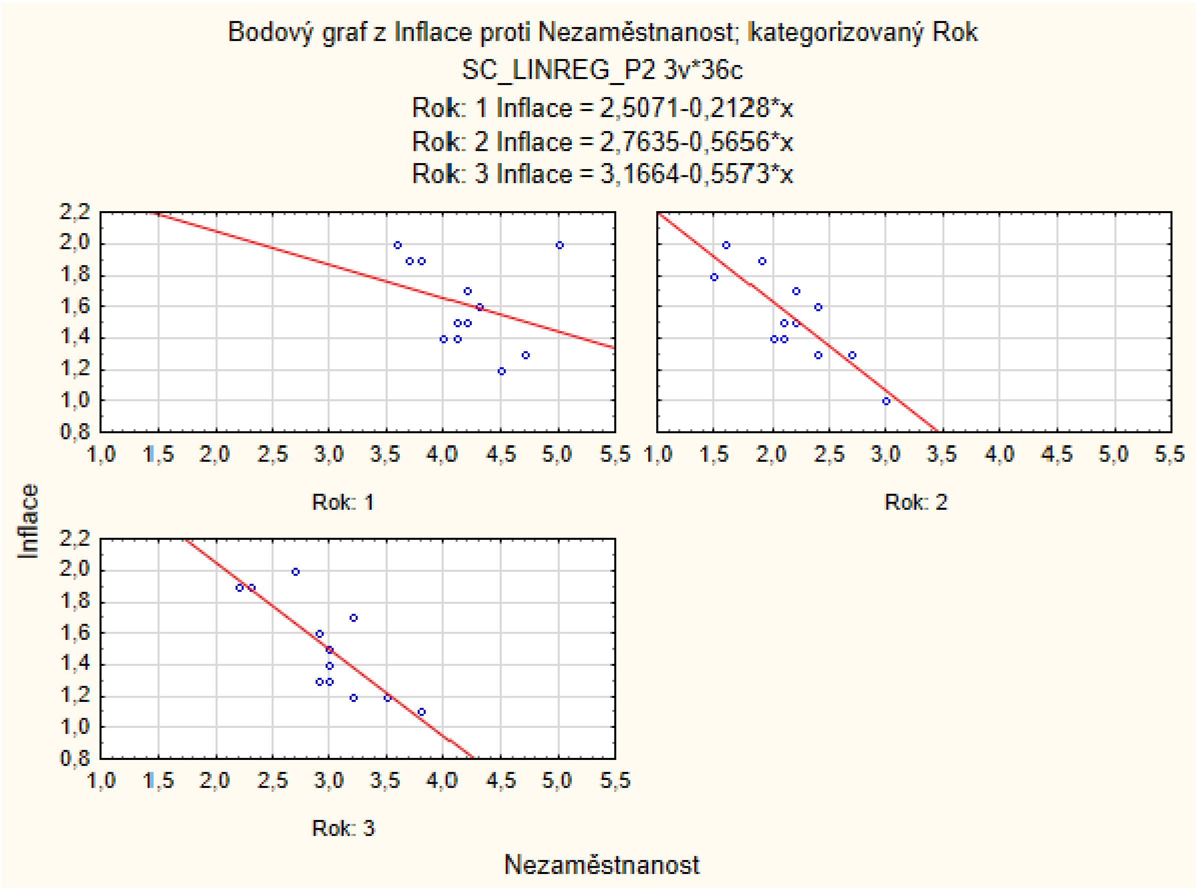
N=12	b*	Sm.chyba z b*	b	Sm.chyba z b	t(10)	p-hodn.
Abs.člen			2,507064	0,867418	2,89026	0,016100
Nezaměstnanost	-0,309951	0,300654	-0,212844	0,206460	-1,03092	0,326867

Výsledky regrese se závislou proměnnou : Inflace (SC_LINREG_P2)
R= ,84146868 R2= ,70806954 Upravené R2= ,67887649
F(1,10)=24,255 p<,00060 Směrod. chyba odhadu : ,16089
Zhrnout podmínku: Rok=2

N=12	b*	Sm.chyba z b*	b	Sm.chyba z b	t(10)	p-hodn.
Abs.člen			2,763524	0,254071	10,87699	0,000001
Nezaměstnanost	-0,841469	0,170860	-0,565605	0,114846	-4,92491	0,000600

Výsledky regrese se závislou proměnnou : Inflace (SC_LINREG_P2)
R= ,80749771 R2= ,65205255 Upravené R2= ,61725781
F(1,10)=18,740 p<,00149 Směrod. chyba odhadu : ,19106
Zhrnout podmínku: Rok=3

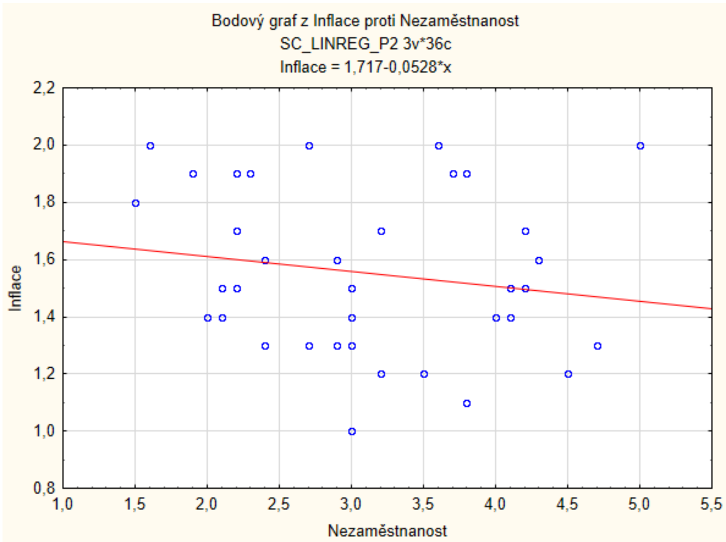
N=12	b*	Sm.chyba z b*	b	Sm.chyba z b	t(10)	p-hodn.
Abs.člen			3,166364	0,386959	8,18268	0,000010
Nezaměstnanost	-0,807498	0,186533	-0,557321	0,128742	-4,32897	0,001492



c) Znaménko je záporné, nicméně parametr β_1 je statisticky nevýznamný.Z tohoto důvodu lze říci, že v dlouhém období neexistuje vztah mezi mírou nezaměstnanosti a inflací.

Výsledky regrese se závislou proměnnou : Inflace (SC_LINREG_P2)
R= ,17216502 R2= ,02964079 Upravené R2= ,00110082
F(1,34)=1,0386 p<,31535 Směrod. chyba odhadu : ,28618

N=36	b*	Sm.chyba z b*	b	Sm.chyba z b	t(34)	p-hodn.
Abs.člen			1,717033	0,168086	10,21521	0,000000
Nezaměstnanost	-0,172165	0,168938	-0,052796	0,051807	-1,01910	0,315351



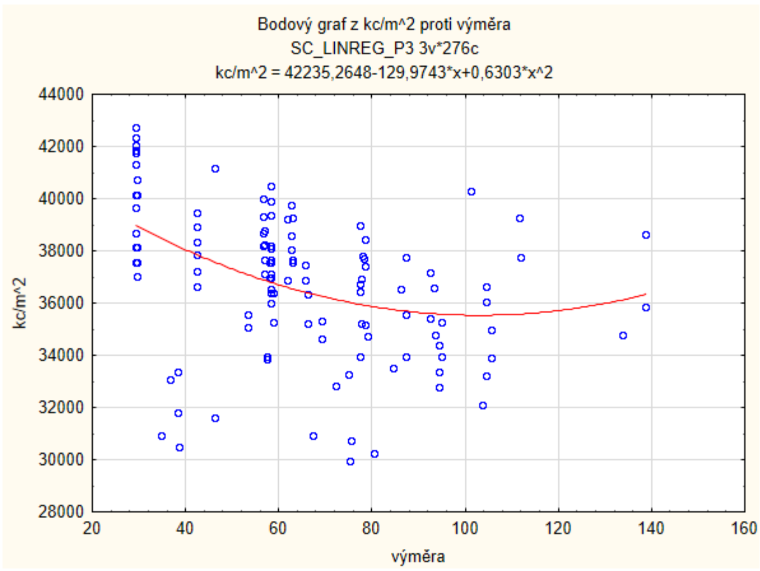
Př. 3

a) Koeficient β_1 by měl nabývat záporných hodnot, koeficient β_2 hodnot kladných. Pozorujeme nelineární vztah. Čím je byt větší tím se cena za metr čtvereční bytu snižuje, ovšem tento pokles není lineární.

b) Znaménka jsou v souladu s očekáváním. F-test vykazuje celkovou statistickou významnost modelu, t-testy ukazují na statistickou významnost jednotlivých proměnných (p hodnoty jsou nižší než hladina významnosti 5 %).

Výsledky regrese se závislou proměnnou : kc/m^2 (SC_LINREG_P3)
R= ,42298173 R2= ,17891355 Upravené R2= ,16618352
F(2,129)=14,054 p<,00000 Směrod. chyba odhadu : 2466,9

N=132	b*	Sm.chyba z b*	b	Sm.chyba z b	t(129)	p-hodn.
Abs.člen			42235,26	1256,695	33,60822	0,000000
výměra	-1,20280	0,345076	-129,97	37,289	-3,48559	0,000672
x_square	0,85149	0,345076	0,63	0,255	2,46753	0,014915



Př. 4 Všechny datové sady vykazují velmi podobné číselné charakteristiky. Pokud jimi proložíme lineární regresní přímku, zjistíme, že koeficienty jsou si velmi podobné. Ovšem při bližším pohledu na data zjistíme, že využívat k proložení regresní přímku je vhodné pouze u jednoho z případů. Proto, je nutné mít na paměti, že je vhodné před samotnou analýzou překontrolovat datový vzorek.