SLEZSKÁ UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ FAKULTA V KARVINÉ

Finanční ekonometrie

Distanční studijní text

Iveta Palečková

Karviná 2020

Obor: Ekonomie

Klíčová slova: Ekonometrický model, model lineární regrese, kointegrace, model korekce chyb, korelační analýza, kauzalita, Grangerova kauzalita, exogenní proměnná, endogenní proměnná, panelová data, časová řada, modely volatility, Box-Jenkinsonova metodologie.

Anotace: Cílem studijní opory Finanční ekonometrie je poskytnout studentům navazujících magisterských oborů informace a znalosti v oblasti finanční ekonometrie. Studijní text je rozdělen do sedmi kapitol. Nejprve jsou teoreticky vymezeny základní pojmy, jako je ekonometrický model a jeho konstrukce a typy dat. V další části je věnována pozornost finančním časovým řadám a jejich charakteristikám a je popsána lineární regrese, včetně popisuje konstrukce lineárního regresního modelu a jeho využití. V další části studijní opory jsou definovány modely jednorozměrných a vícerozměrných časových řad. Čtvrtá kapitola je věnována kauzalitě finančních řad. Následující část studijního textu popisuje kointegraci a modely korekce chyb a dále analýzu panelových dat. Poslední část je věnována nelinearitě finančních časových řad a modelům volatility.

Autor: **Doc. Ing. Iveta Palečková, Ph.D.**

Obsah

ÚVODEM

Předkládaná studijní opora si klade za cíl poskytnout studentům navazujících magisterských oborů zaměřených na ekonomii a finance informace z oblasti finanční ekonometrie. Přestože tento studijní text je primárně určen studentům Obchodně podnikatelské fakulty v Karviné Slezské univerzity v Opavě, její využití je možné i na jiných vysokých školách se zaměřením na ekonomii a finance či bankovnictví.

Jedná se o kurz v oblasti finanční ekonometrie, který je vyučován a navazujícím magisterském studiu, proto je předpokládána základní znalost statistiky a matematických metod. Studijní opora je uspořádána do sedmi kapitol. Struktura jednotlivých kapitol je popsána v části rychlý náhled studijní opory.

Protože studijní opora je zpracovávaná s distančními prvky, je vhodná i pro studenty v kombinované formě studia. Na počátku každé kapitoly najdete rychlý náhled kapitoly, který slouží zejména pro základní orientaci, co bude obsahem dané kapitoly. V části cíle kapitoly najdete, co budete po nastudování kapitoly umět a klíčová slova jsou určena zejména pro rychlou orientaci ve studijní opoře. Tyto distanční prvky slouží mimo rychlejší orientaci v textu, také k přípravě ke zkoušce. Můžete si ověřit, zda umíte definovat pojmy z klíčových slov a zda ovládáte dovednosti uvedené v cílech kapitoly.

V textu kapitoly jsou důležité definice či další důležitá látka uváděna v distančním prvku definice či k zapamatování. Dále je průběžně uváděn prvek průvodce studiem, který provádí studenta danou kapitolou a problematikou. Ve studijní opoře najdete také prakticky popsanou ukázku řešení konkrétních případových studií v ekonometrickém programu EViews, která je uváděna v prvku řešená úloha. Na konci každé kapitoly má student k dispozici otázky a odpovědi k procvičení látky z dané problematiky.

Studijní text je zároveň součástí kurzu v informačním systému a je tedy propojen s dalšími dostupnými studijními materiály. Doufám, že předkládaný studijní text bude užitečnou pomůckou při studiu.

Doufám, že předkládaný text pro Vás bude užitečný při studiu. Uvítám jakékoliv náměty a připomínky k obsahu této publikace.

Iveta Palečková (paleckova@opf.slu.cz)

Karviná, 2020

RYCHLÝ NÁHLED STUDIJNÍ OPORY

Teorie a modely - Cíle finanční ekonometrie. Specifikace a verifikace ekonometrického modelu. Využití ekonometrických programů.

Finanční časové řady a regresní analýza - Deskriptivní statistiky, stacionární a nestacionární časové řady. Regresní analýza, vícenásobná lineární regrese. Metody estimace parametrů modelu, metoda nejmenších čtverců. Evaluace a diagnostická kontrola modelu. Využití regresní analýzy v praktických příkladech, model oceňování kapitálových aktiv. Modely diskrétní volby, modely typu Logit, Probit a Tobit.

Modely jednorozměrných a vícerozměrných časových řad - Autokorelační a parciální autokorelační funkce. Autoregrese, řády autoregresních procesů (AR), model ARMA. Nestacionární časové řady a model ARIMA. Modely sezónních časových řad. Vícerozměrný lineární proces. Modely vektorové autoregrese.

Kauzalita ve finančních a časových řadách - Korelační analýza, její výhody a nedostatky. Grangerova kauzalita. Endogenita a exogenita.

Kointegrace a modely korekce chyb - Trendy a zdánlivá regrese. Testování kointegrace. Testování řádu kointegrace - metoda Johansena. Model Error Correction a vektorový model korekce chyb (VEC).

Analýza panelových dat - Výhody a nevýhody panelové regrese. Statický lineární model. Konstantní a náhodné efekty. Dynamický lineární model. Využití panelové regrese v praktických příkladech.

Nelinearita finančních časových řad a modely volatility - Testování nelinearity časových řad. Modely volatility. ARCH, GARCH modely. Asymetrické modely typu EGARCH a TARCH.

1 TEORIE A MODELY

RYCHLÝ NÁHLED KAPITOLY

V této kapitole bude vymezena definice a cíle finanční ekonometrie. Pozornost bude věnována specifikaci a verifikaci ekonometrického modelu. Základní kroky tvorby ekonometrického modelu jsou klíčovou oblastí, která je důležitá pochopit pro další kapitoly. Tato oblast bude důležitá pro další části tohoto studijního textu. Následně bude definováno využití ekonometrických modelů.

Abychom se mohli věnovat podstatě ekonometrii, bude ekonometrie a ekonometrický model nejprve vymezen. Ekonometrii bychom mohli zjednodušené popsat jako vědní disciplínu aplikující statistické nástroje a techniky v oblasti ekonomie. Ekonometrie ale navíc propojuje a rozšiřuje poznatky ekonomické teorie, matematiky a statistiky. Ekonomie, když budeme konkrétní, tak oblast financí, je věda plná nezodpovězených otázek a problémů, na které ekonometrie dokáže pomoci odpovědět. Příkladem takových otázek může být:

- Ovlivní změna úrokových sazeb množství poskytnutých úvěrů finančními institucemi v Česku?
- Jaké jsou determinanty vývoje úvěrů poskytnutými finančními institucemi?
- Jaký je dlouhodobý vztah mezi vývojem cenové hladiny a úrokovými sazbami?
- Jak predikovat korelaci mezi akciovými indexy na vybraných burzách?
- Jaké jsou determinanty vývoje kurzů akcií kótovanými na vybrané burze?

CÍLE KAPITOLY

- Definovat finanční ekonometrii.
- Definovat ekonometrický model.
- Popsat kroky konstrukce ekonometrického modelu.
- Vymezit typy dat.

KLÍČOVÁ SLOVA KAPITOLY

Ekonometrie, ekonometrický model, endogenní proměnná, exogenní proměnná, náhodná složka, verifikace modelu, časová řada, průřezová data, panelová data.

1.1 Ekonometrie

Ekonometrie je vědní disciplína aplikující statistické nástroje a techniky v oblasti ekonomie. Propojuje v sobě a rozšiřuje zejména poznatky ekonomické teorie, matematické ekonomie, ekonomické statistiky a matematické statistiky. Ekonometrie je disciplína, která umožňuje popsat vztahy ekonomických veličin. Tato vědecká disciplína má ovšem širší poslání než jen hledat vztahy. Používá se i jako nástroj pro empirické ověřování již postulovaných ekonomických zákonitostí.

K ZAPAMATOVÁNÍ

Ekonometrie se tedy zabývá vztahy mezi ekonomickými veličinami s cílem kvantitativně vyjádřit, ověřit a aplikovat ekonomické hypotézy na základě konkrétních statistických údajů, a to použitím matematicko-statistických metod.

Podstatou ekonometrie je měřit a popisovat co nejlépe závislosti a vztahy mezi ekonomickými veličinami, popřípadě i jinými veličinami.

Ekonometrii můžeme rozdělit na teoretickou a aplikovanou a také na klasickou a Beyesiánskou. Bayesiánský přístup vychází z několika základních pravidel pravděpodobnosti, což je jedna z jeho hlavních výhod. Ze základních pravidel pravděpodobnosti se vychází jak při odhadování parametru modelu, tak při predikci, tak i při dalších charakteristikách, což činí Bayesiánský přístup velmi univerzálním. V rámci kurzu Finanční ekonometrie bude pozornost věnována klasické ekonometrii. Beyesiánská analýza je určena zejména pro zájemce o studium.

1.1.1 Finanční ekonometrie

Pojem finanční ekonometrie se používá pro jakoukoli kvantitativní analýzu finančních dat jak na mikroekonomické, tak na makroekonomické úrovni. Přitom kvantitativní analýzou se rozumí především statistické zpracování finančních dat (tj. klasická popisná statistická analýza, statistická identifikace, odhad a verifikace příslušného modelu, statistické testování různých finančních hypotéz a konstrukce předpovědí v rámci zkonstruovaného modelu) s využitím dostupných ekonomických informací a pak následná diskuse získaných výsledků z hlediska jejich kompatibility s praxí a dopadu v rámci dané finanční reality, tedy finančních trhů, finančního sektoru apod. (Cipra, 2014).

PRŮVODCE STUDIEM

Většinu praktických výpočtů v rámci finanční ekonometrie se doporučuje provádět s využitím vhodného softwaru. Pro praktické ukázky jednotlivých příkladů v tomto studijním textu bude využit ekonometrických software EViews. Studentská verze je ke stažení zdarma na webových stránkách www.eviews.com. Další software, který lze využít je STATA (www.stata.com). Jako další software, který lze pro studium využít je R. Jedná se o open source, tedy software zdarma dostupný. Lze doporučit např. R-project (www.r-project.org) nebo R studio (www.r-studio.com).

Využití EViews je zejména z toho důvodu, že je uživatelsky přátelský, jedná se o grafické prostředí a jsou zde dostupné veškeré operace, které budou ve studijním textu popsány. Tento software je také často používán v praxi. Jakmile se naučíte pracovat s jedním software, tak pomocí nápovědy či základního tutoriálu není problém využít jiný. Jde zejména o pochopení základních nástrojů, které lze prakticky aplikovat.

1.2 Ekonometrický model

Pro definici ekonometrického modelu je nejprve nutné vymezit ekonomický model. Podstatou ekonometrie je systematické hledání kvantitativních odpovědí na ekonomické otázky a problémy. Vše se odvíjí od ekonomické teorie, která je spojená s problémem, o který se zajímáme. V ekonomii jsou vyjádřeny vztahy mezi ekonomickými veličinami v podobě matematických funkcí.

Ekonometrický model je vlastním nástrojem ekonometrie. Ekonomický model umožňuje a usnadňuje matematickou a statistickou formalizaci teoretických poznatků, dává do souvislosti ekonomické veličiny, jež jsou předmětem zájmu. Ekonomické modely tedy popisují vztahy mezi ekonomickými proměnnými. Ekonomické modely jsou používány v rámci ekonomické analýzy. Ekonometrický model převádí ekonomický model do podoby, která je s pomocí ekonometrických nástrojů a technik analyzovatelná. Ekonometrický model algebraicky vyjadřuje konkrétní vztah mezi veličinami využívající matematický aparát. Ekonometrickým modelem rozumíme takový ekonometricko-matematický model, který má charakter statistického modelu, tzn. Že má přesně specifikovanou funkční formu, přičemž jsou navíc statisticky definovány jeho náhodné složky (proměnné), které představují náhodné chyby (odchylky) funkčních rovnic (Hušek a Walter, 1976).

1.3 Základní kroky sestavování ekonometrického modelu

Obrázek 1 zachycuje algoritmus, který se většinou dodržuje při konstrukci ekonometrického modelu. Avšak je důležité zdůraznit, že konstrukce ekonometrického modelu je iterativní proces, kdy častou až metodou pokusů a omylů lze dospět ke konečnému modelu. Také dva různí statistikové, kteří provádějí nezávisle na sobě analýzu stejného problému, mohou dospět k formálně velmi odlišným výsledkům, které by však měly implikovat podobné praktické závěry (Cipra, 2014).

K ZAPAMATOVÁNÍ

Kvantitativní analýza na základě ekonometrického modelu je vícestupňovou abstrakcí:

1. Formulace finančního problému
2. Formulace a specifikace modelu a sběr dat
3. Kvantifikace ekonometrického modelu
4. Verifikace ekonometrického modelu
5. Interpretace a praktické využití ekonometrického modelu

1.3.1 FORMULACE FINANČNÍHO PROBLÉMU

Zjednodušeně můžeme říct, že prvním krokem by vždy měla být jednoznačná formulace finančního problému, který chceme řešit. Tedy konkrétně určit, co je cílem naší práce a co konkrétně chceme zjišťovat. K tomu je nutné vymezit příslušný teoretický rámec, provedení rešerše dostupné literatury řešící daný problém. Tomuto kroku sice nebudeme věnovat ve studijním textu větší pozornost, avšak je to stěžejní krok, který je nutné vždy provést. Při zpracovávání jakékoli práce či zadaného úkolu je tento krok jeden z nejdůležitějších.

1.3.2 FORMULACE A SPECIFIKACE MODELU

Ve druhém kroku následuje formulace a specifikace ekonometrického modelu a je nutné provést adekvátní sběr dat. Data můžeme mít z veřejně dostupných zdrojů, jako jsou databáze, statistiky centrálních bank, výroční zprávy jednotlivých finančních institucí apod. Data také mohou pocházet z dotazníků a vlastního šetření. Nejprve je nutno data zpracovat, statisticky testovat a připravit k použití pro ekonometrický model.

Konkrétní specifikace a formulace ekonometrického modelu spočívá ve spojení teoretických poznatků s informacemi o konkrétním problému, který je předmětem kvantitativní analýzy (Hušek, 1999). Formulaci ekonometrického modelu můžeme dále rozdělit do tří kroků. Specifikace ekonometrického modelu pak obsahuje tyto tři kroky:

- Určení závisle a nezávisle proměnných, zahrnutých do zkoumání.
- Stanovení předpokládaných znamének a očekávaných hodnot parametrů modelu.
- Volbu typu a analytické formy modelu a jeho jednotlivých rovnic.

1. Určení závisle a nezávisle proměnných, zahrnutých do zkoumání

Tyto proměnné se v ekonometrii rozdělují především na endogenní a exogenní. **Endogenní proměnné** jsou takové proměnné, kdy jejich hodnoty jsou určeny, neboli generovány modelem. Většina ekonomických proměnných je endogenního typu. Endogenní proměnné mají zpravidla postavení vysvětlovaných proměnných, avšak v některých případech mohou v modelu vystupovat i jako vysvětlující proměnné.

Exogenní proměnné jsou takové proměnné, kdy jejich hodnoty nejsou modelem determinovány, ale jsou dány mimo modelovaný systém. Exogenní proměnné působí na zkoumaný systém, ale samy jím nejsou ovlivňovány. Exogenní proměnné mají vždy charakter proměnných vysvětlujících.

V dynamických ekonometrických modelech se vyskytují často časově **zpožděné proměnné**. Důvodem zahrnutí zpožděných hodnot je, že v ekonomice reagují jevy s určitým zpožděním na daný podnět. Jedná se o takové vysvětlující proměnné, které vyjadřují efekt jedné jednotky určité endogenní proměnné v některém z předcházejících časových období na úroveň stejné nebo jiné vysvětlované endogenní proměnné v běžném období.

Do modelu se zahrnují obvykle pouze nejdůležitější vysvětlující proměnné a vliv všech nepodstatných činitelů se zahrnuje spolu s působením náhodných faktorů do **náhodné složky** či proměnné, jejíž hodnoty nelze získat pozorováním. Náhodné proměnné jsou tedy složky jednotlivých strukturálních rovnic modelu, které vyjadřují teoretické příčinné vztahy a interakce mezi ekonomickými proměnnými ověřované na statistických údajích Náhodná proměnná (složka) v určité strukturální rovnici je tedy rovna odchylce skutečné hodnoty vysvětlované endogenní proměnné od její teoretické hodnoty vyjádřené funkcí hodnot vysvětlujících proměnných. Na rozdíl od endogenních i exogenních proměnných nelze hodnoty náhodných proměnných zjistit, proto je vždy nutné učinit o nich pouze určité předpoklady, jejichž platnost je nutné ověřit.

2. Stanovení předpokládaných znamének a očekávaných hodnot parametrů modelu

Znaménka jednotlivých parametrů se určují na základě teorie nebo se k tomu využívají informace získané z jiných kvantitativních analýz či studií. Obdobně na základě teoretických závěrů nebo apriorních informací lze předem usuzovat i na očekávané hodnoty parametrů modelu.

3. Volbu typu a analytické formy modelu a jeho jednotlivých rovnic.

Stejně jako při klasifikaci proměnných modelu nedává ekonometrická teorie konkrétní návod ani u volby formy modelu a počtu rovnic či jejich vzájemných vztazích. Teorie řík pouze to, zda zkoumaná závislost ekonomických veličin je přímá či nepřímá. V ekonometrické analýze používáme tyto základní typy modelů:

a) jednorovnicový model – má nejčastěji charakter stochastického regresního modelu, tj. vyjadřuje závislost jedné vysvětlované endogenní proměnné na jedné nebo na více vysvětlujících exogenních proměnných a na náhodné složce
b) vícerovnicový model – představovaný soustavou rovnic, přičemž každou z nich lze zkoumat buď odděleně jako zvláštní jednorovnicový model, nebo všechny rovnice jako celek
c) simultánní model – tvořený soustavou simultánně závislých rovnic. Simultánní závislost spočívá v tom, že endogenní proměnné vystupují simultánně, tj. současně jak ve funkci vysvětlovaných, tak i vysvětlujících proměnných a jsou soustavou rovnic simultánně určeny.

1.3.3 Kvantifikace ekonometrického modelu

Kvantifikace neboli odhad ekonometrického modelu slouží k získání numerických hodnot jeho parametrů. Jedná se o odhad parametrů vhodnou metodou. Jednotlivým metodám bude věnována pozornost v dalším textu studijní opory.

Kvantifikace modelu začíná již shromážděním vhodných statistických dat, tedy dostatečného počtu pozorovaných hodnot všech proměnných obsažených v modelu. Vlastní odhad parametrů modelu spočívá ve výběru adekvátního odhadovaného postupu obvykle na všechna pozorování jednotlivých proměnných. Při volbě konkrétního odhadovaného postupu přihlížíme k řadě faktorů, jako jsou např. vlastnosti odhadů, získaných aplikací určité metody odhadu, náročnost jednotlivých způsobů odhadu, kvalita a kvantita dat, která jsou k dispozici, účel, pro který má být odhadovaný model použitý apod.

Na základě odhadnutých parametrů ekonometrického modelu a pomocí napozorovaných hodnot vysvětlujících proměnných stanovíme teoretické hodnoty všech vysvětlovaných endogenních proměnných, čímž dospějeme k hledanému řešení ekonometrického modelu.

Je nutné si však uvědomit, že vysvětlení změn endogenních proměnných, k němuž jsme dospěli na základě odhadnutého ekonometrického modelu, je neúplné a zároveň podmíněné hodnotami těch proměnných, které nebyly pro zjednodušení do modelu zahrnuty. Dokonce i v případě, že je model rozsáhlý, si zachovává svůj podmíněný charakter v tom, že neobsahuje některé proměnné, které jsou vysvětleny nebo určeny mimo rámec uvažovaného modelu.

1.3.4 Verifikace ekonometrického modelu

Dalším krokem formulace ekonometrického modelu je jeho verifikace. Verifikace ekonometrického modelu spočívá v ověření toho, zda odhadnuté parametry jsou teoreticky správné a současně i statisticky významné. Zda hodnota odhadnutých parametrů vůbec odpovídá ekonomické teorii.

- Ekonomická verifikace modelu – ověření teoreticky předpokládaných znamének a apriorních omezení numerických hodnot odhadnutých parametrů, tj. ekonomických konstant.
- Statistická verifikace – slouží k posouzení statistické významnosti odhadnutých parametrů i ekonometrického modelu jako celku. Je založena na statistických kritériích neboli statistických testech (nazvaných testy prvního řádu).
- Ekonometrická verifikace modelu (testy druhého řádu) – používají se k ověřování splnění předpokladů potřebných k aplikaci konkrétních ekonometrických metod a technik. Spočívají tedy v testování statistických testů nebo pomocí ekonometrických kritérií se ověřuje platnost či oprávněnost použití statistických kritérií, zejména v případě malého rozsahu napozorovaných dat.

1.3.5 INTERPRETACE A PRAKTICKÉ VYUŽITÍ EKONOMETRICKÉHO MODELU

Posledním krokem konstrukce ekonometrického modelu je způsob jeho využití. Jedná se tedy o aplikaci modelu. Numerické odhady parametrů modelu slouží převážně k analýze, tj. verifikaci výchozí ekonomické teorie, nebo k prognózování budoucích hodnot vysvětlovaných endogenních proměnných a k výběru hospodářské politiky pro potřeby optimálního řízení. Jedná se tedy o nejdůležitější část.

V prvním kroku jsme definovali finanční problém, který jsme chtěli řešit a nyní bychom měli znát odpověď na naši otázku. Je nutné ekonomicky interpretovat získané poznatky a provést aplikaci do praxe. Jednotlivé kroky, přestože jim byla věnována výrazná pozornost, slouží pouze k tomu, abychom měli nástroj k nalezení odpovědi na naši položenou otázku či formulovanou hypotézu.

1.4 Typy dat

V rámci popisu konstrukce ekonometrického modelu bylo zmíněno, že velmi důležitým krokem je sběr dat. Data jsou klíčová pro provedení ekonometrické analýzy.

K ZAPAMATOVÁNÍ

V rámci finanční ekonometrie lze klasifikovat analyzovaná data do tří skupin:

1. Časová data
2. Průřezová data
3. Panelová data

1.4.1 ČASOVÁ DATA

Data ve tvaru časových řad, tj. hodnoty určité veličiny pozorované v určitém časovém intervalu s určitou frekvencí záznamu. Jedná se o pozorování proměnných pro nějakou jednotku (stát, finanční instituce, domácnost apod). v čase.

Pro časová data je důležité jejich chronologické uspořádání v čase, které nelze přerovnávat. Data jsou seřazena podle času tak, jak byla získána. Frekvencí pozorování se rozumí velikost intervalu mezi jednotlivými pozorováními (např. kalendářní měsíc), nebo pravidelnost s jakou je záznam pořizován (např. každý obchodní den). Časové řady se mohou lišit podle toho, s jakou frekvencí je získáváme. Obvykle pracujeme s daty ročními (kdy hodnota příslušné proměnné je zaznamenávána pravidelně každý rok), s daty čtvrtletními (údaj získáváme čtyři krát do roka), s daty měsíčními a s daty denními.

Příkladem časové řady je například množství poskytnutých úvěrů klientům Komerční bankou od 1. 1. 2000 do 31. 12. 2019. Pokud budeme mít data z výročních zpráv Komerční banky, budeme mít roční data. Dalším příkladem může být vývoj kurzu akcie společnosti Coca-Cola HBC AG v indexu FTSE 100 v období od 1. 1. 2010 do 31. 12. 2019. Tyto data jsou dostupná s denní frekvencí.

1.4.2 PRŮŘEZOVÁ DATA (CROSS-SECTION)

Často můžeme pracovat s daty, která nemají časový rozměr. Tyto data jsou vztažena ke speciálním jednotkám, kterými mohou být například podniky, státy, domácnosti, finanční instituce apod. Průřezová data jsou data ve tvaru průřezového výběru, tj. hodnoty určité veličiny (nebo veličin) pozorované v tentýž časový okamžik přes určitý populační soubor. Pro průřezová data není obvykle důležité jejich uspořádání (abecední, regionální atd.), takže je většinou lze libovolně přerovnávat.

Příkladem takových dat je např. množství poskytnutých úvěrů všech bank na území Česka k 31. 12. 2019. Budeme mít tedy tolik hodnot, kolik bank operuje v českém bankovním sektoru. Dalším příkladem je vývoj kurzu všech společností kótovaných na londýnské burze k 20. 3. 2020.

Dalším aspektem rozdělení dat je dělení na data:

- kvantitativní data
- kvalitativním datům

Kvantitativní data jsou například kurz akcie dané společnosti obchodované na burze, výše zisku vybrané banky apod. V případě, že máme například data z dotazníkového šetření mezi zaměstnanci, mohou být odpovědi také „ano“ a „ne“. Například v dotazníkovém šetření se ptáme, zda jsou zaměstnanci spokojeni se současným systémem jejich hodnocení. Zda mají zaměstnanci zájem o stravování na pracovišti apod. Tento typ dat se může objevit také ve finanční ekonometrii. Můžeme zohledňovat vliv finanční krize na určitou veličinu.

Tedy máme odpověď „ano" či „ne" na otázku, zda v dané ekonomice byla finanční krize. Dále můžeme zohledňovat fakt, zda daná společnost má akcie veřejně obchodované na burze, opět máme odpověď „ano" nebo „ne" apod. Tento typ získaných údajů se nazývá kvalitativní data.

V ekonometrii se kvalitativní data převádějí do formy numerické. V případě, že máme opověď „ano" zaměníme ji za hodnotu jedna a naopak odpověď „ne" zaměníme za hodnotu nula (či naopak dle požadovaného zadání řešeného problému). Tato proměnná, která nabývá pouze hodnot 0 či 1 nazýváme umělou či dummy proměnnou (někdy se můžete setkat také s pojmem binární proměnná).

1.4.3 PANELOVÁ DATA

V praxi mají data, se kterými pracujeme jak časovou, tak i průřezovou složku dohromady. Tento typ dat se nazývá panelová data. Jedná se tedy o kombinaci časových a průřezových dat. Panelová data jsou data, kde jsou charakteristiky za jednotlivá pozorování zjišťovány za více časových období. Soubor panelových dat tvoří vzorek obsahující N průřezových jednotek, které jsou pozorovány v různých T časových obdobích. Panelová data obsahují opakované zjišťování stejných jednotek (jednotlivci, domácnosti, firmy apod.) shromážděné v průběhu několika období. Data musí být opět shromážděna chronologicky.

Příkladem je vývoj poskytnutých úvěrů deseti českých bank v období 1. 1. 2000 až 31. 12. 2019. Nebo hrubý domácí produkt všech zemí Evropské unie v letech 2000 až 2020. Výše hrubého zisku všech bank operujících v českém bankovním sektoru za posledních deset let apod.

K ZAPAMATOVÁNÍ

Můžeme tedy velmi zjednodušeně shrnout typy dat následovně:

- Časová data – pozorování jednoho jevu pozorovaného přes několik časových období.
- Průřezová data – pozorování mnoha jevů v jednom časovém okamžiku.
- Panelová data – pozorování mnoha jevů po několik časových období.

SHRNUTÍ KAPITOLY

První kapitola se věnovala definování ekonometrie, což je vědní disciplína, která aplikuje statistické nástroje a techniky v oblasti ekonomie. Finanční ekonometrie pak aplikuje statistické nástroje a techniky konkrétně v oblasti financí. Pro pochopení problematiky byl definován ekonometrický model, který umožňuje matematickou a statistickou formalizaci teoretických poznatků a dává do souvislosti ekonomické veličiny, které jsou předmětem zkoumání. Znalost kroků sestavení ekonometrického modelu je velmi důležité pro pochopení celé ekonometrie a dalších modelů popsaných v dalších částek studijního textu. Nejprve je nutné formulovat finanční problém, dále formulovat a specifikovat model a sběr dat. Dalším krokem je kvantifikace ekonometrického modelu a následuje jeho verifikace, která je ekonomická, statistická a ekonometrická. V poslední části je nutné interpretovat a prakticky využít ekonometrický model. Jednotlivé kroky sestavení ekonometrického modelu byly podrobně v kapitole popsány. V neposlední řadě byly popsány typy dat, rozlišujeme časová, průřezová či panelová data.

OTÁZKY

1. Proměnné, které jsou generovány modelem se nazývají:
   a) endogenní
   b) exogenní
   c) náhodné
2. K ověření splnění předpokladů potřebných k aplikaci ekonometrických metod a technik slouží:
   a) ekonomická verifikace modelu
   b) statistická verifikace modlu
   c) ekonometrická verifikace modelu
3. Cena akcií společnosti KOFOLA ČS za období od 1. 1. 2015 do 31. 12. 2020 je příkladem dat:
   a) časových
   b) průřezových
   c) panelových
4. Kvantifikace modelu slouží k získání numerických hodnot jeho parametrů. ANO x NE
5. Ve fázi formulace a specifikace ekonometrického modelu je nutné určit závislé a nezávislé proměnné, které jsou zahrnuty do modelu. ANO x NE

ODPOVĚDI

1. a)
2. c)
3. a)
4. NE
5. ANO

2 FINANČNÍ ČASOVÉ ŘADY A REGRESNÍ ANALÝZA

RYCHLÝ NÁHLED KAPITOLY

Ve druhé kapitole studijního textu budou popsány finanční časové řady a jejich charakteristiky. Pro tvorbu ekonometrického modelu je důležitou fází sběr dat. Nejprve bude tedy popsán sběr dat a jejich úprava. Data zpravidla potřebujeme prezentovat, možností prezentace dat je několik, například pomocí grafů nebo deskriptivní statistiky. Pro celou práci ekonometra je nutná znalost testování hypotéz. V kapitole budou teoreticky popsány postupy testování hypotéz, což bude následně prakticky ukázáno v ekonometrickém software EViews. V další části je věnována pozornost lineární regresní analýze. Budou představeny metody pro odhad parametrů lineárního regresního modelu. Největší pozornost bude věnována metodě nejmenších čtverců a předpokladům jejího použití. Pomocí řešeného příkladu bude přiblížen lineární regresní model v ekonometrickém programu EViews. Poslední část kapitoly se bude věnovat modelům diskrétní volby.

CÍLE KAPITOLY

- Popsat sběr dat.
- Popsat možnosti prezentace dat.
- Definovat stacionaritu řad.
- Přestavit testování hypotéz v ekonometrii.
- Definovat lineární regresní model.
- Definovat předpoklady metody nejmenších čtverců.

KLÍČOVÁ SLOVA KAPITOLY

Finanční časové řady, lineární regresní analýza, histogram, deskriptivní statistika, stacionarita, metoda nejmenších čtverců, multikolinearita, normalita, heteroskedasticita, autokorelace. logit a probit model.

2.1 Finanční časové řady a jejich charakteristiky

Finanční trhy jsou jedním ze základních stavebních kamenů finančního systému. Finanční trhy představují základ finančního systému v každé tržní ekonomice a plní řadu významných funkcí. Na finančních trzích se soustřeďuje nabídka dočasně volných finančních prostředků přebytkových ekonomických subjektů a poptávka deficitních subjektů po těchto prostředcích, realizují se procesy jejich směny. Finanční trhy tedy umožnují přesun těchto prostředku v ekonomice od přebytkových subjektů k subjektům deficitním. Tato alokace finančních prostředků se na finančních trzích uskutečňuje prostřednictvím různých druhů finančních dokumentů (Polouček et al., 2009).

Základní informací finančních trhů je cena (cena akcie, cena měny, cena dluhopisu…), ceny jsou sledovány v určité časové frekvenci a tvoří tak časové řady. Tyto časové řady, stejně i řady vycházející z cen nebo charakterizující ceny a jejich vývoj se označují jako finanční časové řady. Základní rys finančních časových řad je vysoká časová frekvence jednotlivých hodnot, nejčastěji jsou tyto hodnoty zaznamenávány v denní frekvenci (blíže viz 1. kapitola).

2.1.1 SBĚR A ÚPRAVA DAT

Podstatou správně sestaveného ekonometrického modelu je správný výběr dat a jejich úprava. Nejprve je nutné definovat, jaká data budeme potřebovat, v jaké frekvenci (denní, měsíční, kvartální, roční), v jaké měně budeme mít data, mezi úpravu dat patří také přepočet na shodné jednotky (tempa růstu apod.), logaritmování dat apod.

PRŮVODCE STUDIEM

V následující řešené úloze bude ve stručnosti popsán postup stažení dat z Yahoo Finance.

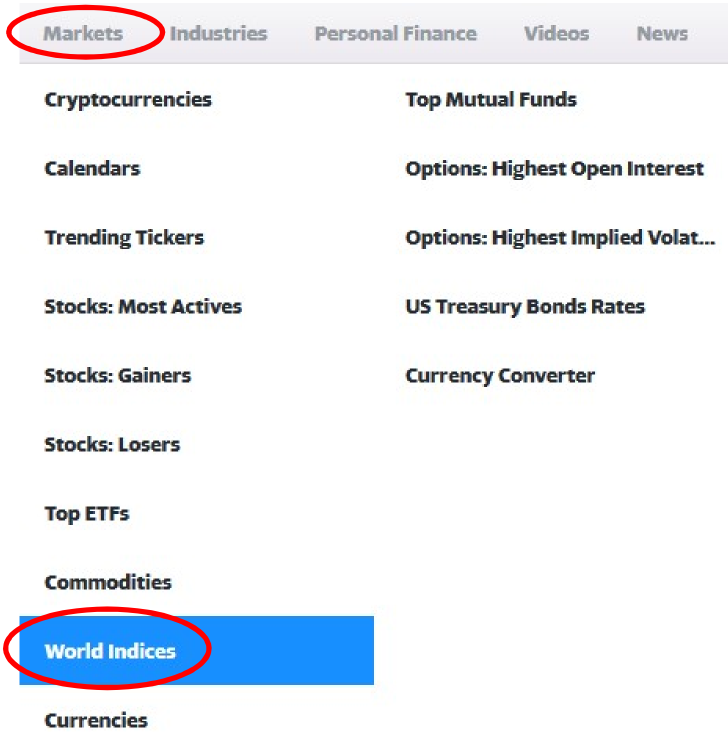
STAŽENÍ A IMPORT DAT

ŘEŠENÁ ÚLOHA

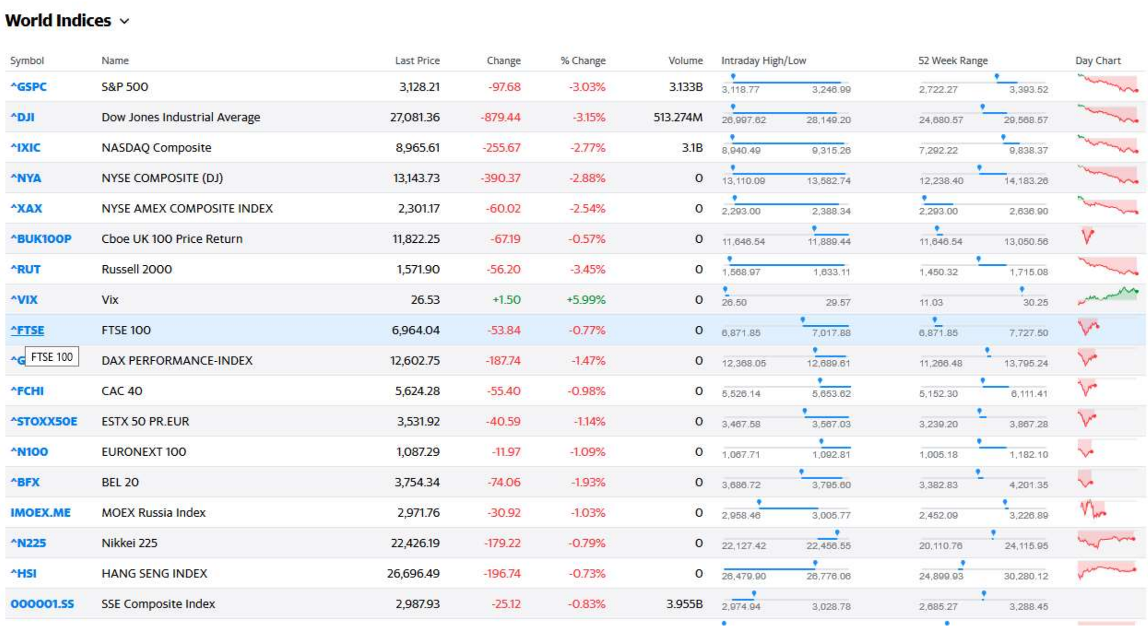
Ukážeme si příklad stažení finančních časových řad z veřejně dostupných zdrojů. S těmito daty můžete dál pracovat a aplikovat na ně jednotlivé úlohy. Stažení dat také využijete pro potřeby samostatné případové studie.

Finanční data můžeme stahovat z placených databází či veřejně dostupných. Jako příklad si uvedeme stažení vývoje kurzu akcií za posledních 5 let. Jedná se tedy o finanční časovou řadu. Data budeme stahovat z webové stránky finance.yahoo.com. V části „Market“ vybereme „World Indices“ (můžete také zvolit měny či komodity).

Obrázek



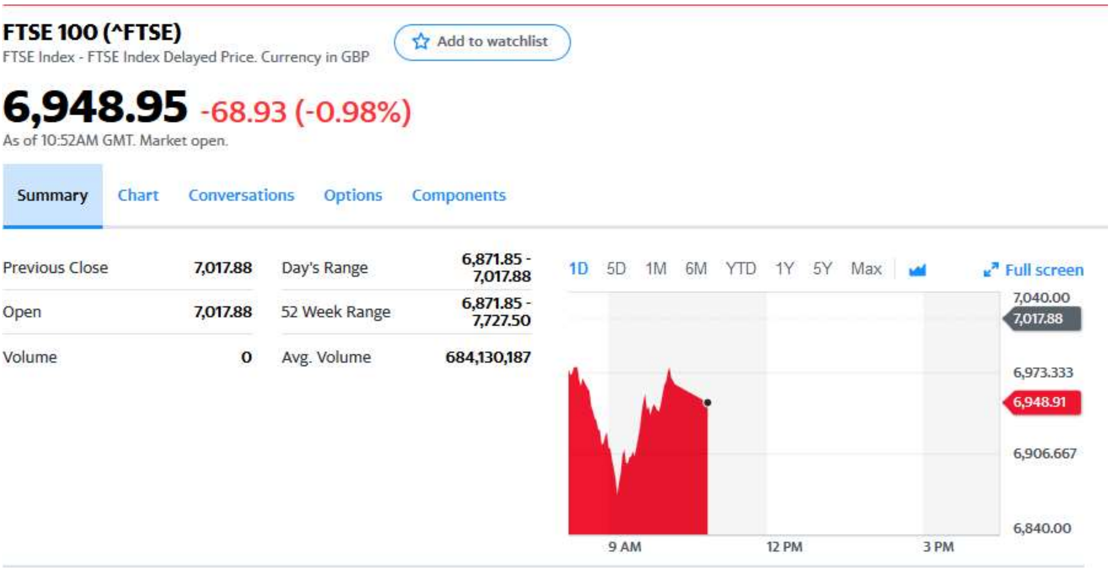
Poté vybereme námi zvolený akciový index. Ukážeme si na příkladu výběru FTSE 100:



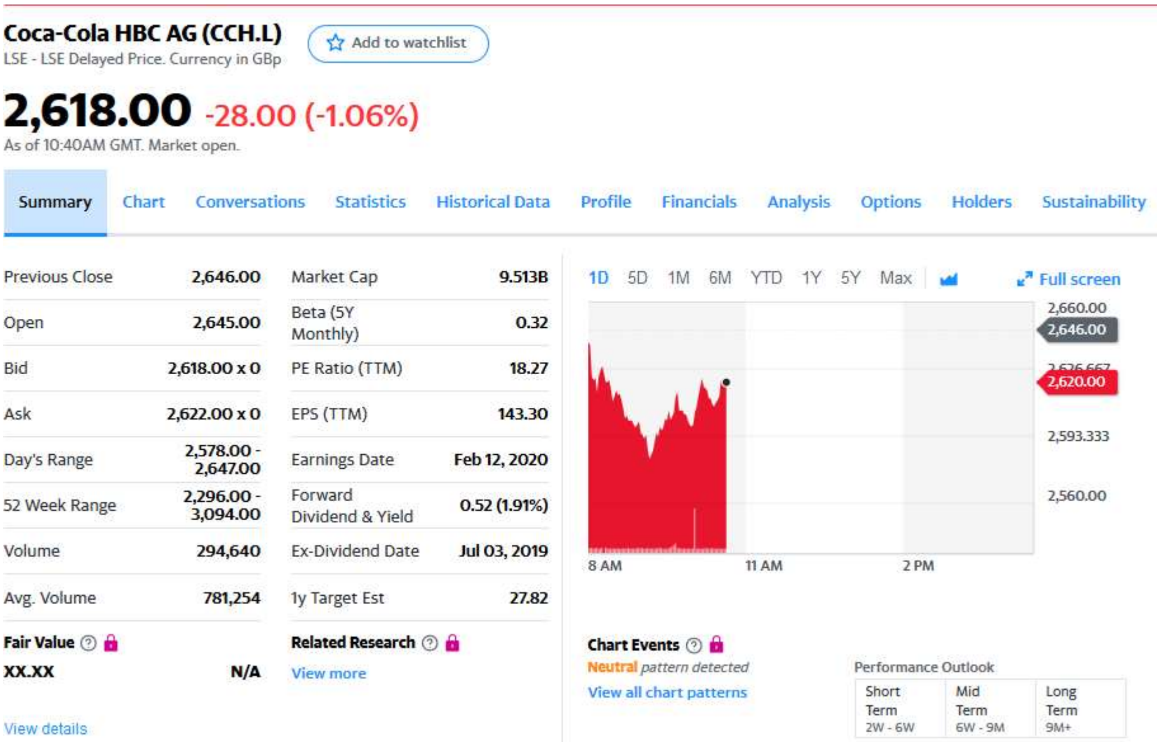
World Indices

Symbol	Name	Last Price	Change	% Change	Volume	Intraday High/Low	52 Week Range	Day Chart
^GSPC	S&P 500	3,128.21	-97.68	-3.03%	3.133B	3,118.77 3,246.99	2,722.27 3,393.52	
^DJI	Dow Jones Industrial Average	27,081.36	-879.44	-3.15%	513.274M	26,997.62 28,149.20	24,680.57 29,568.57	
^IXIC	NASDAQ Composite	8,965.61	-255.67	-2.77%	3.1B	8,940.49 9,315.26	7,292.22 9,838.37	
^NYA	NYSE COMPOSITE (DJ)	13,143.73	-390.37	-2.88%	0	13,110.09 13,582.74	12,238.40 14,183.26	
^XAX	NYSE AMEX COMPOSITE INDEX	2,301.17	-60.02	-2.54%	0	2,293.00 2,388.34	2,293.00 2,636.90	
^BUK100P	Cboe UK 100 Price Return	11,822.25	-67.19	-0.57%	0	11,646.54 11,889.44	11,646.54 13,050.56	
^RUT	Russell 2000	1,571.90	-56.20	-3.45%	0	1,568.97 1,633.11	1,450.32 1,715.08	
^VIX	Vix	26.53	+1.50	+5.99%	0	26.50 29.57	11.03 30.25	
^FTSE	FTSE 100	6,964.04	-53.84	-0.77%	0	6,871.85 7,017.88	6,871.85 7,727.50	
^G FTSE 100	DAX PERFORMANCE-INDEX	12,602.75	-187.74	-1.47%	0	12,368.05 12,689.61	11,266.48 13,795.24	
^FCHI	CAC 40	5,624.28	-55.40	-0.98%	0	5,526.14 5,653.62	5,152.30 6,111.41	
^STOXX50E	ESTX 50 PR.EUR	3,531.92	-40.59	-1.14%	0	3,467.58 3,567.03	3,239.20 3,867.28	
^N100	EURONEXT 100	1,087.29	-11.97	-1.09%	0	1,067.71 1,092.81	1,005.18 1,182.10	
^BFX	BEL 20	3,754.34	-74.06	-1.93%	0	3,686.72 3,795.60	3,382.83 4,201.35	
IMOEX.ME	MOEX Russia Index	2,971.76	-30.92	-1.03%	0	2,958.46 3,005.77	2,452.09 3,226.89	
^N225	Nikkei 225	22,426.19	-179.22	-0.79%	0	22,127.42 22,456.55	20,110.76 24,115.95	
^HSI	HANG SENG INDEX	26,696.49	-196.74	-0.73%	0	26,479.90 26,776.06	24,899.93 30,280.12	
000001.SS	SSE Composite Index	2,987.93	-25.12	-0.83%	3.955B	2,974.94 3,028.78	2,685.27 3,288.45	

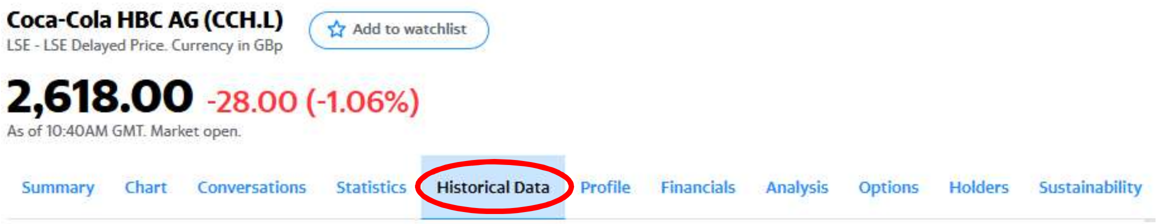
Po otevření vidíme možnosti: souhrn, graf, možnosti a komponenty. Vybereme možnost „Components“, kde se nám zobrazí veškeré komponenty v rámci daného indexu.

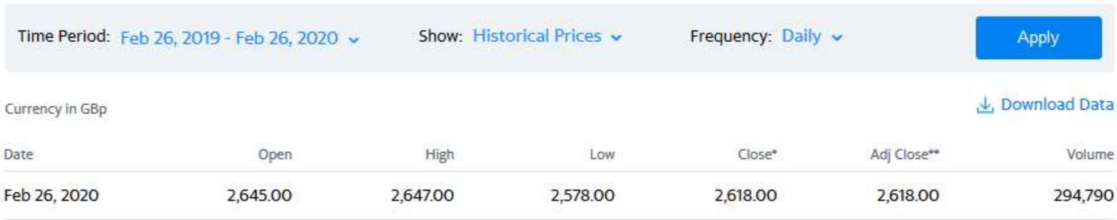


Vybereme zde libovolnou akcii. Ukážeme si na příkladu Coca-Cola HBC AG. Opět vidíme souhrn o dané akcii, graf a další možnosti.



Pokud chceme stáhnout časovou řadu, vybereme možnost „Historical Data“. Poté zvolíme dané časové období a frekvenci a dáme „Apply“. Poté můžeme dát „Download Data“ a finanční řada se stáhne.





Dále je nutné data upravit do požadovaného formátu a seřadit chronologicky od nejstarších k nejnovějším. Jakmile máme data takto upravená, můžeme uložit a pracovat s nimi.

SAMOSTATNÝ ÚKOL

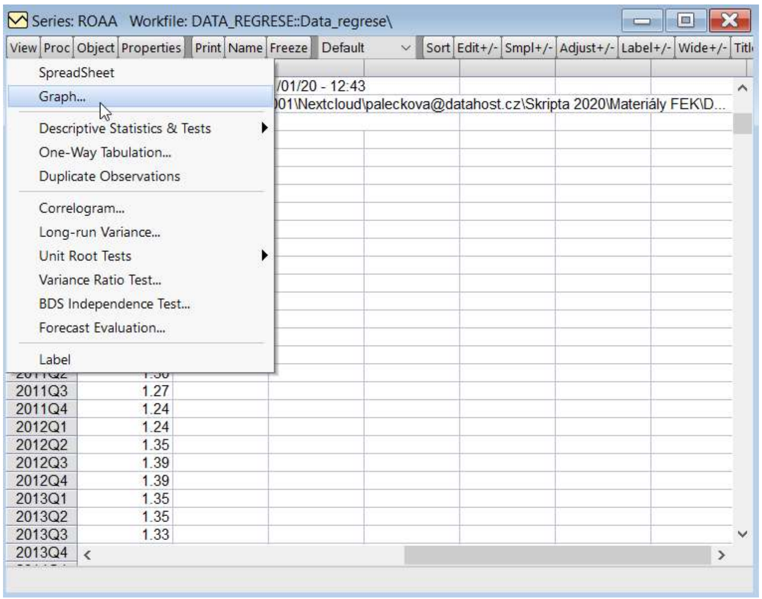
Stáhněte si z libovolného zdroje dvě časové řady a naimportujete převeďte je do MS Excel.

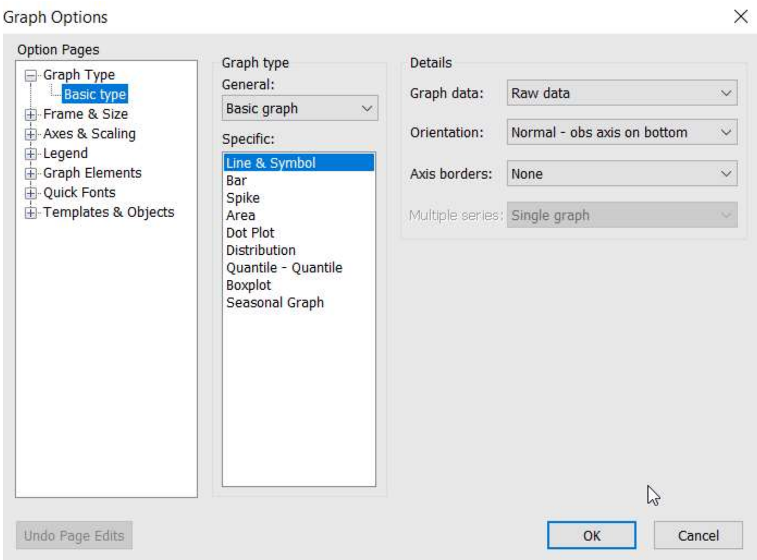
2.1.2 GRAFY

Pracujeme-li s daty, je důležité tato data přehledným a výstižným způsobem prezentovat. Pro čtenáře není atraktivní pročítat celkové položky databáze či tabulky, ale chce mít předložená data do kompaktnější podoby s vyšší informační hodnotou. Proto jsou užitečným způsobem prezentace dat grafy a tabulky. Existuje široká paleta grafů – sloupcové, koláčové atd.

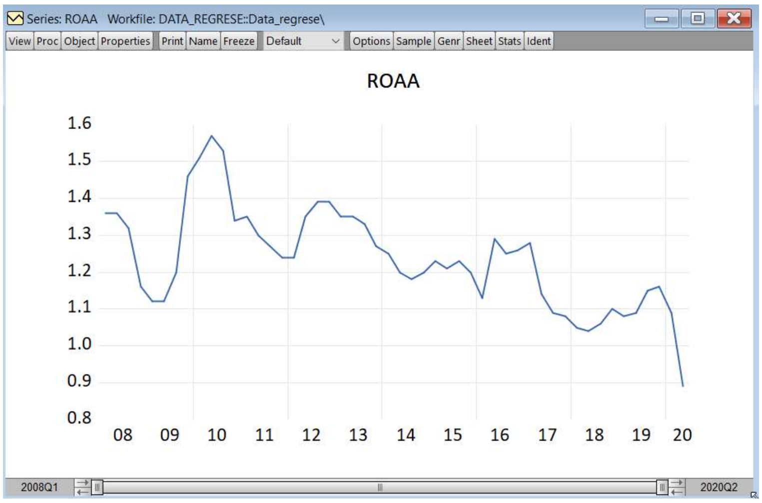
ŘEŠENÁ ÚLOHA

Otevřeme v EViews časovou řadu, ze které chceme vytvořit graf.



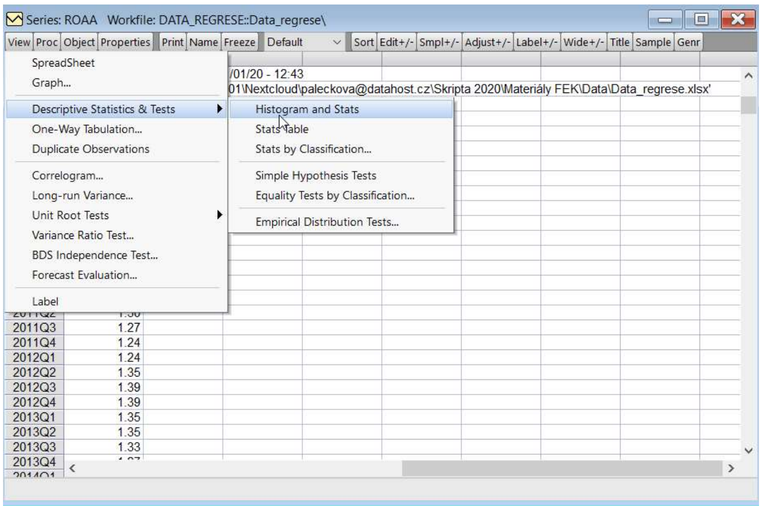


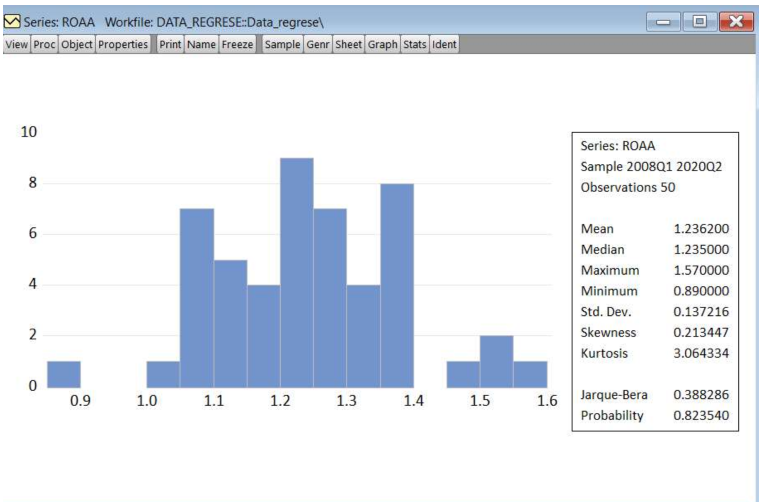
Nyní můžeme vybírat typ grafu (např. spojnicový, boxplot) a jaká data chceme použít. Můžeme také vytvořit graf z více časových řad najednou. Lze si zvolit, zda znázorníme veškerá data v jednom grafu nebo samostatně každou řadu do jiného grafu. Výběr grafu závisí na tom, jaká data modelujeme, zda máme veškerá data ve stejných jednotkách (peněžní vyjádření nebo procenta) apod.



Histogram

Graf časové řady je informativní, pokud jde o vývoj v čase. V případě průřezových dat je tento druh zobrazení naprosto nevhodný a je nutné data prezentovat jiným způsobem. Jedním ze způsobů je využití např. histogramu.



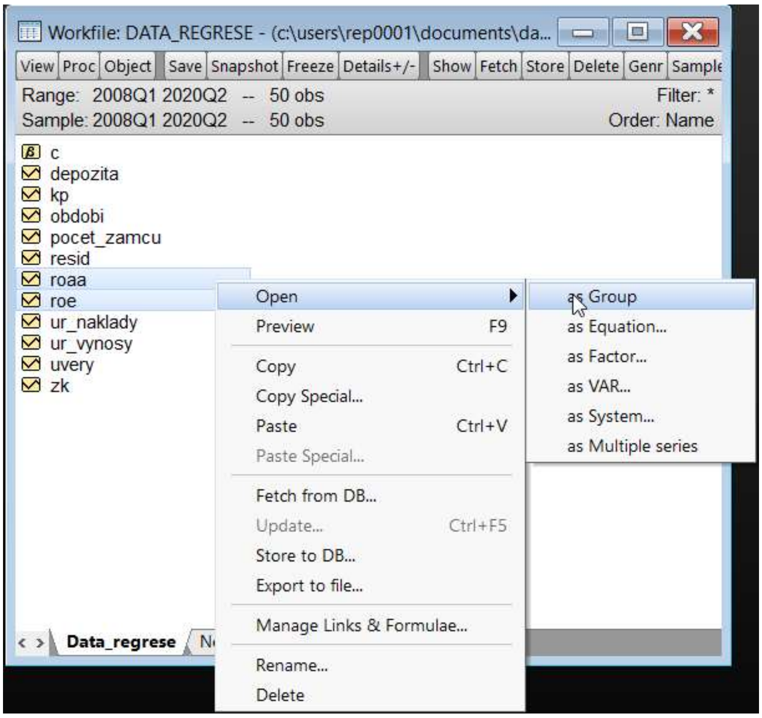


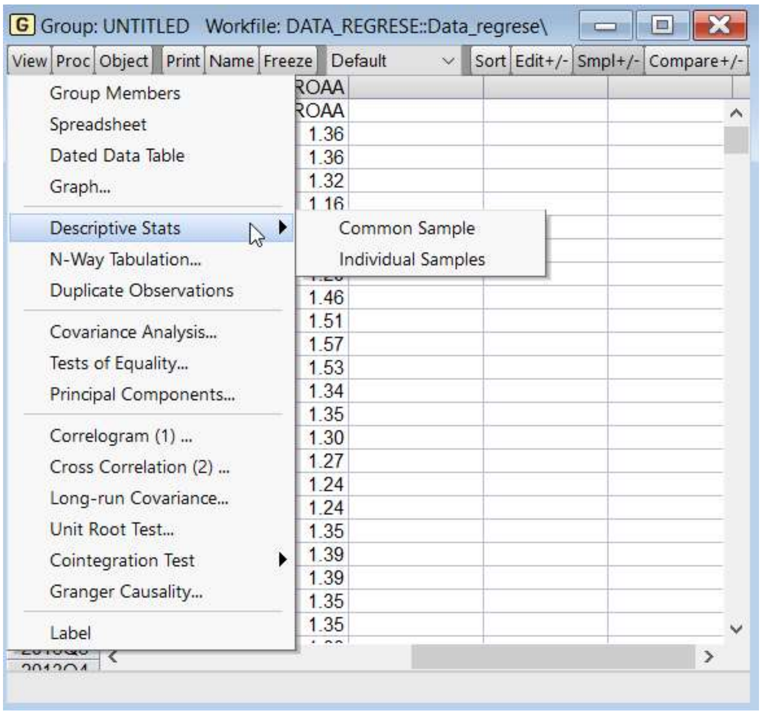
2.1.3 POPISNÉ STATISTIKY, DESKRIPTIVNÍ STATISTIKA

Grafy nám umožňují okamžitý a vizuální pohled na věc. V řadě případů je však důležitá i numerická formulace výsledků našich analýz. Deskriptivní statistika slouží k základnímu popisu dat a zjednodušení komplexity získaných dat. Deskriptivní statistika se tedy snaží několika čísly a obrázky stručně vystihnout podstatné informace o daných datech.

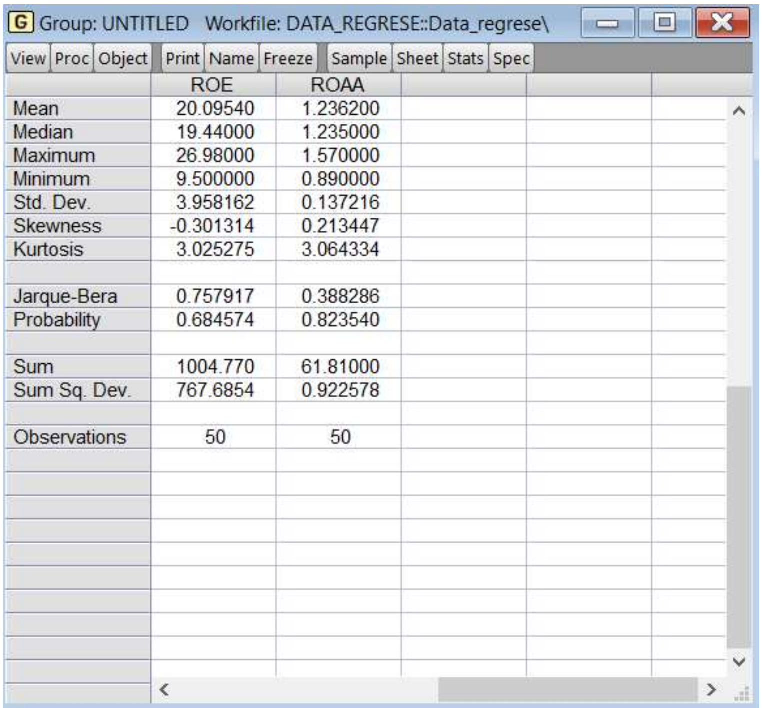
ŘEŠENÁ ÚLOHA

Deskriptivní statistiku můžeme použít k popsání časové řady nebo více časových řad. Pro popis např. dvou časových řad, nejprve otevřeme požadované časové řady jako skupinu.





Zda použijeme „Common Sample“ či „Individual Samles“, záleží na tom, zda chceme popsat data za období, které je shodné pro veškeré (obě) časové řady (Common Sample) či individuálně pro každou časovou řadu (Individual Sample). Je to nutné rozlišovat v případě, kdy máme různě dlouhé období pro dané časové řady. V případě, když máme shodné období pro všechny (obě) řady, jsou výsledky stejné.



	ROE	ROAA
Mean	20.09540	1.236200
Median	19.44000	1.235000
Maximum	26.98000	1.570000
Minimum	9.500000	0.890000
Std. Dev.	3.958162	0.137216
Skewness	-0.301314	0.213447
Kurtosis	3.025275	3.064334
Jarque-Bera	0.757917	0.388286
Probability	0.684574	0.823540
Sum	1004.770	61.81000
Sum Sq. Dev.	767.6854	0.922578
Observations	50	50

SAMOSTATNÝ ÚKOL

Stáhněte si libovolnou časovou řadu, která má alespoň 20 pozorování. V ekonometrickém software (např. EViews či STATA) vytvořte graf dané časové řady a deskriptivní statistiku.

2.1.4 Testování hypotézy

Asi nejdůležitější část, která nás bude provázet celou studijní oporou je testování nulového hypotézy. Na tomto principu je založena celá ekonometrie.

Teoretický rámec statistických testů vypadá tak, že v něm figurují dvě hypotézy:

- nulová hypotéza (H_0), kdy její podstatou je tvrzení, které má být testováno,
- alternativní hypotéza (H_1), která zahrnuje zbývající tvrzení.

Vlastní test hypotézy bývá většinou založen na statistickém porovnání odhadnutého parametru s jeho hypotetickou hodnotou z nulové hypotézy. Jestliže odhadnutý parametr se velmi liší od hypotetické hodnoty, pak nulovou hypotézu zamítneme. Jestliže naopak odhadnutý parametr se od hypotetické hodnoty liší jen málo, pak nulovou hypotézu pravděpodobně nelze zamítnout. Nezamítnutí nulové hypotézy však neznamená její automatické přijetí jako potvrzené, ale spíše nedostatečnou průkaznost pro její zamítnutí. Často se také z pozorovaných dat sestaví vhodný výraz (označovaný jako testová statistika), jehož pravděpodobnostní rozdělení jsme za platnosti nulové hypotézy schopni odvodit, a porovnává se s vhodnou konstantou (tato konstanta se nazývá kritickou hodnotou a nerovnost, na jejímž základě zamítáme nulovou hypotézu se nazývá kritickým oborem testu).

K ZAPAMATOVÁNÍ

Statistický test je možné provést jedním z následujících způsobů:

- pomocí kritického oboru,
- pomocí intervalu spolehlivosti,
- pomocí *p*-hodnoty.

1. Pomocí kritického oboru

Jedná se o klasický přístup k testování hypotéz. Kritický obor se volbou kritické hodnoty nastaví tak, aby tzv. chyba prvního druhu spočívající v zamítnutí nulové hypotézy, přestože nulová hypotéza platí, mohla nastat jen s předem svízelnou (malou) pravděpodobností α označovanou jako hladina významnosti tohoto testu (obvykle se volí $\alpha = 0{,}05$ nebo $\alpha = 0{,}01$, tj. pětiprocentní nebo jednoprocentní hladina významnosti. Např. v případě pětiprocentní hladiny významnosti neoprávněně zamítneme nulovou hypotézu jen v jedné z dvaceti opakovaných simulací vycházejících z našeho teoretického modelu). Chyba druhého druhu naopak spočívá v nezamítnutí nulové hypotézy, přestože nulová hypotéza neplatí, a tedy bychom ji měli zamítnout. Doplněk pravděpodobnosti chyby druhého durhu do jedné (tj. pravděpodobnost zamítnutí nulové hypotézy, když nulová hypotéza neplatí) se nazývá síla testu a kvalitní test by ji měl mít vysokou. Hladina významnosti a síla testu spolu obvykle souvisí. Tedy při redukci hladiny významnosti (např. z 5 % na 1 %) se sice zredukuje pravděpodobnost chyby prvního druhu, ale na druhé straně se také obvykle sníží šance toho, že test zamítne neplatnou nulovou hypotézu, jak by měl (tedy dojde k redukci jeho síly).

2. Pomocí interval spolehlivosti

Jedná se o přeformulování předchozího přístupu. V případě testu nulové hypotézy, že daný parametr nabývá dané hodnoty, se pro něj s využitím jeho odhadu zkonstruuje tzv. interval spolehlivosti. Např. 95 % interval spolehlivosti představuje číselný interval, v němž teoretická (tj. skutečná, ale pro nás neznámá) hodnota tohoto parametru leží s pravděpodobností (označovanou jako spolehlivost) ve výši 0,95. Podle toho, zda testovaná hodnota parametru leží, resp. neleží v tomto intervalu, se nezamítne, resp. zamítne nulová hypotéza, že parametr nabývá testované hodnoty, a to na hladině významnosti, která je doplňkem použité spolehlivosti do jedné. Jestliže tedy např. testovaná hodnota leží vně příslušného 95 % intervalu spolehlivosti, pak testovanou hodnotu pro uvažovaný parametr zamítneme na hladině významnosti 5 %.

3. Pomocí *p*-hodnoty

Tento přístup bývá používán nejčastěji k testování hypotéz. Výstupem tohoto přístupu je dosažená hladina významnosti označovaná jako *p*-hodnota. Jedná se o relativně jednoduchou techniku, kdy není nutné pracovat s pevnou hladinou významnosti (spolehlivosti), protože *p*-hodnota je maximální hladina významnosti, při které bychom ještě příslušnou nulovou hypotézu nezamítly (resp. minimální hladina významnosti, při které bychom tuto nulovou hypotézu zamítli). Také *p*-hodnota se označuje jako přijatelnost nulové hypotézy, protože čím je *p*-hodnota nižší, tím je nulová hypotéza méně přijatelná.

Pokud si stanovíme hladinu významnosti 5 %, tak zjednodušeně můžeme použít toto pravidlo:

- p-hodnota $< 0{,}05 \rightarrow$ zamítáme nulovou hypotézu,
- p-hodnota $> 0{,}05 \rightarrow$ nelze zamítnout nulovou hypotézu.

K ZAPAMATOVÁNÍ

Stručně lze tyto tři postupy shrnout tak, že nulovou hypotézu zamítáme, když 1) testová statistika padne do kritického oboru nebo 2) hypotetická testovaná hodnota není překryta intervalem spolehlivosti nebo 3) $p \leq \alpha$.

Navíc je nutno rozlišovat mezi statistickou a praktickou významností a testovanou hypotézu posuzovat i z jiných hledisek. Proto se nedoporučuje vždy používat standard pětiprocentní hladiny významnosti, protože pak můžeme zamítnou každou hypotézu, pokud budeme mít dostatečně rozsáhlý soubor.

Skutečnost	**Rozhodnutí**	
	nezamítáme H	**zamítáme H**
Hypotéza H platí	správné rozhodnutí pravděpodobnost = 1-a	chyba I. druhu pravděpodobnost = a
Platí alternativa A	chyba II. druhu pravděpodobnost = b	správné rozhodnutí pravděpodobnost = 1-b (síla testu)

Tabulka 2-1 Možné výsledky testování hypotézy

2.1.5 STACIONARITA

Stacionarita časové řady naznačuje, že chování dané řady je stochasticky stále. Stacionární časová řada je taková řada, jejíž základní vlastnosti, například její průměr a rozptyl se nemění v čase. Na rozdíl od nestacionárních časových řad, u kterých se tyto vlastnosti v průběhu času mění. V teorii se rozlišují dva druhy stacionarity:

- striktní,
- slabá.

Striktní stacionarita předpokládá, že chování příslušného náhodného procesu, tj. jeho rozdělení, je invariantní vůči časovým posunům. Naproti tomu slabá stacionarita je méně omezující; požaduje se, aby příslušný náhodný proces měl konstantní střední hodnotu, konstantní rozptyl a aby závislost mezi dvěma libovolnými pozorováními závisela jen na jejich časové vzdálenosti a nikoli na jejich časovém umístění v řadě.

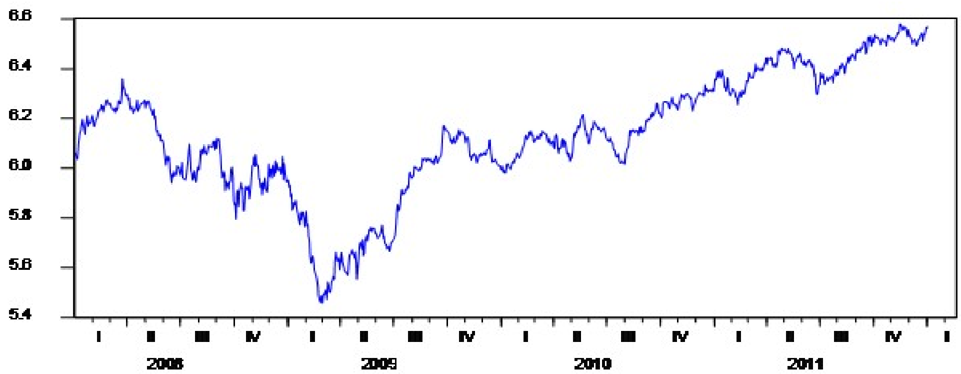
Můžeme tedy říci, že stacionarita časové řady znamená, že chování této řady je v jistém smyslu stochasticky ustálené. Stacionarita časové řady se dá vymezit tak, že chování této řady je v jistém smyslu stochasticky ustálené. Úroveň a rozptyl stacionární řady jsou konstantní v čase a také kovarianční struktura takové řady musí být v čase neměnná. Trend, sezónnost nebo proměnný rozptyl nejsou slučitelné se stacionaritou a tyto jevy musí být z řad odstraněné.

Stacionarita časové řady se testuje pomocí testu jednotkového kořene (unit root test).

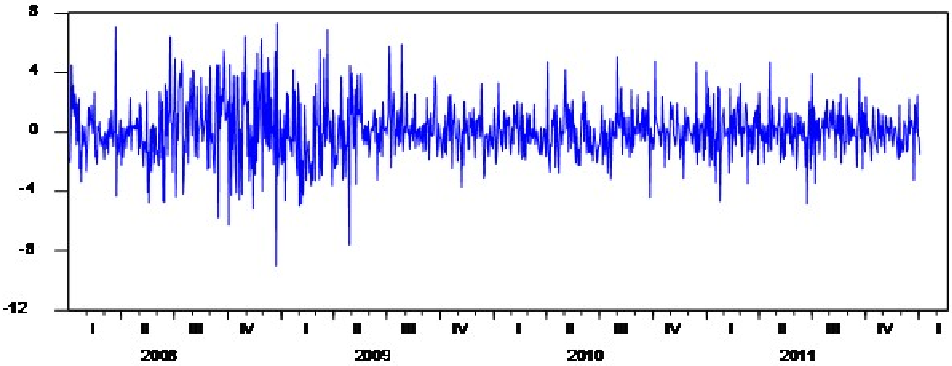
Možnost stacionarizace diferencováním říká o přítomnosti jednotkového kořenu v autoregresním operátoru příslušného modelu. Rozhodnutí o přítomnosti jednotkového kořenu (či vícenásobného jednotkového kořenu) je často klíčovým bodem analýzy. Přítomnost jednotkového kořene může být testována pomocí testu na jednotkový kořen. Další možnost testování slabé stacionarity může poskytnout korelogram autokorelační funkce. Běžně se v praxi k testování stacionarity používají testy jednotkového kořene.

Existuje řada testů stacionarity časových řad, jako příklad si uvedeme následující testy:

- Dickeyův-Fullerův test,
- Rozšířený Dickeyův-Fullerův (ADF) test,
- Phillipsův-Perronův test,
- KPSS test.



Obrázek 2-1 Příklad nestacionární časové řady



Obrázek 2-2 Příklad stacionární časové řady

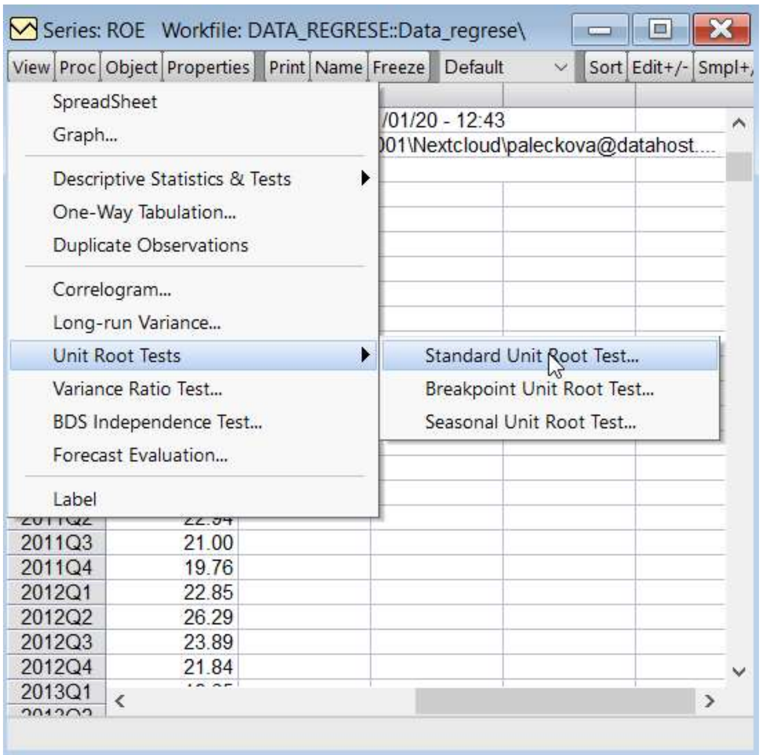
PŘEVOD NESTACIONÁRNÍ ČASOVÉ ŘADY NA STACIONÁRNÍ

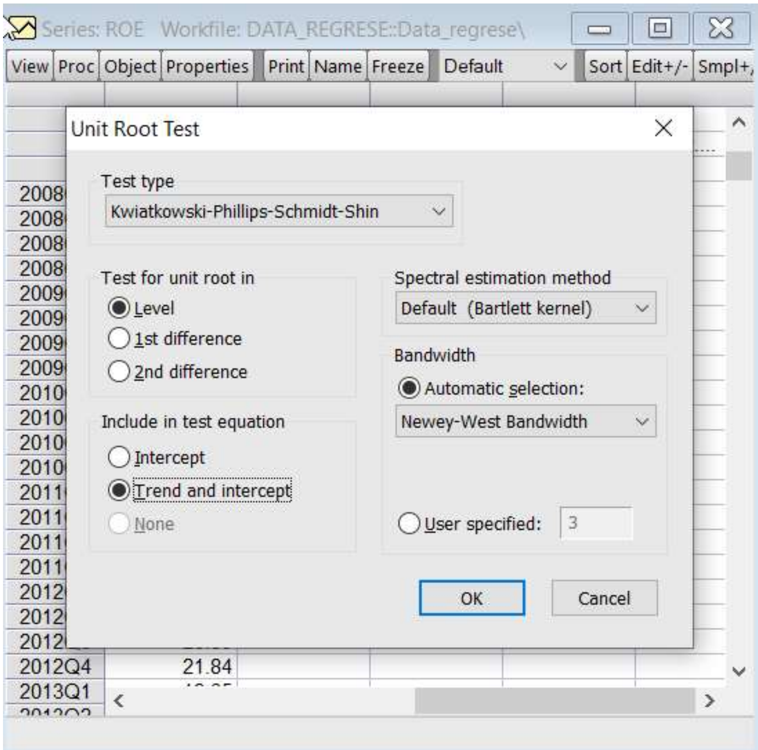
Diferencováním nestacionární časové řady zpravidla vznikne stacionární. Proměnná se nazývá integrovaná řádu *d*, značeno *I(d)*, jestliže musí být diferencovaná *d* krát, aby se stala stacionární. Tudíž stacionární proměnná je integrovaná řádu 0, což se zapisuje jako *I*(0). Nejsou-li časové řady *I*(0), tj. integrovány nultého řádu, dosáhneme stacionárnosti přechodem na jejich logaritmy, resp. použitím 1. diference. Proměnná, která je diferencovaná jednou, aby se stala stacionární, je integrovaná řádu jedna, tedy *I*(1). Nestacionární proměnné jsou obvykle *I*(1), zřídka kdy jsou integrovány vyššího řádu než dva. Tedy nejčastější forma, jak převést nestacionární časovou řadu na stacionární je pomocí diferencování časové řady. Pozor však na interpretaci výsledného modelu.

Případy, které mohou nastat při práci s nestacionárními proměnnými jsou např. zdánlivá regrese a kointegrace. Koitegraci se budeme podrobněji věnovat v 5. kapitole. Pokud bychom použili nestacionární data v modelu lineární regrese, může se jednat o zdánlivou regresi. Zdánlivá regrese má vysoký koeficient determinace R^2 a *t*-statistika se jeví být jako významná, avšak výsledky nejsou ekonomicky významné. Regresní výstup se jeví jako dobrý, protože odhady metody nejmenších čtverců jsou nekonzistentní. Regresní analýze se věnuje podrobněji následující část kapitoly.

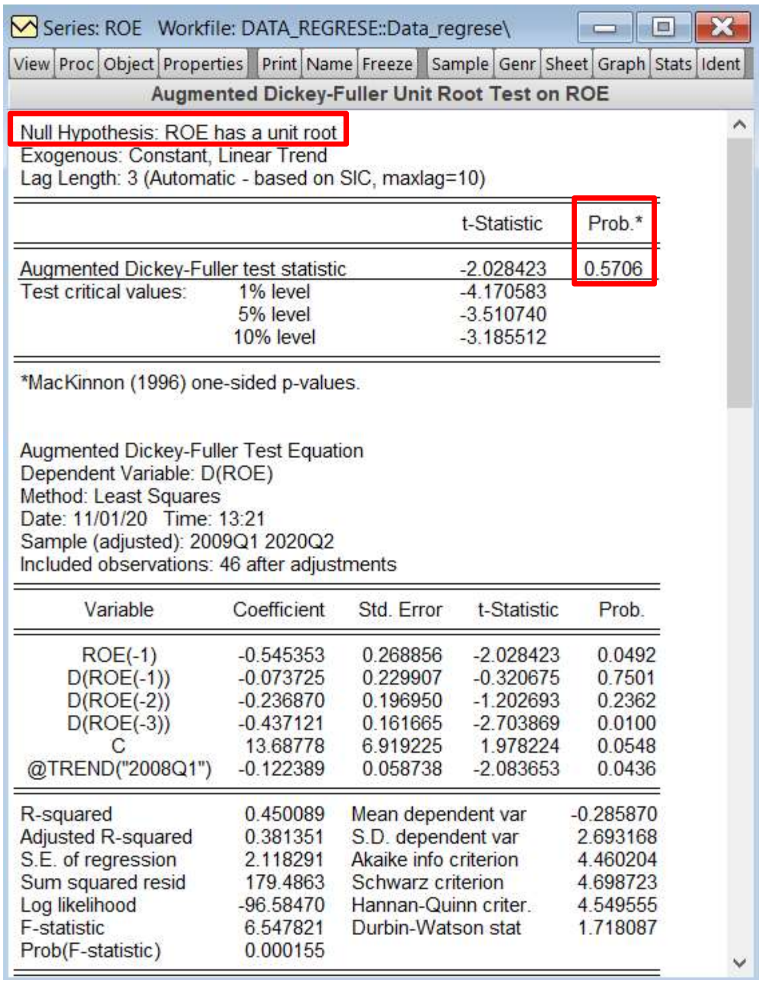
ŘEŠENÁ ÚLOHA

Nejprve otevřeme v EViews časovou řadu, u které chceme zjišťovat stacionaritu.





Nejprve vybereme typ testu (např. Augmented Dickey-Fuller, Phillips-Perron, Kwiatkowski-Phillips-Schmidt-Shin). Vybíráme Augmented Dickey-Fuller (ADF) test. Poté vybeme, zda chceme testovat data na „Levelu“ či na 1. diferenci nebo na 2. diferenci. Protože chceme testovat původní data, zvolíme „Level“. V posledním kroku volíme, co chceme do testu zahrnout. Zde není vhodné použít „None“, proto volíme mezi „Intercept“ (jedná se o méně striktní test) nebo „Trend and intercept“ (striktnější test). Nastavení zpoždění je vhodné ponechat dle doporučení EViews (pravou stranu tedy nemusíme upravovat).



Series: ROE Workfile: DATA_REGRESE::Data_regrese\

View | Proc | Object | Properties | Print | Name | Freeze | Sample | Genr | Sheet | Graph | Stats | Ident

Augmented Dickey-Fuller Unit Root Test on ROE

Null Hypothesis: ROE has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 3 (Automatic - based on SIC, maxlag=10)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-2.028423	0.5706
Test critical values:	1% level	-4.170583	
	5% level	-3.510740	
	10% level	-3.185512	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(ROE)
Method: Least Squares
Date: 11/01/20 Time: 13:21
Sample (adjusted): 2009Q1 2020Q2
Included observations: 46 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
ROE(-1)	-0.545353	0.268856	-2.028423	0.0492
D(ROE(-1))	-0.073725	0.229907	-0.320675	0.7501
D(ROE(-2))	-0.236870	0.196950	-1.202693	0.2362
D(ROE(-3))	-0.437121	0.161665	-2.703869	0.0100
C	13.68778	6.919225	1.978224	0.0548
@TREND("2008Q1")	-0.122389	0.058738	-2.083653	0.0436

R-squared	0.450089	Mean dependent var	-0.285870
Adjusted R-squared	0.381351	S.D. dependent var	2.693168
S.E. of regression	2.118291	Akaike info criterion	4.460204
Sum squared resid	179.4863	Schwarz criterion	4.698723
Log likelihood	-96.58470	Hannan-Quinn criter.	4.549555
F-statistic	6.547821	Durbin-Watson stat	1.718087
Prob(F-statistic)	0.000155		

V prvním řádku je nulová hypotéza: Časová řada má jednotkový kořen. Můžeme využít kritickou hodnotu nebo testujeme nulovou hypotézu pomocí p-hodnoty. Vidíme zde, že nelze zamítnout nulovou hypotézu a řada není stacionární.

2.2 Regresní analýza a estimace parametrů modelu

PRŮVODCE STUDIEM

Regrese je nejdůležitějším nástrojem využívaným v aplikované ekonomii pro analýzu a pochopení vztahu mezi dvěma a více proměnnými.

Nejprve si popíšeme problematiku jednoduchého regresního modelu, tedy modelu s jednou vysvětlující proměnnou. Přestože se v praxi setkáváme s případy použití více proměnných, je nutné si nejprve vysvětlit model s jednou proměnnou a tento model bude sloužit jako základ. Koncept vícenásobné regrese, tedy model, kdy analyzujeme více než dvě proměnné, je pak analogickým a jednoduchým rozšířením jednoduchého regresního modelu.

Lineární regresní analýza využívá pro popis závislostí veličin funkce lineární v parametrech, případně takové funkce, které lze na lineární převést pomocí vhodné transformace. Nelineární regresní analýza pak využívá pro popis závislostí veličin funkce nelineární v parametrech. V rámci studijní opory bude dále popsaná pouze lineární regresní analýza.

Jednoduchá regresní analýza studuje zavilost dvou veličin, naopak vícenásobná regresní analýza studuje závislost jedné veličiny na dvou a více dalších veličinách.

2.2.1 Jednoduchý lineární regresní model

Nejjednodušší je model, který nazýváme jednoduchý lineární regresní model. Regresní model se nazývá lineární, když je lineární ve svých parametrech. Není důležité, jaký je tvar zápisu nezávisle proměnné. Lineární regrese znamená, že funkčním vztahem mezi dvěma veličinami má být přímka. Linearita modelu je důležitá z hlediska praktických výpočtů, jejichž účelem je odhadnout hodnoty parametrů. Pokud se jedná o lineární model, je to mnohem snadnější než u modelu nelineárního, kdy někdy ani odhady parametrů vypočítat není možné. Ve studijním textu se budeme dále věnovat pouze lineárnímu regresnímu modelu.

Cílem jednoduchého regresního modelu je najít model funkční závislosti (spojité) veličiny) Y na jedné (spojité) veličině X (tedy regresoru). Model jednoduché lineární regrese lze zapsat takto:

$$Y = \beta_0 + \beta_1 X + u \tag{1}$$

kde Y představuje vysvětlovanou proměnnou a X je vysvětlující proměnná, β_0 a β_1 jsou parametry modelu a u je náhodná složka.

Vztah mezi závislou a nezávislou veličinou se v rámci regresní analýzy snažíme popsat nějakým funkčním vztahem. Teoretické hodnoty parametrů, které se ve vyjádření závislosti vyskytují, značíme malými řeckými písmeny. Jejich přesnou hodnotu nevíme, ale pokusíme se ji odhadnout. Tedy můžeme říct, že hledáním regresního modelu, resp. regresní funkce rozumíme hledání regresního koeficientu β_1, přičemž typ regresního modelu musíme stanovit sami (na základě zkušeností a vzhledu bodových grafů), hodnoty jednotlivých regresních koeficientů pak nalezne estimační metoda. Tyto metody budou popsány dále, v praxi je nejvíce využívaná metoda nejmenších čtverců, které se budeme podrobně věnovat.

2.2.2 Vícenásobný lineární regresní model

Pokud lineární regresní model rozšíříme o více vysvětlujících proměnných, hovoříme o vícenásobném lineárním regresním modelu. Můžeme tedy zkoumat, jak několik faktorů ovlivňuje určitou veličinu.

Lineární model může být formálně zapsán takto:

$$Y_t = \beta_1 + \beta_2 X_{t1} + \beta_3 X_{t2} + \cdots + \beta_k X_{tk} + u_t \tag{2}$$

Kde Y_t představuje vysvětlované proměnné Y v čase t, X_{t1} až X_{tk} představují hodnoty vysvětlujících proměnných X_1 až X_k pozorovaných v čase t, β_1 až β_k jsou neznámé parametry modelu, kde β_1 představuje tzv. absolutní člen, β_2 až β_k pak změnu závislé proměnné Y_t, danou změnou nezávislé proměnné X_i o jednotku za předpokladu, že ostatní nezávislé proměnné jsou neměnné (*ceteris paribus*) a ut je náhodná neboli také reziduální složka, která zahrnuje:

- souhrn vlivů, explicitně uvedených v modelu,
- chyby měření ekonomických a finančních veličin,
- nekorektní volbu regresního vztahu,
- některé jevy, které mají náhodný charakter.

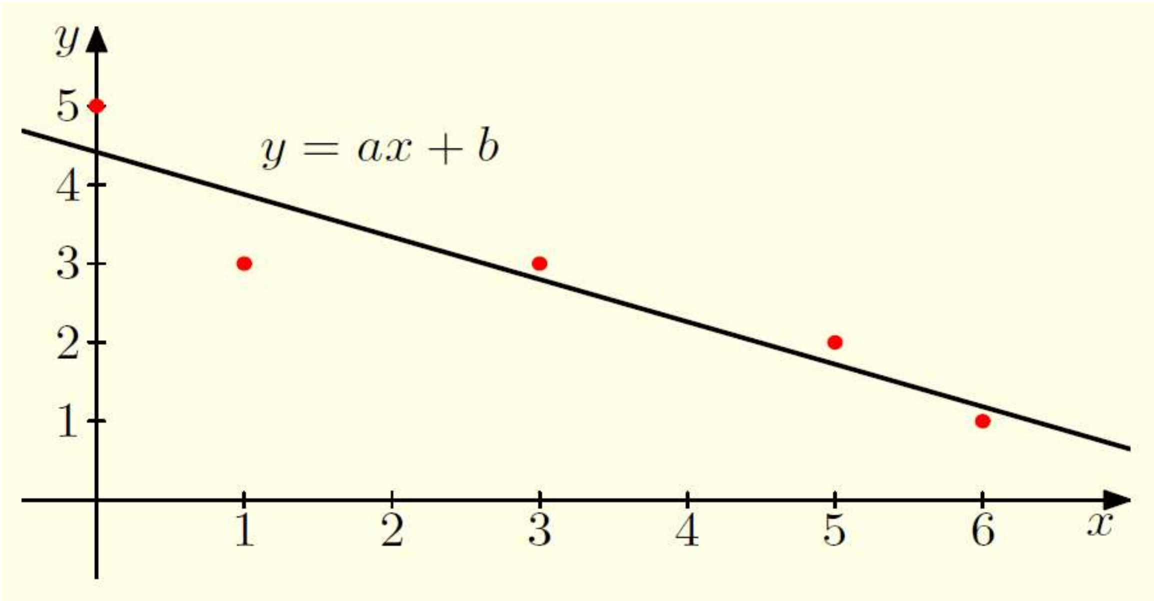
K odhadu parametrů modelu slouží několik metod. Mezi nejvíce využívané modely patří metoda nejmenších čtverců, metoda maximální věrohodnosti nebo momentová metody. V poslední zmiňované metodě se nejčastěji v praxi využívá zobecněná metoda momentů. Nyní si metody blíže specifikujeme, přičemž největší pozornost budeme věnovat metodě nejmenších čtverců. Tato metoda je základní metodou a má v praxi široké využití.

2.2.3 Metoda nejmenších čtverců

Tato metoda představuje nejčastější přístup ke stanovení odhadu parametrů lineárního regresního modelu. Předpokládejme, že mezi veličinami x a y je lineární vztah ve tvaru y = ax + b. Měřením byly pro konkrétní hodnoty veličiny x naměřeny odpovídající hodnoty veličiny y a výsledek byl zanesen do grafu. Body však neleží na jedné přímce, protože měření je vždy zatíženo nějakou chybou a teorie navíc vždy nemusí odpovídat praxi stoprocentně. Máme tedy body v rovině, které leží přibližně v jedné přímce a chceme najít co nejpřesnější matematický model, tj. stanovit koeficienty a, b tak, aby přímka y = ax + b ležela co nejblíže bodům z měření. Snažíme se vystihnout chování bodu pomocí lineární závislosti. Přímka nebude pochopitelně procházet všemi body, chceme tedy alespoň, aby procházela co nejblíže okolo nich. Za optimální přímku považujeme tu, která minimalizuje součet ploch čtverců.

K ZAPAMATOVÁNÍ

Název metoda nejmenších čtverců (ordinary least squares, OLS) se vžil díky grafickému názoru, že totiž opravdu hledáme takovou přímku, pro kterou je součet ploch čtverců odchylek minimální.



Obrázek 2-3 Předpoklady metody nejmenších čtverců

Metoda nejmenších čtverců (ordinary least squares, OLS) je základní a nejpoužívanější metodou odhadu parametrů lineárního regresního modelu. Při metodě nejmenších čtverců se požaduje, aby součet čtverců (druhých mocnin) rozdílů naměřených hodnot yi a funkce f(xi) byl co nejmenší. Odhady parametrů β jsou hledány v regresní rovnici tím, že vzhledem k těmto parametrům je minimalizován součet čtverců. Odvození vlastnosti odhadu OLS je možné jen v případě, že model splňuje určité předpoklady.

Pro použití metody nejmenších čtverců za účelem odhadu regresních parametrů jsou ověřovány následující předpoklady:

- lineární regresní model je lineární v parametrech,
- Xi není stochastická veličina,
- střední hodnota náhodné složky je nulová,
- konstantní rozptyl náhodné složky,
- náhodná složka není sériově závislá (nevyskytuje se autokorelace),
- nulová kovariance mezi náhodnou složkou a vysvětlující proměnnou,
- počet pozorování musí být větší, než počet parametrů regresního modelu,
- regresní model je správně specifikován,

- náhodná složka má normální rozdělení.

Předpoklady charakterizující klasický model lineární regrese jsou často uváděny v následujícím tvaru:

$$E(\varepsilon_t) = 0, \tag{3}$$

$$\mathrm{var}(\varepsilon_t) = \sigma^2 < \infty, \tag{4}$$

$$\mathrm{cov}(\varepsilon_t, \varepsilon_s) = 0 \text{ pro } t \neq s, \tag{5}$$

$$\mathrm{cov}(x_t, \varepsilon_t) = 0, \tag{6}$$

$$\varepsilon_t \sim \mathrm{N}(0, \sigma^2). \tag{7}$$

Výše uvedené předpoklady metody nejmenších čtverců lze tedy definovat následovně:

- 1. Střední hodnota reziduální složky je nulová pro všechna t.
- 2. Rozptyl reziduální složky je konstantní a konečný pro všechna t.
- Předpoklad konstantního rozptylu reziduálních složek se někdy verbálně označuje jako homoskedasticita.
- Jeho porušení (heteroskedasticita) je v ekonometrii poměrně časté a vyžaduje použití speciálních postupů (např. modelů typu GARCH pro podmíněnou heteroskedasticitu).
- 3. Reziduální složky jsou navzájem nekorelované pro všechna s ≠ t.
- Předpoklad vzájemně nekorelovaných reziduálních složek v ekonometrické praxi často neplatí, což si opět vyžádá použití speciálních postupů (tzv. autokorelace reziduí testovaná např. pomocí Durbinova-Watsonova testu).
- 4. Regresory jsou ve stejném čase nebo pro stejnou průřezovou jednotku nekorelované s reziduální složkou pro všechna i a t.
- 5. Reziduální složky mají normální rozdělení, neboli předpoklad normality.

Vlastnosti odhadované funkce nejmenších čtverců jsou následující. Odhad parametrů je:

- nestranný, jestliže hodnota odhadovaného parametru je rovna jeho střední hodnotě (v opačném případě je odhad vychýlený),
- vydatný (eficientní), jestliže nemá větší rozptyl jako odhad téhož parametru,
- konzistentní, jeslitže při rostoucím rozsahu výběru její výběrové rozdělení konverguje ke skutečné hodnotě odhadovaného parametru.

Koeficient determinace

Následně po provedení modelu lineární regrese je potřeba posoudit, zda je kompatibilní s použitými daty. Můžeme to provést sofistikovaně pomocí statistických testů, nebo také orientačně na základě tzv. koeficientu determinace. Tento koeficient značíme jako R^2. Základní myšlenkou tohoto koeficientu je rozložení úplného součtu čtverců pozorované a střední hodnoty Y na její reziduální část a na část vysvětlovanou regresí, kdy se zjišťují relativní podíly (Hančlová, 2012).

Koeficient determinace vyjadřuje stupeň vysvětlení celkové změny vysvětlované proměnné Y regresí neboli prostřednictvím působení lineárního vztahu vysvětlující proměnné. Můžeme říct, že jde o kritérium, které měří pomocí regresní přímky shodu napozorovaných dat a odhadů. Koeficient determinace nám tedy říká, jaký podíl variability vysvětlované proměnné jsme podchytili prvním faktorem – vysvětlující proměnnou.

Vlastnosti koeficientu determinace jsou následující:

- $0 \leq R^2 \leq 1$, pak se koeficient determinace pohybuje mezi nulou a jedničkou,
- $R^2=1$, pak všechna výběrová pozorování leží přímo na vyrovnané regresní přímce,
- $R^2=0$, znamená, že se ani jedno pozorování nenachází na regresní přímce. To znamená, že veškeré informace jsou nevysvětleny a jsou součástí reziduální části. Odhadnutý regresní model nemá smysl (Hančlová, 2012).

Koeficient determinace může nabývat hodnot od nuly do jedné. Je-li nulový, pak je chování vysvětlované proměnné v podstatě náhodné, nezávislé na vysvětlující proměnné. Je-li roven jedné, pak náhoda naopak žádnou roli nehraje – hodnoty vysvětlované proměnné jsou naprosto přesně určeny hodnotami proměnné vysvětlující. Čím vyšší je koeficient determinace, tím více se body představující jednotlivá pozorování pohybují poblíž regresní přímky.

Koeficient determinace má své nedostatky a to takové, že adekvátně nereaguje na změny počtu pozorování v regresním modelu a nebere v potaz rozšíření počtu vysvětlujících proměnných v daném regresním modelu (Hančlová, 2012).

Multikolinearita

Silné závislosti mezi vysvětlujícími proměnnými, tedy vysoká vzájemná korelovanost regresorů. Lineární závislost sloupců regresní matice X. Příznakem multikolinearity je vysoká hodnota korelačního koeficientu mezi regresory. Pozor – korelovanost mezi vysvětlovanou proměnnou a regresorem se již v žádném případě za multikolinearitu nepovažuje.

$$\underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{bmatrix}}_{\mathbf{y}} = \underbrace{\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1j} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2j} & \cdots & x_{2m} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{i1} & x_{i2} & \cdots & x_{ij} & \cdots & x_{im} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nj} & \cdots & x_{nm} \end{bmatrix}}_{\mathbf{X}} \cdot \underbrace{\begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_j \\ \vdots \\ \beta_m \end{bmatrix}}_{\boldsymbol{\beta}} + \underbrace{\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_i \\ \vdots \\ \varepsilon_n \end{bmatrix}}_{\boldsymbol{\varepsilon}}$$

Vektory matice X musí být **skutečně navzájem nezávislé** (jejich párové R musí být nulové nebo statisticky nevýznamné). Pokud tomu tak není, dochází k **multikolinearitě**, která způsobuje početní i statistické problémy.

Perfektní kolinearita znamená, že sloupce v matici vysvětlujících proměnných jsou perfektně lineárně závislé. Pak bude determinant této matice nulový a nebude ji možné invertovat, a tedy ani použít MNČ. Pokud jsou některé sloupce „jen silně lineárně závislé,“ jde o multikolinearitu. Determinant matice vysvětlujících proměnných tak sice bude také blízký nule, ale inverzní matici spočítat dokážeme. To může být ale také problém. Odhady sice budou nestranné i vydatné, ale budou velmi nestabilní (tzn. budou citlivé i na malé změny v matici vysvětlujících proměnných). Tedy při jiném výběru z téže populace se odhady mohou velmi lišit. Druhým problémem jsou velké směrodatné chyby odhadnutých parametrů, takže některé parametry se mohou jevit jako statisticky nevýznamné, a pak nevíme, jestli některé proměnné vyřadit a případně které. Příčinou může být to, že kvůli silné vzájemné závislosti proměnných často nedokážeme určit vliv jedné konkrétní proměnné na vysvětlovanou proměnnou, proto se může jevit jako nevýznamná. Dokážeme určit jen společný vliv všech proměnných.

Příčinou multikolinearity je:

- podobný vývoj makroekonomických veličin (při analýze časových řad);
- souvislost veličin (v průřezové analýze), např. kdybychom vysvětlovali výdaje domácnosti výší příjmu a výší bohatství (tyto dvě proměnné souvisí), nebo objem výstupu počtem pracovníků a hodnotou fixního kapitálu;
- zahrnutí zpožděných hodnot, ty bývají totiž zkorelovány;
- zahrnutí nesprávného počtu umělých proměnných (pak se vyskytne perfektní kolinearita).

Jak postupovat v případě, když je v modelu multikolinearita:

- zvětšit rozsah výběru, pokud to jde, například zahrnout i extrémní pozorování (nové hodnoty se musí lišit od těch v původním výběru);
- přidat dodatečnou apriorní informaci (apriorní omezení), díky kterému pak snížíme směrodatné chyby odhadových funkcí (např. omezení parametrů produkčních či poptávkových funkcí atd.);
- použít tzv. „smíšený odhad modelu“ –zkombinovat průřezová data a údaje z časových řad;
- změnit specifikaci, tedy vypustit některé proměnné, které se jeví jako nevýznamné, ale pozor na to, že tím může vzniknout specifikační chyba;
- transformovat původní proměnné (například pomocí prvních diferencí nebo podílů, ale často tak do modelu zase dostaneme autokorelaci nebo heteroskedasticitu);
- použít metody vícerozměrné statistiky (metodu hlavních komponent, faktorovu analýzu, hřebenovou regresi).

První tři řešení bychom měli preferovat, protože měnit specifikaci modelu může být riskantní a nesprávná specifikace modelu má horší dopady než multikolinearita. Ta v podstatě ani moc nevadí, pokud chceme model použít k prognózování, v případě, že vztahy mezi proměnnými se nebudou nějak dramaticky měnit a proměnné budou i nadále lineárně závislé.

Heteroskedasticita

Problém heteroskedasticity vzniká v případě porušení předpokladu homoskedasticity. Tedy jestliže reziduální složky nemají konstantní rozptyl (tj. jestliže množství náhodnosti obsažené ve výstupu y_t může být pro každé pozorování různé), pak se označují jako heteroskedastické. Reziduální složky ε_t nemají konstantní rozptyl $\sigma^2 k_t$ s neznámými kladnými hodnotami k_t a jsou vzájemně nekorelované.

Heteroskedasticita znamená, že pro různé hodnoty vysvětlujících proměnných má náhodná složka, a tedy i vysvětlovaná proměnná, jiný rozptyl. Setkáme se s ní většinou:

- v průřezových datech, zejm. mikroekonomických analýzách (příklad: výdaje domácností budou mít větší rozptyl u bohatších domácností než u chudších; objem produkce firmy bude mít větší rozptyl v případě větších než menších firem)
- při chybné specifikaci modelu (jde o tzv. „kvaziheteroskedasticitu“), například pokud vynecháme významnou proměnnou;
- kvůli chybám měření (s rostoucí hodnotou vysvětlované proměnné se kumulují chyby a zvětšuje se rozptyl reziduí);
- při použití skupinových průměrů či jiných agregovaných údajů.

Důvody, proč nepoužít model s heteroskedasticitou:

- odhad pomocí MNČ bude sice nestranný a konzistentní, ale nebude vydatný ani asymptoticky vydatný;

- odhad směrodatných chyb parametrů a odhad rozptylu modelu bude vychýlený, takže se nelze spolehnout na zkonstruované intervaly spolehlivosti ani závěry testů hypotéz.

Testování heteroskedasticity:

- Spermanův test korelace pořadí zjišťuje, zda se v modelu vyskytuje lineární závislost směrodatné odchylky náhodné složky na vysvětlující proměnné.
- Golfeldův-Quandtův test je použitelný tehdy, když je směrodatná odchylka monotónní funkcí některé vysvětlující proměnné.
- Glejserův test postupuje tak, že odhaduje závislost absolutní hodnoty reziduí na vysvětlující proměnné.
- Parkův test je podobný jako Glejserův test: také sestavuje pomocnou regresi
- Whiteův test je obecnější, protože není potřeba vycházet z žádného předpokladu o tvaru heteroskedasticity ani z předpokladu závislosti rozptylu reziduí jen na jedné proměnné.
- Breusch-Paganův test sestavuje pomocnou regresi čtverců reziduí původního modelu na vysvětlujících proměnných (některých či všech) či jejich funkcích (třeba druhých mocninách).

AUTOKORELACE NÁHODNÉ SLOŽKY

Autokorelace je chápána jako závislost mezi posloupností hodnot jedné proměnné, uspořádaných v čase, někdy v prostoru. Většina údajů časových řad ekonomických a finančních proměnných vykazuje značnou setrvačnost, takže jejich pozorování za několik po sobě jdoucích období nejsou závislá, ale sériově zkorelovaná. Autokorelaci lze potvrdit pomocí Durbin-Watsonova testu.

Autokorelace se nejčastěji vyskytuje:

- v časových řadách: hodnoty různých ekonomických veličin apod. vykazují obvykle pozitivní autokorelaci kvůli setrvačnosti, takže nezahrnutí zpožděných proměnných do modelu se projeví autokorelací náhodné složky (kvaziautokorelace);
- chybná specifikace (např. lineární funkce místo kvadratické) se též může projevit autokorelací;
- chyby měření;
- zprůměrovaná, extrapolovaná či jinak upravená data;
- zahrnutí zpožděných endogenních či exogenních proměnných do modelu (náhodné složky pak mohou být sériově závislé)

Autokorelace způsobuje obdobné problémy jako heteroskedasticita: odhady parametrů nejsou vydatné a standardní chyby a odhad rozptylu modelu jsou vychýlené. Při pozitivní autokorelaci jsou obvykle odhady rozptylu podhodnocené.

Důsledky použití MNČ v případě autokorelace náhodné složky jsou obdobné heteroskedasticitě, tedy:

- Odhady parametrů MNČ zůstávají nestranné a konzistentní, ale nemají minimální rozptyl (nejsou vydatné) ani asymptoticky vydatné.
- 2) Odhadnuté rozptyly náhodných složek a standardní chyby (zpravidla podhodnocené při pozitivní autokorelaci) jsou vychýlené, špatně určené intervaly spolehlivosti, nelze použít běžné testy (výsledky klasických testů jsou nereálné).

Platí: $|\rho| < 1$, protože jinak by náhodné složky měly explozivní charakter (nebyla by homoskedasticita a bylo by to nereálné)

Jestliže:

- $\rho > 0$ pozitivní autokorelace 1.řádu – posloupnost několika kladných složek se střídá s posloupností několika záporných náhodných složek
- $\rho < 0$ negativní autokorelace 1. řádu – po sobě jdoucí složky pravidelně střídají znaménka
- $\rho = 0$ sériová nezávislost
- $\rho = 1$silná kladná (pozitivní) autokorelace
- $\rho = -1$ silná záporná (negativní) autokorelace

Testování autokorelace 1. řádu

Test reziduí pomocí Durbinovy-Watsonovy statistiky (DW) po aplikaci MNČ. Podíl součtu čtverců rozdílů sousedních reziduí a reziduálního součtu čtverců, platí E(d) = 2. Rezidua mohou být kladná i záporná.

DW statistika testuje nulovou hypotézu $H_0: \rho = 0$ a alterativní hypotézu $H_1: \rho \neq 0$. V případě nulové hypotézy reziduály v čase $t - 1$ jsou na sobě nezávislé. V případě zamítnutí nulové hypotézy, lze konstatovat, že existuje důkaz o vztahu mezi reziduály. DW kritérium má tvar:

$$DW = \frac{\sum_{t=2}^{T}(\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=2}^{T}{\varepsilon_t}^2}, \tag{8}$$

kde ε_t jsou reziduály a může nabýt hodnot z intervalu ⟨0, 4⟩. Jednoduchá algebra ukazuje, že:

$$DW \approx 2 - 2\hat{\rho}, \tag{9}$$

kde označení aproximace je vzhledem k malým rozdílům v pozorováních, přes které jsou brány sumarizace.

Mohou nastat následující situace:

- Úplná pozitivní autokorelace – sousední hodnoty reziduí mají stejná znaménka, jejich rozdíl je nulový, taktéž jejich čtverce, a tudíž $d = 0$ a $\rho = 1$,
- Pozitivní autokorelace – sousední hodnoty reziduí mají stejná znaménka, jejich rozdíl je malý, taktéž jejich čtverce ve srovnání se čtverci reziduí, a tudíž $d \to 0$a $\rho \to 1$, obecně $\rho \geq 0$. Platí, že $0 \leq d \leq dL$
- Perfektní nezávislost – sousední hodnoty reziduí střídají náhodně znaménka, při grafickém zobrazení nevidíme žádný vzor ani pravidla. Platíd=2a ρ=0.
- Nezávislost – sousední hodnoty reziduí střídají náhodně znaménka, při grafickém zobrazení nevidíme žádný vzor ani pravidla. Platí $d \to 2$a $\rho \to 0$, a tudíž $dU \leq d \leq 4 - dU$.
- Negativní autokorelace – sousední hodnoty reziduí mají opačná znaménka, jejich rozdíl je velký (téměř dvojnásobný ve srovnání s hodnotami reziduí), taktéž jejich čtverce (čtyřnásobné) ve srovnání se čtverci reziduí, a tudíž $d \to 4$a $\rho \to -1$, obecně $\rho \leq 0$. Platí, že $4 - dL \leq d \leq 4$.
- Úplná negativní autokorelace – sousední hodnoty reziduí mají opačná znaménka, jejich rozdíl je dvojnásobný ve srovnání s hodnotami reziduí, jejich čtverce jsou pak čtyřnásobné ve srovnání se čtverci reziduí, a tudíž d=4 a $\rho = -1$.

DW	Výsledek
$4 - d_L < \text{DW} < 4$	H_0 se zamítá - autokorelace
$4 - d_U < \text{DW} < 4 - d_L$	Nelze se rozhodnout, je třeba zvýšit T
$2 < \text{DW} < 4 - d_U$	H_0 se přijímá – není autokorelace
$d_U < \text{DW} < 2$	H_0 se přijímá – není autokorelace
$d_L < \text{DW} < d_U$	Nelze se rozhodnout, je třeba zvýšit T
$0 < \text{DW} < d_L$	H_0 se zamítá - autokorelace

Poznámka: dU a dL označují horní a dolní kritické hodnoty, které závisí na počtu pozorování T a na počtu regresorů k.

Tabulka 2-2 Výsledky Durbinova-Watsonova testu

Zdroj: Arlt et al. (2002, s. 28)

Závěry Durbinova-Watsonova testu jsou znázorněny v Tabulce 2-2. Pro zjednodušení Vogelvang (2005) ukazuje tři základní charakteristiky DW statistiky:

- a) neexistuje autokorelace reziduálů: $\rho = 0 \implies DW \approx 2$,
- b) pozitivní autokorelace reziduálů: $\rho > 0 \implies DW < 2$,
- c) negativní autokorelace reziduálů: $\rho < 0 \implies DW > 2$.

Normalita

O normálním modelu mluvíme v situaci, kdy k výše uvedeným předpokladům modelu lineární regrese přidáme další předpoklad, tedy reziduální složky jsou normálně rozdělené pro všechna t (hodnoty momentů tohoto normálního rozdělení, totiž nulová střední hodnota a konstantní rozptyl σ^2, byly již stanoveny v předchozích předpokladech).

Normalita je nejjednodušším předpokladem v situaci, kdy je nutné specifikovat pravděpodobnostní rozdělení reziduální složky (aby bylo možné statisticky testovat různé hypotézy v modelu, konstruovat spolehlivostní a předpovědní intervaly, nalézt maximálně věrohodné odhady parametrů atd.). I když lez normalitu často zdůvodnit pomocí centrální limitní věty, může být někdy ve finanční ekonometrii až příliš zjednodušující. Normalitu modelu lze statisticky testovat. Pro testování normálního rozdělení reziduální složky lze použít grafické analýzy nebo neparametrické testy normality.

Neparametrické testy vycházejí z nulové hypotézy normálního rozdělení reziduální složky. K nejpoužívanějším testům pro testování normality reziduální složky se používají zejména chi–test dobré shody, Jarque–Bera test a Kolmogorovův–Smirnovův test (Hančlová, 2012).

Jedním z často využívaných testů normality je test Jarque-Bera, který bývá běžnou součástí statistických a ekonometrických softwarů.

Normální rozdělení je plně střední hodnotou μ a rozptylem σ^2, ale jsou pro něj charakteristické také dvě další vlastnosti související s šikmostí a špičatostí:

1) Normální rozdělení je symetrické, takže jeho koeficient šikmosti (skewness) je nulový. Jedná se o koeficient, který popisuje odlišnost četností menších a větších hodnot.

2) Normální rozdělení je mesokurtické, tj. jeho koeficient špičatosti (kurtosis) je nulový. Koeficient špičatosti popisuje odlišnost četností prostředních a krajních hodnot.

Test normality Jarque-Bera formalizuje vlastnosti normálního rozdělení. Nulová hypotéza v případě testu normality zní: Řada má normální rozdělení.

V případě zjištěné nenormality je nutné:

- přijmout nenormalitu jako odůvodněnou a reagovat na ni použitím specifického modelu (např. modelu typu GARCH s podmíněnou heteroskedasticitou pro finanční časové řady),
- transformovat proměnné či celý model vhodným způsobem,
- použít robustní modely, které jsou necitlivé na typ pravděpodobnostního rozdělení,
- modelovat odlehlá pozorování (outliers), které jsou někdy příčinou nenormality, pomocí kvalitativních proměnných (dummy proměnných): nenormalita se tak vysvětlí v systematické části modelu a rozdělení reziduálních složek zůstane normální.

Testy pro jednotlivé parametry

Za předpokladu normality modelu v něm můžeme ověřovat různé hypotézy, a to jak individuálně (pro jednotlivé parametry), tak souhrnně (pro více parametrů najednou).

Nebudeme se věnovat veškerým testům, pro zájemce lze dohledat v literatuře (např. Cipra, 2014). Nejčastěji se v ekonometrii provádí testy významnosti regresních parametrů s $\beta_i^* = 0$, tedy testujeme, zda i-tý regresor $x_i (i = 1, \dots k)$ opravdu do modelu patří. Nulová hypotéza, kterou testujeme, zde zní: $H_0: B_i = 0$.

Souhrnné testy pro více parametrů

Opět se na souhrnné testy podíváme pouze velmi zjednodušeně a definujeme si pouze *F*-statistiku, která je ve výstupu příslušného testu v EViews. Tato procedura se označuje zjendodušeně jako *F-test*. Důležitým případem aplikace *F*-testu je pro nulovou hypotézu ve tvaru:

$$H_0: \beta_2 = \beta_3 = \cdots = \beta_k = 0 \tag{10}$$

Pokud tuto hypotézu nelze zamítnout, tak to znamená, že žádný z regresorů modelu není schopen vysvětlit změny vysvětlované proměnné.

ŘEŠENÁ ÚLOHA

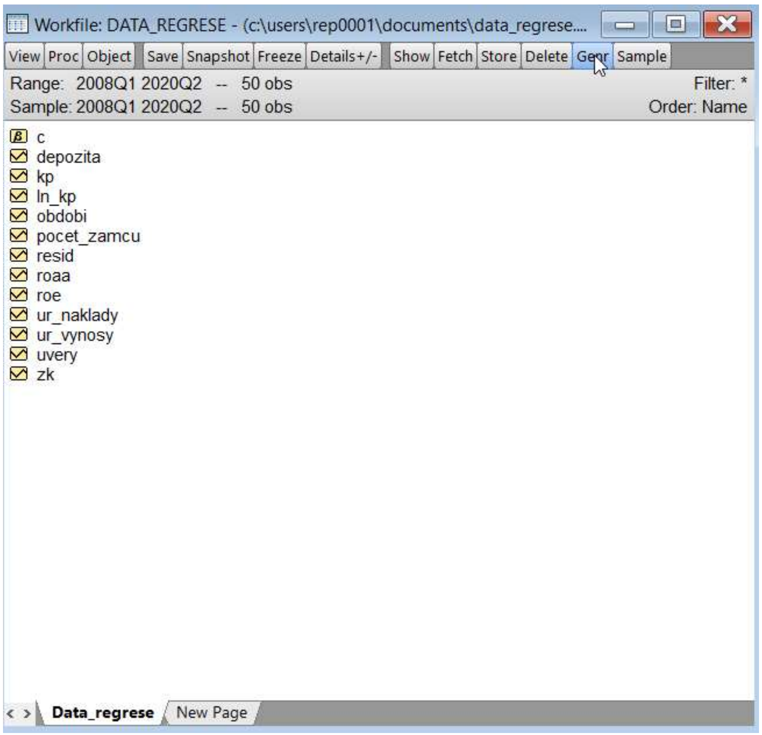
V řešené úloze bude proveden konkrétní odhad regresní analýzy pomocí metody nejmenších čtverců včetně testování předpokladů metody nejmenších čtverců. Zadání příkladu:

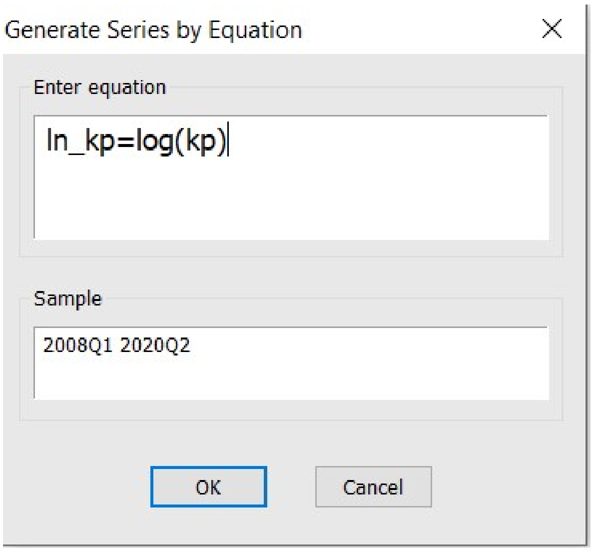
Odhadněte jednotlivé determinanty ukazatele ROE (návratnosti kapitálu) v českém bankovním sektoru vyjádřena v procentech. Jako endogenní proměnná je tedy proměnná ROE za český bankovní sektor v období od 1. čtvrtletí 2008 do 2. čtvrtletí 2020. Jako exogenní proměnné jsou zvoleny následující proměnné:

- celková klientská depozita (v mil. Kč),
- celkové poskytnuté úvěry (v mil. Kč),
- základní kapitál (v mil. Kč),
- počet zaměstnanců v bankovním sektoru,
- návratnost průměrných aktiv - ROAA (v %),
- úrokové náklady (v mil. Kč),
- úrokové výnosy (v mil. Kč),
- kapitálová přiměřenost (v %).

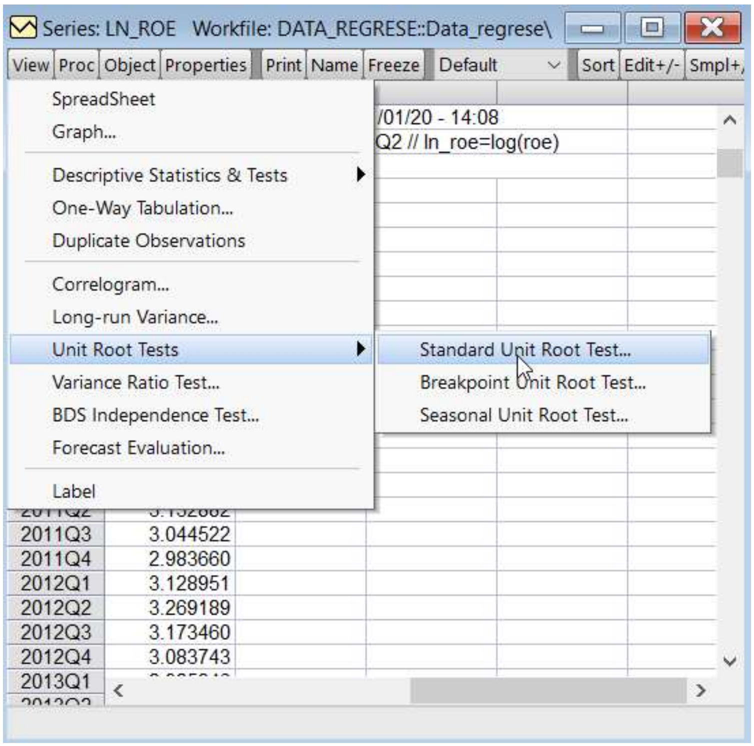
Veškeré údaje jsou za období 1. čtvrtletí 2008 - 2. čtvrtletí 2020.

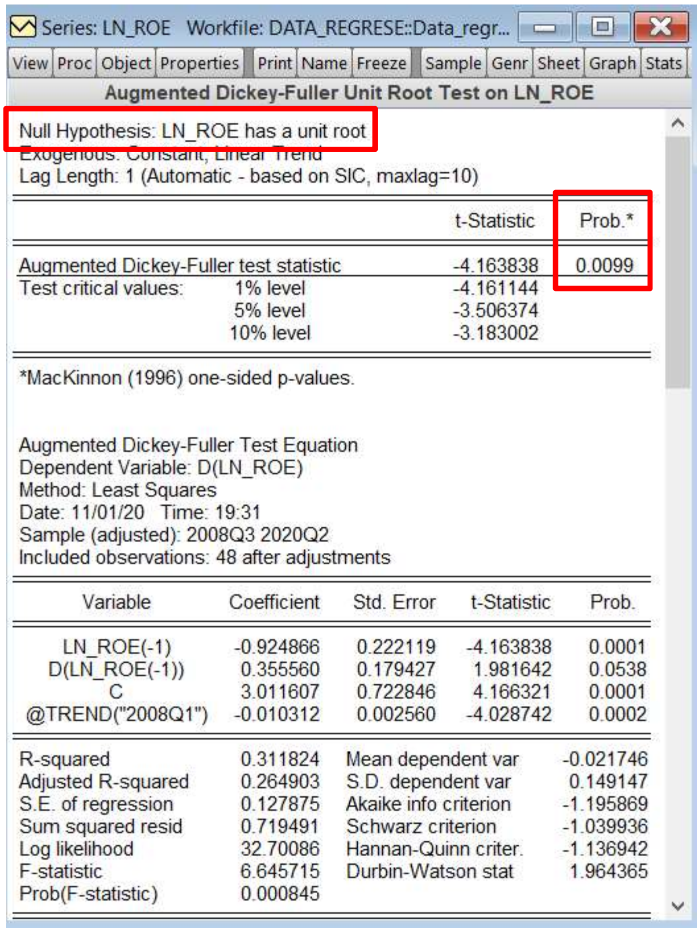
Dále veškerá data zlogaritmujeme a pracujeme s logaritmovanými daty. Logaritmujeme nejen vybrané čtyři vysvětlující proměnné, ale i vysvětlovanou proměnnou (ROE). Postup logaritmování časové řady (základního kapitálu, zkratka ZK):





V dalším kroku otestujeme stacionaritu veškerých časových řad. V textu bude znázorněn test stacionarita pro dvě časové řady a stejným způsobem otestujeme stacionaritu všech časových řad.





Series: LN_ROE Workfile: DATA_REGRESE::Data_regr...

View Proc Object Properties Print Name Freeze Sample Genr Sheet Graph Stats

Augmented Dickey-Fuller Unit Root Test on LN_ROE

Null Hypothesis: LN_ROE has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 1 (Automatic - based on SIC, maxlag=10)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-4.163838	0.0099
Test critical values:	1% level	-4.161144	
	5% level	-3.506374	
	10% level	-3.183002	

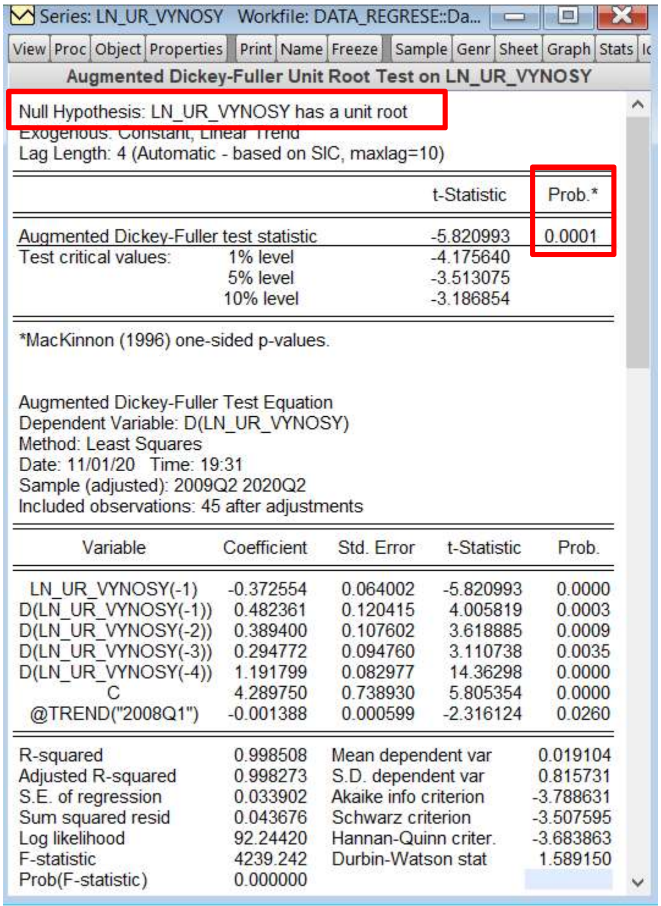
*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LN_ROE)
Method: Least Squares
Date: 11/01/20 Time: 19:31
Sample (adjusted): 2008Q3 2020Q2
Included observations: 48 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LN_ROE(-1)	-0.924866	0.222119	-4.163838	0.0001
D(LN_ROE(-1))	0.355560	0.179427	1.981642	0.0538
C	3.011607	0.722846	4.166321	0.0001
@TREND("2008Q1")	-0.010312	0.002560	-4.028742	0.0002

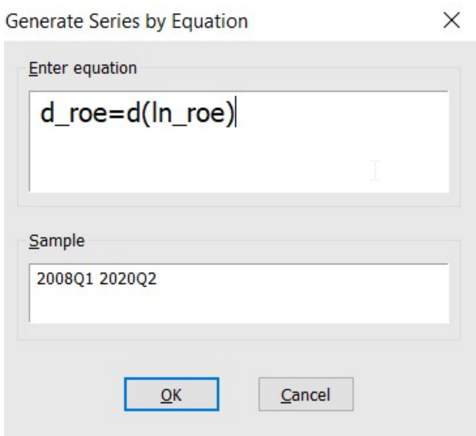
R-squared	0.311824	Mean dependent var	-0.021746
Adjusted R-squared	0.264903	S.D. dependent var	0.149147
S.E. of regression	0.127875	Akaike info criterion	-1.195869
Sum squared resid	0.719491	Schwarz criterion	-1.039936
Log likelihood	32.70086	Hannan-Quinn criter.	-1.136942
F-statistic	6.645715	Durbin-Watson stat	1.964365
Prob(F-statistic)	0.000845		

Časová řada ROE není stacionární. Pomocí p-hodnoty nelze zamítnout nulovou hypotézu, že časová řada má jednotkový kořen.

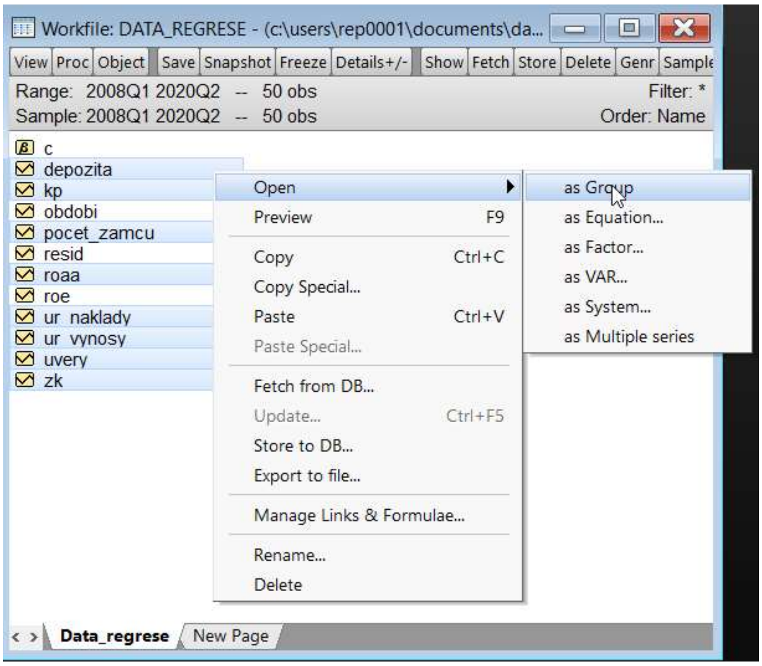


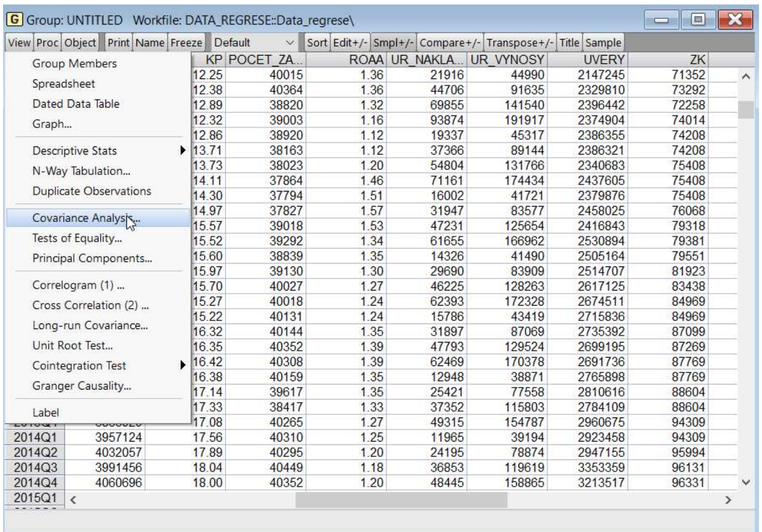
Časová řada úrokové výnosy je stacionární. Pomocí *p*-hodnoty zamítáme nulovou hypotézu, že časová řada má jednotkový kořen.

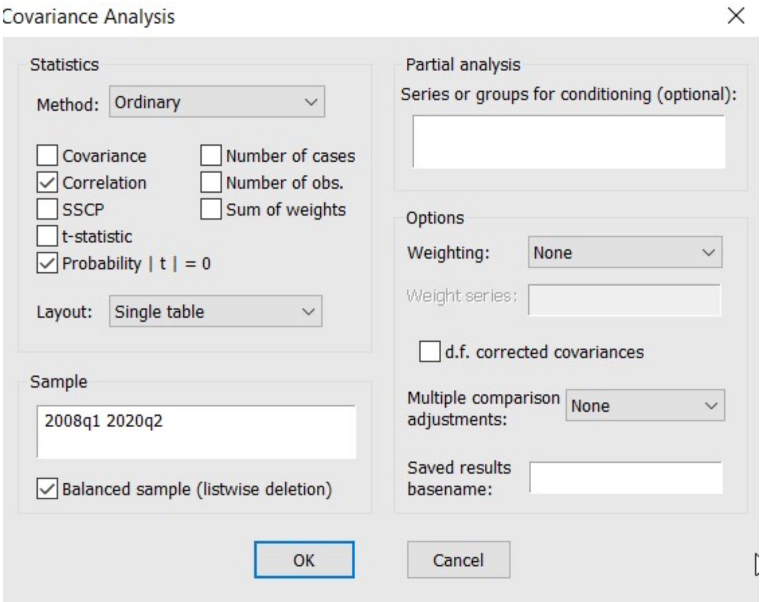
Nestacionární časové řady je nutné převést na stacionární, např. převedením na první diferenci:



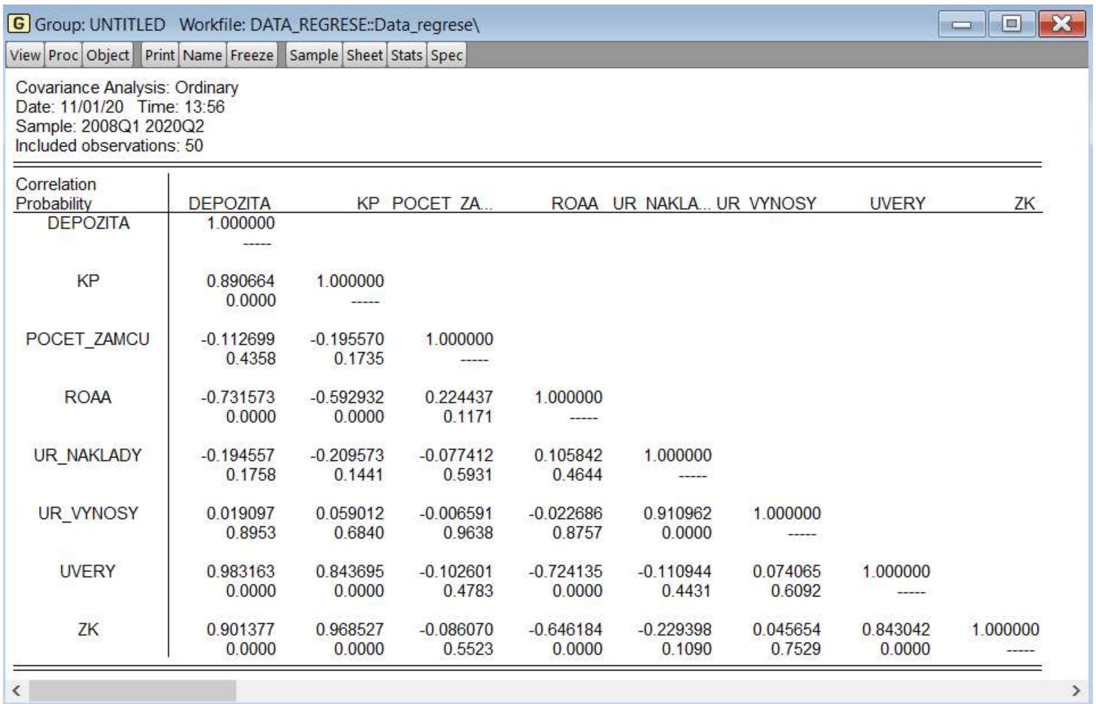
Nyní budeme testovat předpoklady metody nejmenších čtverců. Nejprve budeme testovat multikolinearitu. Otevřeme veškeré exogenní proměnné jako skupinu v EViews.







Zvolíme korelační analýzu. Výsledky korelační matice:



Group: UNTITLED Workfile: DATA_REGRESE::Data_regrese\

View | Proc | Object | Print | Name | Freeze | Sample | Sheet | Stats | Spec

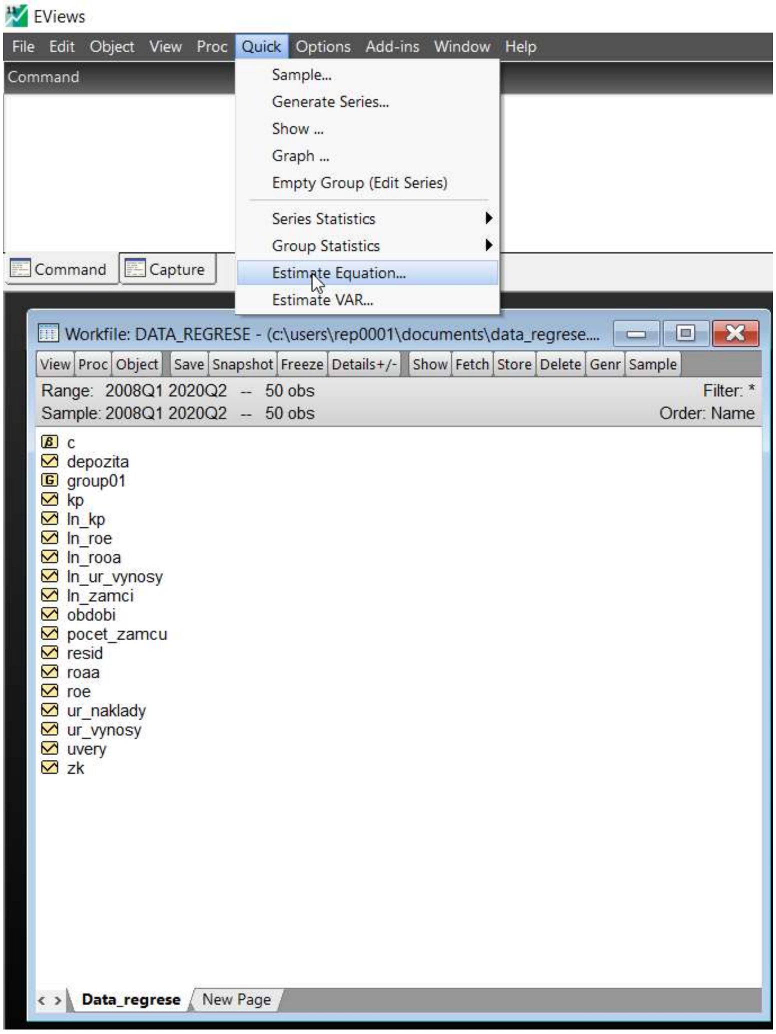
Covariance Analysis: Ordinary
Date: 11/01/20 Time: 13:56
Sample: 2008Q1 2020Q2
Included observations: 50

Correlation Probability	DEPOZITA	KP	POCET_ZA...	ROAA	UR_NAKLA...	UR_VYNOSY	UVERY	ZK
DEPOZITA	1.000000 -----							
KP	0.890664 0.0000	1.000000 -----						
POCET_ZAMCU	-0.112699 0.4358	-0.195570 0.1735	1.000000 -----					
ROAA	-0.731573 0.0000	-0.592932 0.0000	0.224437 0.1171	1.000000 -----				
UR_NAKLADY	-0.194557 0.1758	-0.209573 0.1441	-0.077412 0.5931	0.105842 0.4644	1.000000 -----			
UR_VYNOSY	0.019097 0.8953	0.059012 0.6840	-0.006591 0.9638	-0.022686 0.8757	0.910962 0.0000	1.000000 -----		
UVERY	0.983163 0.0000	0.843695 0.0000	-0.102601 0.4783	-0.724135 0.0000	-0.110944 0.4431	0.074065 0.6092	1.000000 -----	
ZK	0.901377 0.0000	0.968527 0.0000	-0.086070 0.5523	-0.646184 0.0000	-0.229398 0.1090	0.045654 0.7529	0.843042 0.0000	1.000000 -----

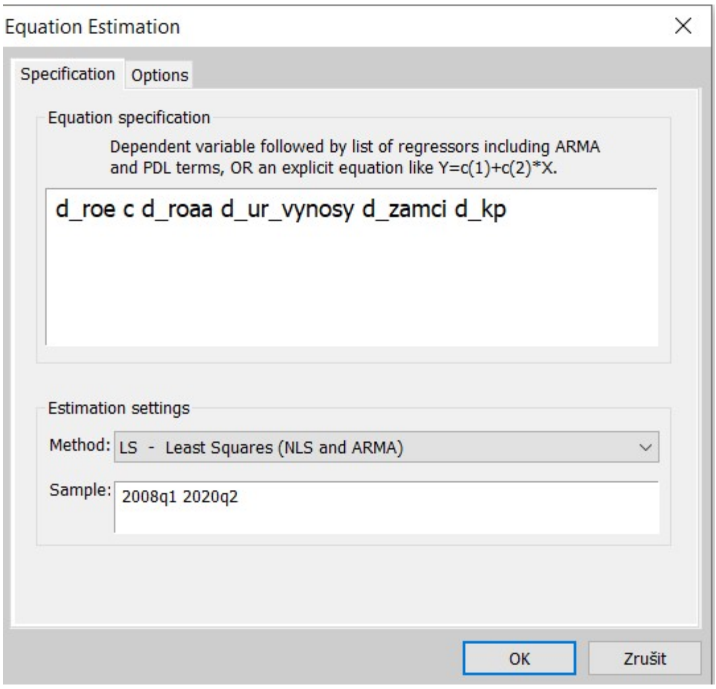
Z korelační matice je patrné, že existuje několik silných závislostí. Silnou závislost vidíme mezi depozity a kapitálovou přiměřeností, mezi úvěry a depozity, mezi úvěry a kapitálovou přiměřeností, dále mezi úrokovými náklady a úrokovými výnosy, základním kapitálem a depozity či mezi základním kapitálem a kapitálovou přiměřenosti. Je nutné odstranit multikolinearitu. Je vhodné se rozhodnout, která data potřebujeme v modelu ponechat a která můžeme vyřadit. Závisí to samozřejmě na cíli práce, na tom, co zkoumáme a co chceme zjistit. Nyní máme pouze teoretický příklad a zabýváme se ekonometrickou stranou (ne ekonomickou), proto vyloučíme libovolné proměnné, které mají mezi sebou silnou vazbu. V tomto příkladu budeme dál pracovat s proměnnými:

- kapitálová přiměřenost,
- počet zaměstnanců,
- ROAA,
- úrokové výnosy.

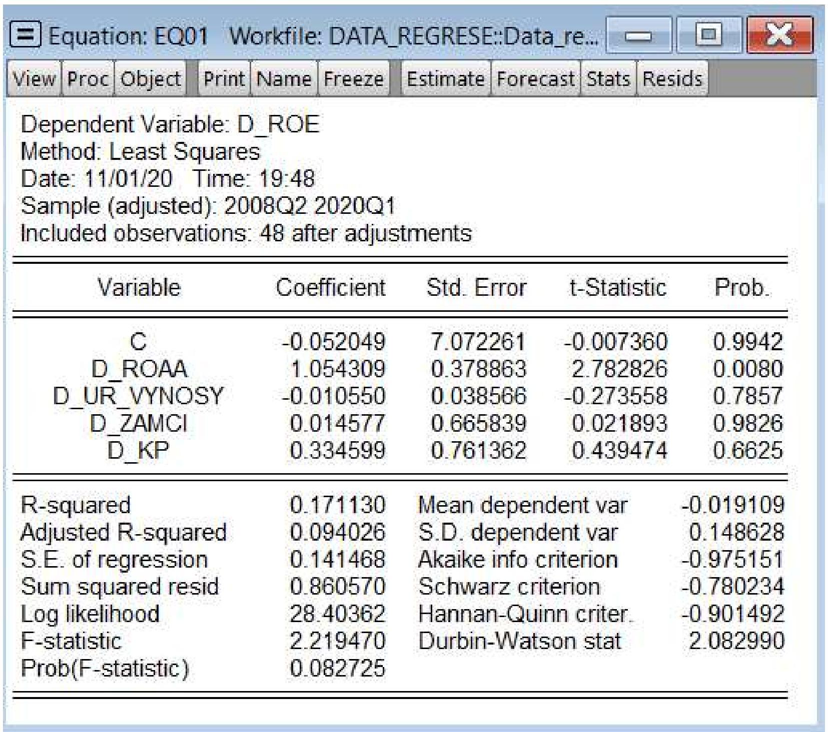
Nyní vytvoříme regresní rovnici. Existuje několik způsobů, jak vytvořit regresní rovnici.



V dalším kroku zapíšeme regresní rovnici a vytvoříme lineární regresní model v EViews. Vždy je nutné na prvním místě zapsat vysvětlovanou proměnnou, poté není důležité pořadí vysvětlujících proměnných a do rovnice doplníme také konstantu (c). Volíme metodu nejmenších čtverců a provádíme odhad regresní rovnice za celé zkoumané období.



Výstup regresní rovnice:

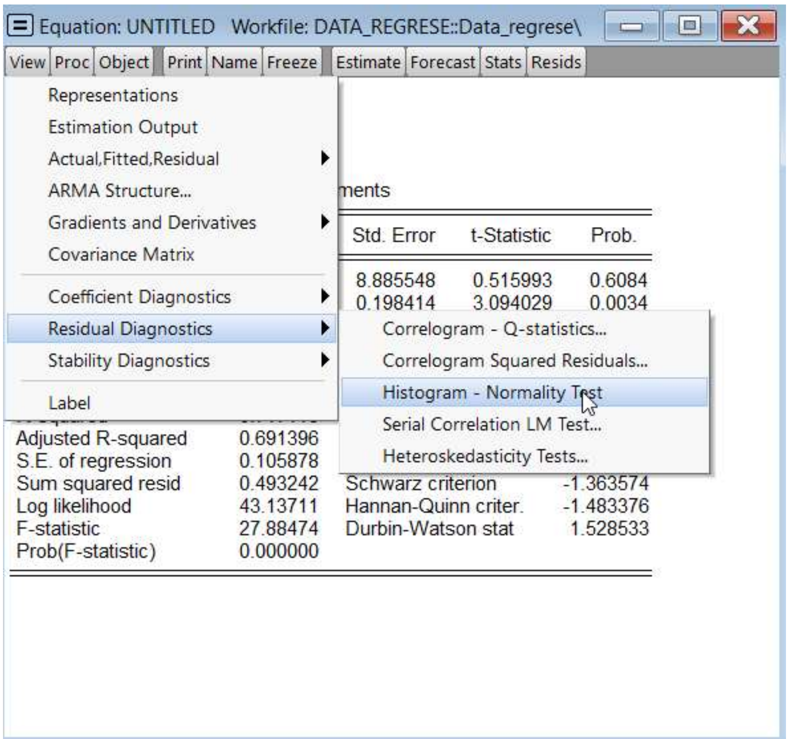


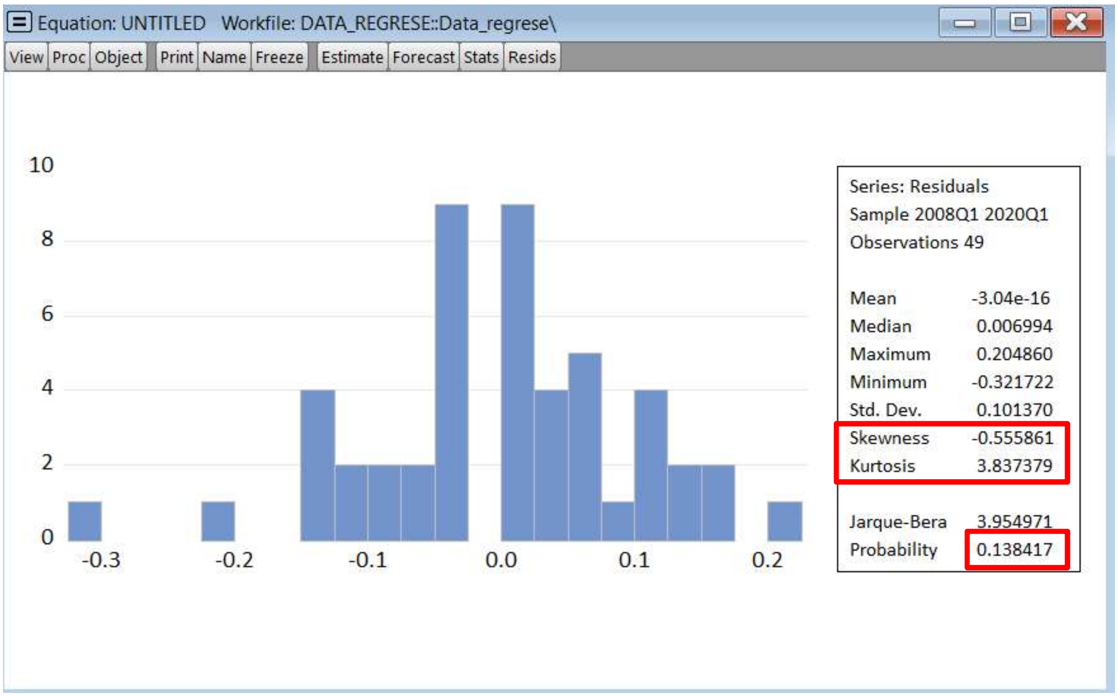
Dependent Variable: D_ROE
Method: Least Squares
Date: 11/01/20 Time: 19:48
Sample (adjusted): 2008Q2 2020Q1
Included observations: 48 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.052049	7.072261	-0.007360	0.9942
D_ROAA	1.054309	0.378863	2.782826	0.0080
D_UR_VYNOSY	-0.010550	0.038566	-0.273558	0.7857
D_ZAMCI	0.014577	0.665839	0.021893	0.9826
D_KP	0.334599	0.761362	0.439474	0.6625

R-squared	0.171130	Mean dependent var	-0.019109
Adjusted R-squared	0.094026	S.D. dependent var	0.148628
S.E. of regression	0.141468	Akaike info criterion	-0.975151
Sum squared resid	0.860570	Schwarz criterion	-0.780234
Log likelihood	28.40362	Hannan-Quinn criter.	-0.901492
F-statistic	2.219470	Durbin-Watson stat	2.082990
Prob(F-statistic)	0.082725		

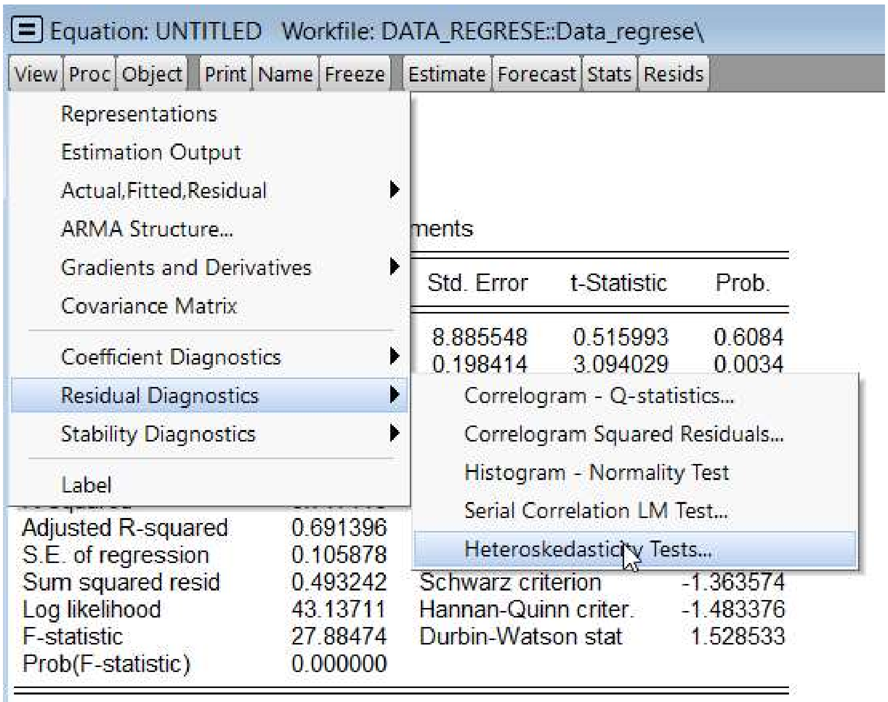
Nejprve je nutné otestovat zbývající předpoklady metody nejmenších čtverců, protože se jedná o testování normality reziduální složky, heteroskedasticity reziduální složky či autokorelace reziduí, provádí se tyto testy až po vytvoření regresního modelu. Jak bylo uvedeno, jedná se o diagnostiku reziduální složky, najdeme všechny testy v části diagnostika reziduálu. Začneme např. testem normality:



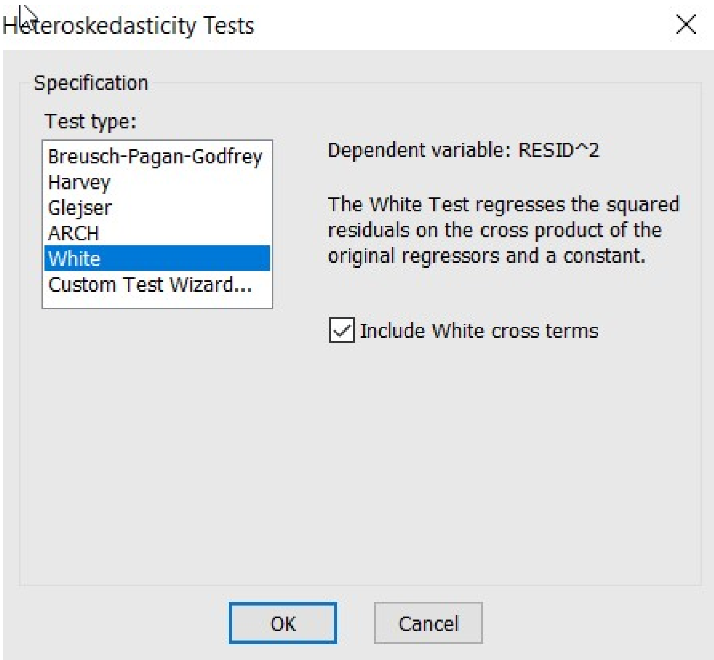


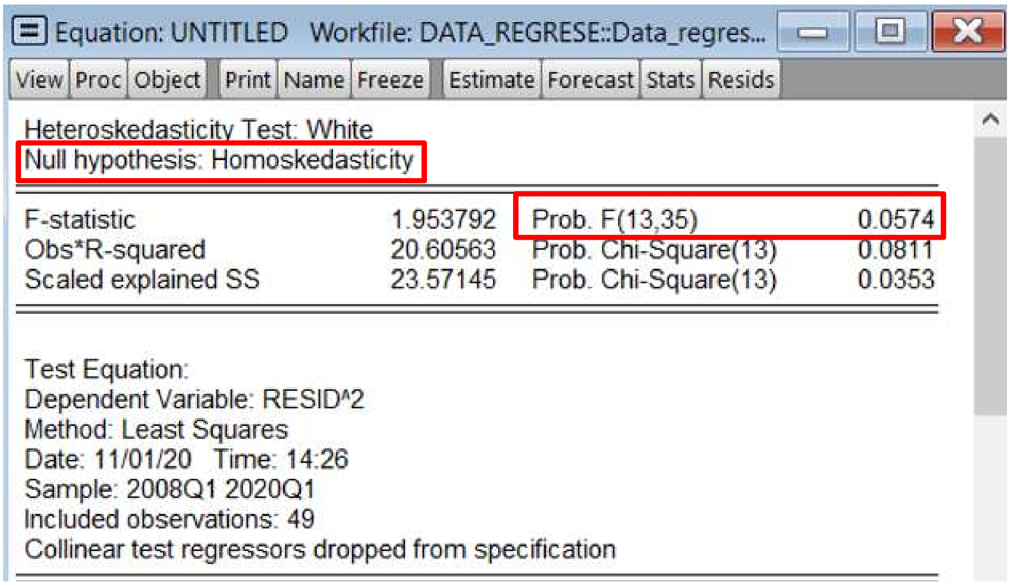
Dle grafického rozdělení, stejně jako z hodnoty šikmosti a špičatosti není jednoznačně patrné, zda reziduální složka má normální rozdělení. Dle Jarque-Bera testu není možné zamítnout nulovou hypotézu a můžeme tedy říct, že reziduální složka má normální rozdělení.

Dále budeme testovat heteroskedasticitu, resp. homoskedasticitu:



Vybereme typ testu (např. White test):





Equation: UNTITLED Workfile: DATA_REGRESE::Data_regres...

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

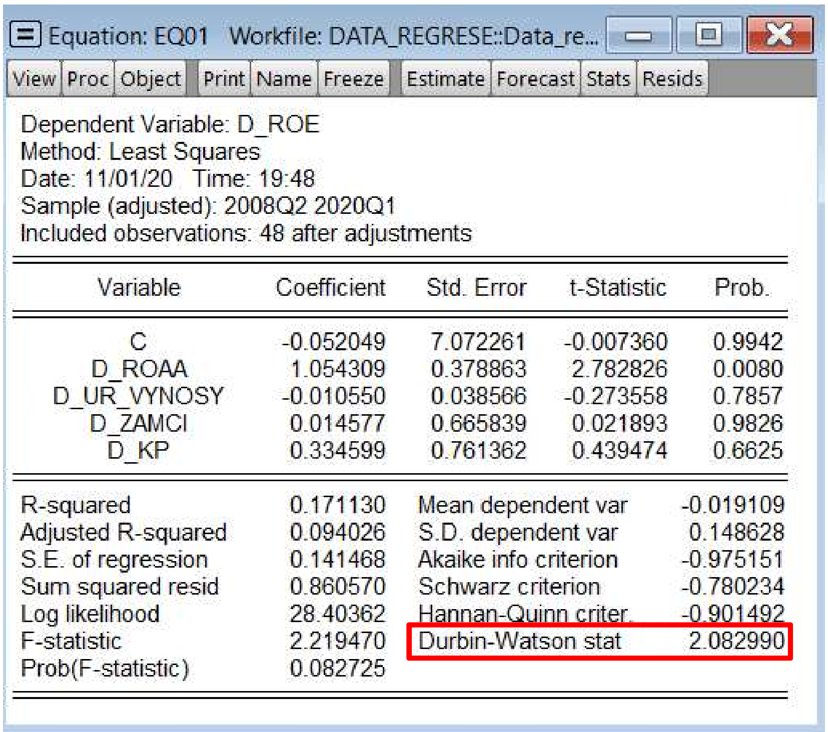
Heteroskedasticity Test: White
Null hypothesis: Homoskedasticity

F-statistic	1.953792	Prob. F(13,35)	0.0574
Obs*R-squared	20.60563	Prob. Chi-Square(13)	0.0811
Scaled explained SS	23.57145	Prob. Chi-Square(13)	0.0353

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 11/01/20 Time: 14:26
Sample: 2008Q1 2020Q1
Included observations: 49
Collinear test regressors dropped from specification

Pomocí *p*-hodnoty testujeme nulovou hypotézu, která je zde uvedena (H0: Homoskedasticita). Použitím *p*-hodnoty nelze zamítnout nulovou hypotézu, tedy je zde homoskedasticita rezidální složky. Tento předpoklad je splněn.

Posledním testovaným předpokladem použití metody nejmenších čtverců je autokorelace reziduální složky. Využijeme zjednodušeného pravidla u Durbin-Watson testu z regresního výstupu:



Equation: EQ01 Workfile: DATA_REGRESE::Data_re...

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: D_ROE
Method: Least Squares
Date: 11/01/20 Time: 19:48
Sample (adjusted): 2008Q2 2020Q1
Included observations: 48 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.052049	7.072261	-0.007360	0.9942
D_ROAA	1.054309	0.378863	2.782826	0.0080
D_UR_VYNOSY	-0.010550	0.038566	-0.273558	0.7857
D_ZAMCI	0.014577	0.665839	0.021893	0.9826
D_KP	0.334599	0.761362	0.439474	0.6625

R-squared	0.171130	Mean dependent var	-0.019109
Adjusted R-squared	0.094026	S.D. dependent var	0.148628
S.E. of regression	0.141468	Akaike info criterion	-0.975151
Sum squared resid	0.860570	Schwarz criterion	-0.780234
Log likelihood	28.40362	Hannan-Quinn criter.	-0.901492
F-statistic	2.219470	Durbin-Watson stat	2.082990
Prob(F-statistic)	0.082725		

Je vidět, že DW statistika je kolem 2 a je splněn i předpoklad, že reziduální složka nemá autokorelaci.

Uvedli jsem si jednotlivé předpoklady použití metody nejmenších čtverců. Z předchozí části řešené úlohy vidíme, že máme splněny všechny předpoklady použití metody nejmenších čtverců. Proto můžeme daný model interpretovat. V dalším kroku bude ukázán příklad, jak se dá výsledný regresní model interpretovat.

Equation: EQ01 Workfile: DATA_REGRESE::Data_re...

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: D_ROE
Method: Least Squares
Date: 11/01/20 Time: 19:48
Sample (adjusted): 2008Q2 2020Q1
Included observations: 48 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.052049	7.072261	-0.007360	0.9942
D_ROAA	1.054309	0.378863	2.782826	0.0080
D_UR_VYNOSY	-0.010550	0.038566	-0.273558	0.7857
D_ZAMCI	0.014577	0.665839	0.021893	0.9826
D_KP	0.334599	0.761362	0.439474	0.6625

R-squared	0.171130	Mean dependent var	-0.019109
Adjusted R-squared	0.094026	S.D. dependent var	0.148628
S.E. of regression	0.141468	Akaike info criterion	-0.975151
Sum squared resid	0.860570	Schwarz criterion	-0.780234
Log likelihood	28.40362	Hannan-Quinn criter.	-0.901492
F-statistic	2.219470	Durbin-Watson stat	2.082990
Prob(F-statistic)	0.082725		

Koeficient determinance (Adj. R-squared) je ve výši 0,09.

Equation: EQ01 Workfile: DATA_REGRESE::Data_re...

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: D_ROE
Method: Least Squares
Date: 11/01/20 Time: 19:48
Sample (adjusted): 2008Q2 2020Q1
Included observations: 48 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.052049	7.072261	-0.007360	0.9942
D_ROAA	1.054309	0.378863	2.782826	0.0080
D_UR_VYNOSY	-0.010550	0.038566	-0.273558	0.7857
D_ZAMCI	0.014577	0.665839	0.021893	0.9826
D_KP	0.334599	0.761362	0.439474	0.6625

R-squared	0.171130	Mean dependent var	-0.019109
Adjusted R-squared	0.094026	S.D. dependent var	0.148628
S.E. of regression	0.141468	Akaike info criterion	-0.975151
Sum squared resid	0.860570	Schwarz criterion	-0.780234
Log likelihood	28.40362	Hannan-Quinn criter.	-0.901492
F-statistic	2.219470	Durbin-Watson stat	2.082990
Prob(F-statistic)	0.082725		

Dále řádek po řádku interpretujeme vliv dané proměnné na ROE. Konkrétně logaritmovaná hodnota ROAA pozitivně (dle znaménka před hodnotou koeficientu) ovlivňuje vývoj ROE. Statistickou významnost parametru testujeme pomocí *p*-hodnoty (Prob.) a vidíme, že parametr ROAA je statisticky významný (dle hodnoty Prob., která je nižší než hladina významnosti, tedy 0,05). Můžeme tedy říci, že vývoj ROAA v českém bankovním sektoru pozitivně ovlivňuje vývoj ROE. Stejným způsobem interpretujeme úrokové výnosy.

U počtu zaměstnanců v bankovním sektoru je vidět, že parametr dané proměnné není statistiky významný:

Equation: EQ01 Workfile: DATA_REGRESE::Data_re...

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: D_ROE
Method: Least Squares
Date: 11/01/20 Time: 19:48
Sample (adjusted): 2008Q2 2020Q1
Included observations: 48 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.052049	7.072261	-0.007360	0.9942
D_ROAA	1.054309	0.378863	2.782826	0.0080
D_UR_VYNOSY	-0.010550	0.038566	-0.273558	0.7857
D_ZAMCI	0.014577	0.665839	0.021893	0.9826
D_KP	0.334599	0.761362	0.439474	0.6625

R-squared	0.171130	Mean dependent var	-0.019109
Adjusted R-squared	0.094026	S.D. dependent var	0.148628
S.E. of regression	0.141468	Akaike info criterion	-0.975151
Sum squared resid	0.860570	Schwarz criterion	-0.780234
Log likelihood	28.40362	Hannan-Quinn criter.	-0.901492
F-statistic	2.219470	Durbin-Watson stat	2.082990
Prob(F-statistic)	0.082725		

Dle *p*-hodnoty testujeme nulovou hypotézu, zda je parametr počtu zaměstnanců statisticky významný. Konkrétně: Prob. je větší než 5% hladina významnosti, tedy 0.9826>0.05, nelze tedy zamítnou nulovou hypotézu a parametr dané proměnné není statisticky významný. Nelze tedy určit, zda počet zaměstnanců ovlivňuje vývoj hodnoty ROE v českém bankovní sektoru.

2.3 Modely diskrétní volby, modely typu Logit, Probit a Tobit

Vysvětlované proměnné však nemusí nabývat pouze kvantitativních spojitých hodnot (výše úrokové sazby, cena akcií apod.) Ve finanční ekonometrii se můžeme setkat také z praxí, kdy vysvětlované proměnné mohou nabývat hodnot, které jsou diskrétní či kategoriální (např. ratingové hodnocení) nebo jsou pozorovatelné jen určitém rozsahu.

Ve finanční ekonometrii často potřebujeme pracovat s kvalitativními proměnnými. Tedy zejména proměnnými, které mají odpověď ano či ne. Tyto kvalitativní proměnné lze do modelu zakódovat jako lineární kombinaci vhodných nula-jedničkových proměnných (tj. binárních proměnných nabývajících jen hodnot 0 nebo 1), které označujeme jako dummy proměnné (umělé proměnné).

Velmi častým typem kategoriální (nespojité) vysvětlované proměnné je binární proměnná, která nabývá hodnot pouze dvou kategorií, tedy jedničky či nuly. Tento typ proměnné se vyskytuje především v těchto případech:

- Jedná se o dummy proměnnou, tj. proměnnou, která nabývá kvůli své podstatě pouze dvou hodnot. Může to být logická proměnná, odpověď v anketě ano či ne atd.
- Jedná se o proměnnou, která je vytvořená z jiné jejím zjednodušením, např. cena výrobku vyšší, či nižší než padesát korun, výše hypotečního úvěru do 1 milionu nebo nad 1 milion atd.

Pravá část rovnice vysvětlující tuto binární proměnnou je stejná jako v případě lineárního regresního modelu.

2.3.1 Modely typu Logit, Probit a Tobit

Probit a logit modely jsou vhodné v případě, kdy je volba omezena na dvě věci (např. investuji či neinvestuji do dané akcie, přísluší daná společnost k finančnímu konglomerátu či nikoli apod.). Jedná se o modely binární volby. V řadě případů však je možnost volby z více alternativ. V některých otázkách volby lze zařadit i další volbu (např. investice do akcií společnosti A, společnosti B, společnosti X atd. (a tak rozšířit paletu alternativ na tři). Z toho důvodu je tedy možné modely probit a logit rozšířit i pro tyto případy a výsledné modely se nazývají multinomiální probit a multinomiální logit modely.

Logitový regresní model

Odlišnost mezi regresním modelem typu Logit a lineárním regresním modelem je takový, že závislá proměnná v regresním modelu typu Logit je diskrétní, často binární tj. nula nebo jednička (dichotomická). Klasický lineární model předpokládá, že závislá proměnná je spojitá. Tento rozdíl se projevuje jak ve výběru modelu, tak i v jeho předpokladech. Metody analýzy využívající logitovou regresi, se řídí stejnými všeobecnými pravidly využívanými v lineární regresi (Hosmer a Lemeshow, 2000).

Předpokládejme, že máme *n* nezávislých pozorování vyjádřené jako dvojici (xi, yi), i = 1,2,.,n, kde yi znamená dichotomickou závislou proměnnou a xi je nezávislá proměnná pro i-té pozorování. Dále předpokládejme, že závisle proměnná nabývá hodnot nula nebo jedna, vyjadřující absenci, respektive přítomnost dané charakteristiky. K formulaci logistického regresního modelu z rovnice (1) je třeba mít adekvátní množství dat, které umožní odhad neznámých parametrů β0 a β1.

V lineární regresi se často používá k odhadu parametrů metoda nejmenších čtverců. V této metodě jsou voleny takové hodnoty β0 a β1, které minimalizují hodnotu čtvercové odchylky pozorované hodnoty Y od predikované hodnoty v tomto modelu. Metoda nejmenších čtverců nám poskytuje odhad s optimálními statistickými vlastnostmi, avšak když se MNČ použije na model s dichotomickou závisle proměnnou, pak tento odhad již tyto vlastnosti postrádá.

Další metodou odhadu parametrů je metoda maximální věrohodnosti (MMV). Metoda maximální věrohodnosti poskytuje hodnoty neznámých parametrů, které maximalizují pravděpodobnost získání zkoumaných dat. Odhad parametrů pomocí metody maximální věrohodnosti je vybrán tak, aby byla maximalizována hodnota této funkce. Výsledný odhad je co nejblíže napozorovaným hodnotám (Hosmer a Lemeshow, 2000).

Logitová regrese je ekonometrická metoda, která patří mezi takzvané modely diskrétní volby. Znamená to, že vysvětlovaná proměnná, označme ji y, nabývá pouze dvou stavů, 0 nebo 1. Máme k dispozici n realizací vysvětlované proměnné takových, že platí:

$y_i = 1$ s pravděpodobností p_i a $y_i = 0$ s pravděpodobností $1 - p_i$.

Podíl $\frac{p_i}{1-p_i}$ porovnává pravděpodobnost hodnoty 1 a pravděpodobnost hodnoty 0 a nazývá se šance. Pro tento podíl platí: $\frac{p_i}{1-p_i} \epsilon \langle 0; +\infty)$. K získání hodnot z celého intervalu $(-\infty; +\infty)$ je nutné použit logaritmickou transformaci: $\eta_i = \ln\left(\frac{p_i}{1-p_i}\right)$. Pro vytvoření logitového modelu předpokládáme, že η_i lineárně závisí na vysvětlujících proměnných $x_1, x_2, x_3, \dots, x_k$. Daný vztah pak vypadá následovně:

$$\eta_i(y_i) = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_k x_{ki}, \text{ kde } i = 1, \dots, n. \tag{11}$$

PROBITOVÝ MODEL

Dalším nelineárním modelem binární volby je probitový model. Je-li náhodná složka v regresi standardně normálně rozdělené, potom se jedná o probitový model.

TOBIT MODEL

V některých případech se lze setkat se situací, kdy data jsou nějakým způsobem cenzorována. Jako příklad můžeme uvést data z dotazníkového šetření. Tobit model předpokládá, že závislá proměnná má limitovanou/omezenou formu, tj. její hodnoty se rovnají (v tomto případě) nule ve významném počtu pozorování. Tento model shrnuje selekci a alokaci pomoci do jedné regresní rovnice a je pro něj zásadní zejména předpoklad normality a homoskedasticity latentní proměnné. Nevýhodou modelu je hlavně omezení, že všechny vysvětlující proměnné působí ve stejném směru na rozhodnutí o poskytnutí (nenulové) pomoci a na rozhodnutí o výši alokované pomoci.

SHRNUTÍ KAPITOLY

Druhá kapitola studijní opory se věnovala finančním časovým řadám a jejím charakteristikám. Pokud chceme sestavit ekonometrický model, nejprve potřebujeme vhodná data. Proto je na počátku kapitoly popsán sběr a úprava dat. Prezentace dat je zpravidla pomocí grafů nebo deskriptivní statistiky. Celou ekonometrii bude provázet testování hypotéz, jsou tedy popsány způsoby testování hypotéz, konkrétně pomocí kritického oboru, intervalu spolehlivosti či p-hodnoty. Dalším důležitým pojmem je stacionarita, což znamená, že chování dané řady je stochasticky ustálené.

Lineární regresní analýza je jedním z nejdůležitějších nástrojů v aplikované ekonomii pro analýzu vztahu mezi dvěma a více proměnnými. Cílem regresního modelu je najít funkční závislost dané vysvětlované veličiny na jedné či více vysvětlujících veličinách. K odhadu parametrů modelu slouží několik metod, nejvíce využívaná je metoda nejmenších čtverců, maximální věrohodnosti či momentová metoda. Při použití metody nejmenších čtverců je nejprve nutné splnit její předpoklady:

- Střední hodnota reziduální složky je nulová pro všechna t.
- Rozptyl reziduální složky je konstantní a konečný pro všechna t.
- Reziduální složky jsou navzájem nekorelované pro všechna $s \neq t$.
- Regresory jsou ve stejném čase nebo pro stejnou průřezovou jednotku nekorelované s reziduální složkou pro všechna i a t.
- Reziduální složky mají normální rozdělení neboli předpoklad normality.

Pokud data nabývají hodnot, které jsou diskrétní či kategoriální nebo jsou pozorovatelné jen v určitém rozsahu hovoříme o modelech diskrétní volby.

OTÁZKY

1. Závislost mezi posloupností hodnot jedné proměnné, uspořádaných v čase, někdy v prostoru se nazývá:

a) autokorelace náhodné složky

b) autokorelace vysvětlující proměnné

c) nekorelovanost závislé složky

2. Koeficient determinace vyjadřuje podíl variability vysvětlované proměnné podchycený vysvětlujícími proměnnými. ANO x NE

3. Mulikolinearita vyjadřuje:

a) závislost mezi vysvětlující a vysvětlovanou proměnnou

b) závislost mezi vysvětlujícími proměnnými

c) nezávislost mezi vysvětlujícími proměnnými

4. Pokud hodnota Durbin-Watsonovy statistiky je 1,98, znamená to, že pro náhodnou složku:

a) existuje pozitivní autokorelace

b) existuje negativní autokorelace

c) není autokorelace

5. Lineární regresní analýza vyžaduje stacionární data. ANO x NE

ODPOVĚDI

1. a)
2. ANO
3. b)
4. c)
5. ANO

3 MODELY JEDNOROZMĚRNÝCH A VÍCEROZMĚRNÝCH ČASOVÝCH ŘAD

RYCHLÝ NÁHLED KAPITOLY

Ve třetí kapitole studijní opory budou definovány modely jednorozměrných stacionárních a nestacionárních řad a modely vícerozměrných časových řad. Nejprve budou popsány základní pojmy jako je autokorelační a parciální autokorelační funkce. Dále budou definovány ARMA a ARIMA modely. V další části kapitoly budou popsány modely vektorové autoregrese, budou popsány kroky konstrukce modelu a následná diagnostika.

CÍLE KAPITOLY

- Definovat modely jednorozměrných časových řad.
- Popsat Boxovu-Jenkinsonovu metodologii.
- Definovat ARMA modely, jejich konstrukci a diagnostiku.
- Definovat modely vícerozměrných časových řad.
- Popsat konstrukci VAR modelu.

KLÍČOVÁ SLOVA KAPITOLY

Autokorelační funkce, korelogram, parciální autokorelační funkce, AR, MA, ARMA, ARIMA, sezónní časová řada, VAR model.

3.1 Analýza časových řad

S časovými řadami se můžeme setkávat ve finanční ekonometrii velmi běžně. Časovou řadou rozumíme posloupnost hodnot proměnné uspořádaných v čase. Zaznamenávání hodnot určitého jevu v čase není samoúčelné. V současnosti je v podstatě nemožné provádět důležitá ekonomická rozhodnutí bez znalosti minulého vývoje základních ukazatelů. Na podrobnou analýzu vývoje těchto ukazatelů je kladen značný důraz. Z vývoje minulých zaznamenaných hodnot můžeme daný jev nejen analyzovat, ale navíc i předpovídat budoucí vývoj. K tomu je zapotřebí dosavadní průběh časové řady vhodným způsobem modelovat.

Autoři Box a Jenkins (1970) publikovali nový přístup k modelování časových řad. Hlavním přínosem je vytvoření principů, které lze v praxi používat zejména v situacích, kdy nelze použít klasickou dekompoziční analýzu časových řad. Základním prvkem konstrukce modelu časové řady v tomto pojetí je reziduální složka, která může být tvořena korelovanými náhodnými veličinami. Boxova–Jenkinsova teorie je založena na myšlence, že časová řada může být chápána jako řada stochastického charakteru, na jehož základě je možno modelovat systematičnosti v reziduální složce, a doplnit dekompoziční analýzu. Existuje řada oblastí, kde je nezbytné využívat moderní metody analýzy časových řad. Boxově-Jenkinsonově teorii se budeme podrobně věnovat v podkapitole 3.2. Nejprve si však ukážeme dekompozici časové řady, která by zpravidla měla předcházet dalším postupům při analýze časových řad.

3.1.1 DEKOMPOZICE ČASOVÉ ŘADY

K ZAPAMATOVÁNÍ

V klasické dekompozici předpokládáme, že časovou řadu můžeme rozložit na jednotlivé složky:

- trendovou (T_t),
- sezónní (S_t),
- cyklickou (C),
- náhodnou (ε_t).

Trend odráží dlouhodobé změny v průměrném chování časové řady. Vzniká v důsledku působení sil, které systematicky působí ve stejném směru. **Sezónní složka** popisuje periodické změny v časové řadě, které se odehrávají během jedné sezóny (např. kalendářního roku) a každou další sezónu se opakují. Sezónní složka je způsobena takovými faktory jako např. střídání ročních dob, v ekonomické oblasti např. lidské zvyky apod. **Cyklická složka** je periodická složka, která je někdy popisována jako fluktuace okolo trendu (střídání fází růstu a poklesu). Délka a intenzita cyklů může být proměnlivá. Cyklická složka může být důsledkem vnějších vlivů, vytipování příčin ale může být obtížné. Typickým příkladem cyklické složky je obchodní cyklus v ekonomice (konjunktura – krize – deprese). **Náhodná složka** je tvořena náhodnými pohyby (fluktuacemi) v průběhu časové řady, které nemají rozpoznatelný systematický charakter. Pokrývá také chyby měření nebo chyby výpočetní(zaokrouhlování). Z pohledu statistiky se jedná o bílý šum s normálním rozdělením.

Zatímco trend, sezónní a cyklická složka jsou funkcí času, náhodná složka je posloupností náhodných veličin. Časovou řadu si tedy můžeme představit jako trend, na který jsou „nabaleny“ periodické složky (sezónní a cyklická) a bílý šum.

Zatímco první tři složky jsou deterministické a lze je popsat matematickou hodnotou, poslední složka je stochastická a předpokládáme o ní, že má nulovou střední hodnotu, konstantní rozptyl a jednotlivé její hodnoty jsou nekorelované. Tyto předpoklady nám zaručují, že tato náhodná složka neovlivňuje systematicky hodnotu časové řady. Cyklická složka je v krátkodobých časových řadách bezpředmětná.

Při analýze lze použít aditivní model:

$$Y_t = T_t + S_t + C_t + \varepsilon_t \tag{12}$$

a multiplikativní model

$$Y_t = T_t . S_t . C_t . \varepsilon_t \tag{13}$$

Nejdůležitější složkou časových řad je především složka trendová, přičemž ve většině případů vystačíme s nejjednoduššími křivkami – přímkou a parabolou. Sezónní složku lze modelovat například v případě čtvrtletních, měsíčních nebo denních časových řad.

Pro dekompozici je k dispozici několik přístupů, např. klasické dekompoziční metody, které kladou důraz na systémové složky a jednotlivá pozorování se obvykle berou jako nekorelovaná. Naopak Boxova-Jenkinsova metodologie dekompozice časové řady bere za základní prvek konstrukce modelu časové řady reziduální (náhodnou) složku, která může být tvořena korelovanými(závislými) náhodnými veličinami.

Klouzavé průměry

Metoda klouzavých průměrů se řadí mezi adaptivní přístupy k modelování trendové složky. Tyto modely se ukazují jako vhodné, protože rychle reagují na změny v časové řadě. Při konstrukci klouzavých průměrů lze vyrovnávat úseky řady polynomickými křivkami, což ve svém důsledku vede k aplikaci vážených klouzavých průměrů.

Exponenciální vyrovnávání

Exponenciální vyrovnávání je dalším adaptivním přístupem k trendové složce, který lze úspěšně použít. Oproti metodě klouzavých průměrů má exponenciální vyrovnávání tu výhodu, že pracuje se všemi dostupnými minulými pozorováními v časové řadě. Při tvorbě modelu se metoda nejmenších čtverců modifikuje tak, aby váhy jednotlivých čtverců v minimalizovaném součtu směrem do minulosti exponenciálně klesaly. Jedná se tedy o váženou metodu nejmenších čtverců s exponenciálně klesajícími váhami. Váha jsou dány tzv. vyrovnávací konstantou β, přičemž se předpokládá, že $0 < \beta < 1$. Problém je však s vhodnou volbou vyrovnávací konstanty. Většina statistického a ekonometrického software umožňuje automatické vyhledání této konstanty.

3.2 Boxova-Jenkinsonova metodologie

Metody označované jako Boxovy-Jenkinsovy jsou vhodné v situacích, kdy nelze použít některé z výše uvedených metod. Modelují se jimi takové časové řady, které se chovají naprosto nesystematicky, rychle mění svůj charakter a úroveň, a nelze v nich vysledovat žádnou systematickou složku. Tyto metody kladou, na rozdíl od výše uvedených metod, vyšší důraz na práci s náhodnou složkou, o které se předpokládá, že je tvořena navzájem korelovanými hodnotami. Tato metoda je založena na práci s autokorelační a parciální autokorelační funkcí. Tyto dvě funkce vykazují zcela charakteristický průběh pro různé typy modelů. Celá Boxova-Jenkinsonova metodologie je náročná po stránce teoretické a vyžaduje velmi dobré znalosti z oblasti pravděpodobnosti a matematické statistiky.

Boxova-Jenkinsonova metodologie se vymyká klasickému dekompozičnímu pojetí modelu časové řady. Tato metoda patří mezi tzv. adaptivní přístup, to znamená, že je schopná přizpůsobovat se změnám v průběhu časových řad, je způsobem modelování časových řad i v ekonomikách procházejících rychlými změnami, které běžné a jednodušší metody nejsou schopny zachytit. Tyto modely patří mezi stochastické, které mohou modelovat trend i sezonnost. Jsou flexibilnější než klasické dekompoziční modely, avšak jejich nevýhodou je potřeba poměrné dlouhé časové řady (minimálně 50 pozorování) a náročnost zpracování a zejména velmi náročná interpretace výsledných modelů.

Boxova-Jenkinsonova metodologie vychází z přístupu, který říká, že časovou řadu je možno modelovat pomocí vzájemně nekorelovaných náhodných veličin s nulovou střední hodnotou a konstantním rozptylem, což je označováno jako bílý šum (white noise). V Boxově-Jenkinsonově metodologii lze modelovat pouze stacionární časové řady. Další požadovanou vlastnosti je invertibilita procesů (vzájemná zaměnitelnost různých procesů). Mezi důležité ukazatele v tomto modelu jsou autokorelační a parciální autokorelační funkce a jejich chování.

Autokorelace znamená závislost mezi členy jedné časové řady. Protože ale pro konkrétní časovou řadu skutečnou (teoretickou autokorelační funkci neznáme, používáme odhadnutou autokorelační funkci, která se ve výstupech označuje jako ACF a obdobně získáme i odhadnutou parciální autokorelační funkci značenou PACF).

Pokud je časová řada náhodně uspořádána (je náhodná), pak korelace párových dat y_t a y_{t-1} jsou blízké nule. Tedy autokorelační koeficienty za sebou následujících hodnot časové řady blízké nule znamenají, že časová řada není lineárně závislá. O tom, jestliže je časová řada náhodně uspořádána, se přesvědčíme z grafického znázornění ACF. Testem nulovosti ACF jsou dvojnásobky vypočtené směrodatné chyby odhadu autokorelačních koeficientů, které jsou na grafu vyznačeny čárkovaně.

V Boxově-Jenkinsonově metodologii rozlišujeme při práci s časovou řadou následující kroky:

- transformace dat časové řady tak, abychom odstranily případný trend a sezonnost v datech,
- identifikace modelu, což je rozhodnutí, jaký typ modelu vybrat a explicitně určit řád modelu,
- odhad parametrů modelu,
- diagnostické ověřování modelu, které má potvrdit nebo zamítnout vhodnost sestrojeného modelu,
- výpočet předpovědi pro vybraný a ověřený model.

K ZAPAMATOVÁNÍ

Základní typy Boxových-Jenkinsonových modelů pro časové řady bez sezónnosti jsou:

- MA(q) – proces klouzavých průměrů řádu q,
- AR(p) – autoregresní proces řádu p,
- ARMA(p,q) – smíšený proces řádu p a q,
- ARIMA(p,d,q) – smíšené integrované modely řádu p, d, q, kde d označuje řád diferencování potřebného dosažení stacionarity.

1. Proces klouzavých průměrů MA(1)

Tento proces má tvar:

$$y_t = e_t + \theta_1 e_{t-1} \tag{14}$$

kde:

- y_t jsou pozorování časové řady,
- θ_1 je parametr,
- e_t je náhodná složka.

Proces klouzavých průměrů je vždy stacionární a invertibilní za podmínky, že $|\theta_1| < 1$. ACF procesu MA(1) má pouze 1 významný autokorelační koeficient, což lze identifikovat z grafu ACF. Graf PACF procesu MA(1) má tvar geometricky klesající posloupnosti.

2. Autoregresní proces AR(1)

Tento proces má tvar:

$$y_t = \phi_1 y_{t-1} + e_t \tag{15}$$

kde:

- y_t jsou pozorování časové řady,
- ϕ_1 je parametr,
- e_t je náhodná složka.

Autoregresní proces je invertibitlní a podmínkou stacionarity je $|\phi_1| < 1$. ACF procesu AR (1) má tvar geometrické posloupnosti klesající v absolutní hodnotě k nule, PACF má významný pouze jeden parciální autokorelační koeficient.

3. Smíšený proces ARMA (1,1)

Tento proces má tvar:

$$y_t = \phi_1 y_{t-1} + e_t + \theta_1 e_{t-1} \tag{16}$$

Podmínkou stacionarity procesu je $|\theta_1| < 1$ a $|\phi_1| < 1$. ACF stacionárního procesu ARMA (1,1) geometricky klesá k nule a PACF také geometricky klesá k nule.

4. Smíšený integrovaný proces ARIMA (1,1,1)

Modely ARIMA umožňují pracovat s časovými řadami i takových jevů a procesů, u nichž může docházet ke změnám úrovně hodnot časové řady, tzn., že mají trend, a tyto změny mohou dokonce mít i nesystematický charakter. ARIMA modely jsou tedy schopné stochasticky modelovat nejen náhodné výkyvy, ale i trendovou složku. Parametr modelu $d = 1$ pak představuje jedno diferencování, které bylo nutné provést k tomu, abychom původně nestacionární řadu (většinou s trendem) převedli na řadu stacionární.

Vlastní konstrukce modelu ARIMA (1,1,1) probíhá tak, že původní řadu y_t diferencujeme, a tím ji převede man stacionární řadu z_t, pro kterou pak sestrojíme model ARMA (1,1).

V tabulce jsou uvedeny podmínky stacionarity a invertibility a chování funkcí ACF a PACF.

Model	Podmínky stacionarity	Podmínky invertibility	ACF	PACF
ARIMA (1,0,0) nebo AR (1)	$\lvert\phi_1\rvert < 1$	žádné	klesá	1 významný PAK
ARIMA (2,0,0) nebo AR (2)	$\phi_1 + \phi_2 < 1$ $\phi_2 - \phi_1 < 1$ $-1 < \phi_2 < 1$	žádné	klesá	2 významné PAK
ARIMA (0,0,1) nebo MA (1)	žádné	$\lvert\theta_1\rvert < 1$	1 významný AK	klesá
ARIMA (0,0,2) nebo MA(2)	žádné	$\theta_1 + \theta_2 < 1$ $\theta_2 - \theta_1 < 1$ $-1 < \theta_2 < 1$	2 významné AK	klesá
ARIMA (1,0,1) nebo ARMA (1,1)	$-1 < \phi_1 < 1$	$-1 < \theta_1 < 1$	klesá	klesá

Pozn. ACF je teoretická autokorelační funkce, PACF je teoretická parciální (dílčí) autokorelační funkce, AK jsou autokorelační koeficienty, PAK jsou parciální (dílčí) autokorelační koeficienty.

Identifikace modelu

Jednou z nejtěžších úloh při výstavbě Boxových-Jenkinsových modelů je jejich identifikace. Tato úloha spočívá v rozhodnutí, jaký typ modelu vybrat. Jde o relativně náročnou činnost, jenž je v mnoha případech závislá na zkušenostech ekonometra. Identifikace je přitom pouze první fází konstrukce modelů, neboť identifikovaný model je třeba ještě ověřit a upravit. Dále budou popsány jednotlivé kroky identifikace modelu. Uvažujme model ARIMA (p,d,q).

1. Nejprve je vhodné prozkoumat graf časové řady. V mnoha případech je možné na první pohled rozpoznat přítomnost trendu. V této fázi jde především o subjektivní zhodnocení situace.
2. Dalším krokem je výpočet odhadů ACF a PACF původní časové řady. Na jejich základě je možné potvrdit, že časovou řadu je třeba stacionarizovat (v případě, že hodnoty výběrové ACF a PACF v prvním zpoždění jsou velmi blízké jedné a ostatní hodnoty výběrové ACF klesají pomalu).
3. Po stacionarizaci časové řady se použijí výběrové ACF a PACF pro identifikaci modelů AR a MA (nalezení hodnot p a q). Tato identifikace je založena na principu podobnosti výběrových ACF a PACF s teoretickými ACF a PACF.

3.2.1 Sezonní model

Výše uvedený model se týkal jednoduššího případu časových řad ročních údajů, tedy řad bez sezonnosti. V ekonomické oblasti mají význam i časové řady s periodicitou kratší než jeden rok, kde se často objevuje vliv sezonnosti. Tomuto tématu se budeme věnovat pouze okrajově.

V případě analýzy a prognózy časových řad se sezonní složkou je celý postup trochu komplikovanější. Uvažujme sezonní model SARIMA, který je označen SARIMA(p, d, q)x(P, D, Q)s. V tomto modelu řády sezonní složky model P, D, Q mají stejný význam jako nesezonní rády p, d, q, které byly popsány výše a s značí sezonnost. Např. s = 4 je pro řady se čtvrtletní periodicitou a s = 12 je zase pro řady s roční periodicitou.

Identifikace řádů SARIMA modelu je komplikovanější a vyžaduje zkušenosti, zpravidla se provádí na základě grafického zobrazení odhadnuté autokorelační funkce. Potřeba sezonního diferencování je identifikována v případě, že ACF má lokální maxima v bodech, které jsou *S*-násobky sezonnosti bez ohledu na průběh ACF mezi těmito body.

Sezónní očišťování

Pro možnost průběžného srovnávání po sobě jdoucích údajů v časové řadě a pro operativní ekonomickou analýzu je nutné údaje časové řady sezónně očišťovat. Jde o modelové rozdělení časové řady na složku trendovou, sezónní a náhodnou, kdy prvořadým úkolem je zbavit časovou řadu sezónní složky, ale přitom ponechat v modelu složku trendovou, příp. cyklickou. Postup se týká pouze periodicity. Po jejím odstranění ještě sledujeme, zda v ekonomické časové řadě nezbyla složka cyklická.

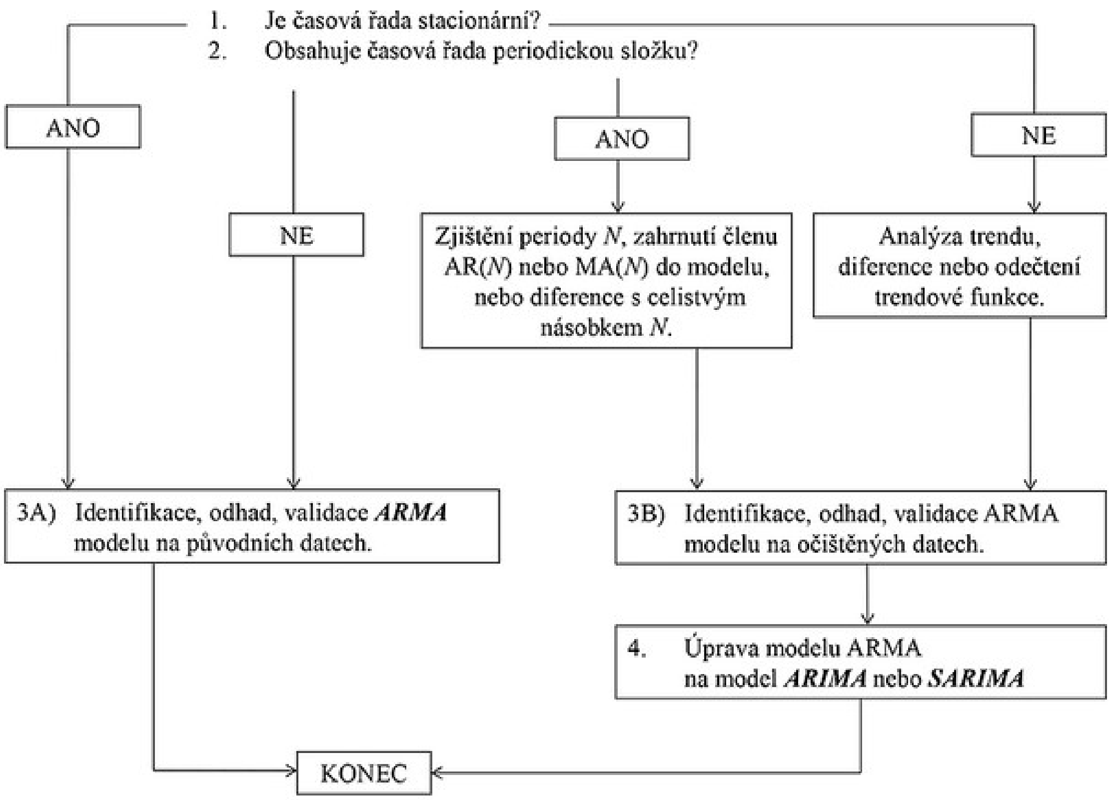
Metod sezónního očišťování je mnoho. Postupy zpravidla vycházejí z různých typů klouzavých průměrů, případně z jejich kombinací. Klouzavý průměr totiž eliminuje výrazně z časové řady ty složky, jejichž perioda nepřesahuje počet pozorování tvořící délku tohoto klouzavého průměru. Kromě toho byly vyvinuty speciální filtry, které jsou schopny časovou řadu dokonale očistit. Např. metoda Census X-12. Tato metoda využívá klouzavé průměry různých délek a na filtrovanou řadu aplikuje několik klouzavých průměrů postupně za sebou. Metoda je také schopna současně vyrovnávat nepravidelnosti způsobené kalendářem a eliminovat extrémní hodnoty, pokud se v řadě vyskytují náhodně.

Nejpoužívanější odhadová kritéria

- Kritérium AIC – název tohoto kritéria se většinou interpretuje na základě anglického Akaike's Information Criterion. Vzhledem ke svému jednoduchému tvaru patří toto kritérium k jednomu z nejpoužívanějších. Při jeho praktických aplikacích a na základě simulací však bylo zjištěno, že toto kritérium ve značném počtu případů vede k přeceňování skutečného řádu modelu.
- Kritérium BIC – název tohoto kritéria je odvozen z anglického Bayesian Information Criterion.

- Kritérium HQ – název kritéria je označen dle autorů, kteří ho navrhli, tedy Hannan a Quinn.
- Kritérium CAT – název tohoto kritéria vznikl z anglického Criterion Autoregressive Transfer function, a bylo navrženo pro stanovení řádu autoregresního modelu.

V praxi bývají mezi autory nejvíce oblíbená zejména kritéria AIC a BIC, která bývají s úspěchem využívána pro hodnocení kvality ARIMA modelů při modelování a předpovídání v časových řadách nejrůznějších ekonomických ukazatelů.



Zdroj: Croarkin et al. (2010)

Úprava modelu ARMA na model ARIMA nebo SARIMA v případech konstrukce modelu na diferencovaných datech. Diferencování časové řady je doporučeno podle Box-Jenkinsovy metodiky v případech, kdy je časová řada nestacionární nebo se v ní vyskytuje významná periodická složka.

DALŠÍ ZDROJE

Řešený vzorový příklad na ARMA, ARIMA či SARMA model najdete ve videonahrávce v LMS Moodle.

3.3 Modely vícerozměrných časových řad

Na rozdíl od modelů simultánně závislých rovnic, vyžadujících v souladu s ekonomickou teorií *apriori* klasifikaci proměnných systému na endogenní a exogenní, modely vektorových autoregresí (VAR) mezi nimi nerozlišují, takže všechny proměnné v nich obsažené jsou endogenní. Jednotlivé endogenní proměnné VAR modelu jsou autoregresními funkcemi svých zpožděných hodnot a funkcemi minulých hodnot ostatních endogenních proměnných modelu. Obsahuje-li VAR model některé exogenní proměnné, jde zpravidla o trendové proměnné, resp. Sezónní umělé proměnné. Specifikace modelu VAR vychází ze skutečnosti, že zahrnuté proměnné mají shodný trend a jsou autokorelované. Předpokládá se, že všechny proměnné vektorových autoregresí jsou stacionární.

VAR modely přestavují zobecnění jednorovnicových autoregresních (AR) modelů na časové řady několika proměnných a jejich předností je názorné vystižení dynamických interakcí proměnných zahrnutých v soustavě. VAR se dají pokládat za alternativu k rozsáhlým modelům strukturních simultánních rovnic, zejména v prognózování (Hušek, 2009).

VAR modely se využívají zejména k predikci a k testování Grangerovy kauzality, při simulaci efektů nástrojů hospodářské politiky, v dynamické analýze funkcí odezvy (impulse response function) a k dekompozici rozptylu chyb předpovědi na období v makroekonomické a finanční analýze.

3.3.1 MODELY VEKTOROVÝCH AUTOREGRESÍ

Vektorová autoregrese je vícerovnicový regresní model, který je spojením modelů časových řad jedné proměnné a strukturních modelů simultánních rovnic. Nejjednodušší vektorovou autoregresí je dvourovnicový VAR model, který obsahuje dvě endogenní proměnné, jejich běžné hodnoty jsou lineárními funkcemi p zpožděných hodnot obou proměnných a náhodných složek. Zvolíme-li délku zpoždění $p = 1$, pak vektorovou autoregresi 1. řádu, nebo-li model VAR (1) lze zapsat jako:

$$y_t = \varpi_1 + B_{11} y_{t-1} + \alpha_{11} x_{t-1} + v_{t1} \tag{17}$$

$$x_t = \varpi_2 + B_{21} y_{t-1} + \alpha_{21} x_{t-1} + v_{t2} \tag{18}$$

Tento zápis je standardní tvar modelu. Rovnice jsou zobecněním modelu AR (1).

Specifikace modelu VAR

Konstrukci VAR modelů lze rozdělit do několik etap, která na sebe navazují. Před vlastní specifikací modelů je nutné u většiny časových řad ekonomických proměnných dosáhnout nejprve jejich stacionárnosti. Obsahují-li časové řady proměnných deterministický trend, zajistíme jejich stacionaritu zahrnutím trendové proměnné jako funkce času do příslušné rovnice modelu. Postup pro testování stacionarity byl popsán v podkapitole 2.1.5.

Zahrnutí diferenčně stacionárních proměnných do modelu VAR však není adekvátní, existuje-li mezi nimi kointegrační vztah (bude popsáno v 5. kapitole), který je I(0), protože vynecháním kointegrační kombinace modelu VAR vzniká specifikační chyba. Správné je proto použít model korekce chyby (viz kapitola 5.2).

Vlastní specifikace modelu VAR spočívá v těchto krocích:

1) Výběr proměnných modelu a volba maximální délky zpoždění

V této fázi specifikace se využívá apriorní ekonomické nebo finanční informace a vhodná testovací kritéria. Při volbě optimální délky zpoždění VAR modelu odhadneme metodou nejmenších čtverců rezidua modelů pro různou délku zpoždění od 1 do k. K výběru optimální délky zpoždění p, využíváme zobecněná vícerozměrná informační kritéria, modifikovaná pro vícerozměrné modely za předpokladu normality. Avšak informační kritéria jsou platná za předpokladu správně specifikovaného modelu.

2) Zjednodušení a restrikce modelu

V konkrétních VAR modelech musíme apriori aplikovat některá omezení, protože existují zpravidla limity, a to jak počtu proměnných zahrnutých do modelu, tak i maximální délky zpoždění. Při volbě počtu proměnných (m) i maximální délky zpoždění (p) je nutno korigovat snahu ekonometra o co nejvyšší hodnoty m i p vzhledem k rozsahu disponibilního výběru pozorování, protože s rostoucím počtem proměnných a zvětšující se délkou maximálního zpoždění roste rychle počet odhadovaných parametrů VAR modelu a tím klesá počet stupňů volnosti. Proto v průběhu specifikace je nutno v modelech VAR apriori omezit počet zahrnutých proměnných i délku zpoždění. Optimální délka zpoždění zároveň vede k odhadům reziduí, nezatížených významnou autokorelací a zlepšuje i predikční vlastnosti VAR modelů.

3) Ortogonalizace náhodných složek

Vektor náhodných složek neomezeného VAR(p) modelu obsahuje dle předpokladu obvykle vzájemně zkorelované náhodné složky, které však mohou vykazovat i autokorelaci, takže jejich kovarianční matice není diagonální. Ortogonalizace je transformace proměnných (či náhodných složek) na nekorelované.

K ZAPAMATOVÁNÍ

Konstrukce modelu VAR

1. Jak již bylo uvedeno výše, před specifikací modelu je nutné dosáhnout stacionárních časových řad.

2. V dalším kroku následuje identifikace řádu modelu. V praxi se k určení řádu p využívají identifikační procedury pomocí statistických testů nebo informačních kritérií.

3. Samotný odhad modelu, kdy konzistentní a asymptoticky normálně rozdělené odhadované funkce parametrů lze získat klasickou metodou nejmenších čtverců nebo lze použít metodu maximální věrohodnosti.

4. Diagnostická kontrola modelu, kdy je nejdůležitější kontrola stacionarity odhadnutého modelu. Jedná se tedy o splnění podmínky, kdy hodnoty kořenů odhadnutého autoregresního polynomu leží vně jednotkového kruhu v komplexní rovině (či převrácené kořeny odhadnutého autoregresního polynomu leží uvnitř jednotkového kruhu). Dále na časové řady aplikujeme testy nekorelovanosti v odhadnuté reziduální složce tzn. ve vypočteném bílém šumu.

PŘEDNOSTI A PROBLÉMY MODELŮ VAR

K pozitivním stránkám VAR modelů patří především jejich jednoduchost. Při specifikaci VAR není nutno rozhodovat, které proměnné jsou endogenní nebo exogenní, resp. predeterminované. Také odhady jejich parametrů lze získat běžně používanou MNČ aplikovanou na každou rovnici soustavy zvlášť, přičemž při dodržení požadavků Gausse a Markova jsou zachovány optimální vlastnosti odhadů.

Jednou z výhod modelů VAR je i to, že umožňují testování směru kauzality, konkrétně Grangerovu kauzalitu, jejíž existence a využití vede к zpřesnění ekonometrických, tj. kvantitativních předpovědí.

Obecně platí, že předpovědi ekonomických veličin na základě VAR modelů jsou v řadě případů reálnější nežli např. při predikci pomocí rozsáhlých a složitějších modelů simultánně závislých rovnic.

Analýza impulzních odezev umožňuje sledování reakce libovolné proměnné na exogenní jednotkový šok zvolené proměnné. Ortogonalizované funkce odezvy, spočtené z VAR modelů umožňují při dynamické analýze efektů nástrojů hospodářské politiky stanovit odezvu (reakci) libovolné jedné proměnné modelovaného systému na exogenní jednotkový šok (impuls) jiné proměnné v čase. Odezvy na jednotkové šoky v jednotlivých okamžicích tvoří posloupnost, kterou lze znázornit při využití řady softwarových programů graficky a sledovat tak způsob i rychlost odeznění vyvolaných reakcí.

Na druhé straně můžeme zmínit i problémy spojené s využíváním modelů VAR. Jednou z výhrad je, že v důsledku svého a teoretického charakteru nejsou VAR modely vhodné např. k volbě nástrojů ekonomické politiky. Dalším negativem VAR modelu je technická náročnost aplikace tohoto modelu. Dále můžeme uvést, že problémy s určením řádu modelu, velký počet parametrů i při nižších řádech modelu a úprava proměnných kvůli předpokladu stacionarity může způsobit ztrátu důležitých vypovídacích schopností mezi jednotlivými časovými řadami. Další kritikou VAR modelů je nevhodnost modelů k volbě nástrojů hospodářské politiky z titulu převážně empirického charakterů a nezahrnutí teoretických východisek. Dalším negativem může být neovlivňování proměnných navzájem při aplikaci ortogonalizovaného šoku v jedné proměnné. Pro častý případ výběrů malých rozsahů bývají VAR modely přeparametrizovány, což snižuje potřebný počet stupňů volnosti a výsledkem jsou statisticky nevýznamné odhady parametrů. Testy kauzality zase nedokáží interpretovat význam odhadnutých parametrů ani délku jednotlivých impulzů.

Přes tyto výhrady jsou VAR běžně využívány jako protiváha a konkurence k systémům simultánně závislých rovnic, zejména pro účely predikce a při využití a analýze nestacionárních časových řad.

Využití VAR modelů

Vektorové autoregresní procesy jsou v praxi často využívány ekonometry i ekonomy k popisu a analýze makroekonomických časových řad pro svou flexibilitu, jednoduchý způsob odhadu a obvykle dobrou shodu s makroekonomickými daty. Jejich obliba a atraktivnost jsou hlavně díky skutečnosti, že umožňují kombinovat informace z dat dlouhodobého a krátkodobého charakteru, a sice využitím vlastnosti kointegrace ekonomických proměnných.

VAR modely jsou v praxi využívány v řadě případů. Významné využití mají VAR modely při analýze impulzní odezvy. Dále je možné VAR modely použít k předpovědím, Grangerove kauzality či rozkladu rozptylu.

Rozklad rozptylu zkoumá vliv inovací v jednotlivých rovnicích na vybranou proměnnou. Je to metoda, která určuje kolik z rozptylu chyby předpovědi o h krocích dopředu pro vybraný vysvětlovaný determinant, je vysvětleno inovací v jednotlivých rovnicích (h = 1, 2, ...).

3.3.2 Analýza a testy Grangerovy kauzality

Jednou z možností aplikace VAR modelů je jejich využití k testování směru kauzální závislosti. V ekonometrii je chápána kauzalita jako schopnost určité proměnné predikovat jinou proměnnou. Tento koncept kauzality zavedli Granger (1969) a Sims (1972). V jejich pojetí je testování kauzality ověřením toho, zda změny určité proměnné předcházejí změně jiné proměnné a ne, která proměnná je příčinou a která je následkem.

Granger (1969) rozpracoval poměrně jednoduchý test, pomocí kterého lze definovat jeho pojetí kauzality tak, že pro dvě stacionární proměnné y_t a x_t je např. proměnná x_t ve vztahu Grangerovy kauzality k proměnné y_t, zvyšuje-li využití zpožděných hodnot x_t jako regresorů přesnost predikce y_t. Nezlepší-li zahrnutí zpožděných hodnot x_t do regrese přesnost predikce y_t, resp. Nezmění-li se předpověď proměnné y_t po vynechání minulých hodnot x_t, pak x_t nepůsobí ve smyslu Grangerovy kauzality na y_t.

K ZAPAMATOVÁNÍ

Konkrétně za Grengoru kauzalitu dvou simultánně závislých stacionárních časových řad označíme takový jejich vztah, kdy běžná hodnota jedné řada (proměnná y_t) je zkorelována se zpožděnými hodnotami druhé rady (proměnná x_t), neboli změny x předcházejí změnám y.

Při testování Grangerovy kauzality pro dvojici stacionárních proměnných y_t a x_t odhadneme nejprve dvourovnicový VAR(p) model bez úrovňových konstant

$$y_t = \sum_{i=1}^{p} \alpha_i y_{t-1} + \sum_{i=1}^{p} B_i x_{t-1} + u_{t1}, \tag{19}$$

$$x_t = \sum_{i=1}^{p} \gamma_i y_{t-1} + \sum_{i=1}^{p} \delta_i x_{t-1} + u_{t2,} \tag{20}$$

kdy náhodné složky u_{t1} a u_{t2} jsou nezkorelované.

V dalším kroku ověříme statistickou významnost odhadovaných koeficientů a testujeme nejprve důsledek vynechání zpožděných hodnot proměnné x v rovnici (19) a poté vynecháním minulých hodnot y v rovnici (20).

Pro rovnici (19) ukážeme postup testování. Nulová hypotéza zde zní, že x_t nepodmiňuje v Grangerově pojetí proměnnou y_t. Při testování nulové hypotézy nejdříve vyjádříme např. lineární regresi proměnné y_t na jejich zpožděných hodnotách vztahem

$$y_t = \sum_{i=1}^{p} \alpha_i y_{t-1} + u_{t1}, \tag{21}$$

Pokud Vás zajímá konkrétní teoretický postup a více informací doporučuji nastudovat další zdroje, např. Hušek (2009).

3.3.3 FUNKCE IMPULS-ODEZVY A JEJICH INTERPRETACE

Souhrnnou informací v modelech časových řad jsou AR či VAR modely získáme z funkcí odezvy někdy označované jako odezva na impuls (impulse response function,), které měří reakci neočekávaného šoku na běžnou hodnotu i budoucí hodnoty libovolné proměnné časové řady. Např. neočekávané zvýšení klíčových úrokových sazeb ČBN nebo nabídky peněz má vliv na míru inflace, tak i na ekonomický růst. Velikost těchto reakcí lze aproximativně zjistit právě pomocí odhadnutých funkcí odezvy (Hušek, 2009).

Arlt a Arltová (2009) hovoří o analýze impulsní reakce jako o zkoumání vztahu mezi dvěma jednorozměrnými časovými řadami ve vícerozměrném systému. Jde totiž o rozbor reakce v jedné časové řadě vyvolanou impulsem v jiné časové řadě, ale systém zároveň obsahuje mnoho dalších časových řad. Cipra (2014) definuje odezvu na impuls jako reakci zvolené vysvětlované proměnné na impuls v jiné rovnici VAR modelu.

Jednotlivé veličiny náhodného vektoru jsou korelované, což může být problémem. Nelze uvažovat, že pokud jedna veličina bude mít v daném čase hodnotu jedna, ostatní veličiny budou ve stejném čase nulové. Doporučuje se tedy analýzu impulzních odezev provádět v systému, ve kterém jsou veličiny nesystematické složky nekorelované. Pokud je ortogonalizovaném systému změněna jedna složka vektoru, nebude to mít vliv na jiné složky.

DALŠÍ ZDROJE

Řešený vzorový příklad na VAR model včetně možností jeho využití naleznete ve videonahrávce v LMS Moodle.

SHRNUTÍ KAPITOLY

V kapitole byly popsány modely jednorozměrných a vícerozměrných časových řad. Nejprve byly popsány jednorozměrné stacionární a nestacionární časové řady. Časová řada znamená posloupnost hodnot proměnné uspořádaný v čase. Boxova-Jenkinsonova metodologie se využívá k modelování časových řad, které se chovají nesystematicky, mění svůj charakter a nelze v nich vysledovat systematickou složku. Tyto metody kladou důraz na práci s náhodnou složkou, o které se předpokládá, že je tvořena korelovanými hodnotami. Boxova-Jenkinsonova metodologie vychází z přístupu, který říká, že časovou řadu je možno modelovat pomocí vzájemně nekorelovaných náhodných veličin s nulovou střední hodnotou a konstantním rozptylem. Modelovat je možné pouze stacionární časové řady.

Modely vícerozměrných časových řad (VAR modely) je vícerovnicový (zpravidla dvourovnicový) regresní model. Konstrukce těchto modelů je zpravidla rozdělena do několika kroků, nejprve je nutný výběr proměnných modelu a určení délky zpoždění. Je nutné mít stacionární časové řady. Následuje identifikace řádu modelu a odhad modelu. V poslední části je nutná diagnostika odhadnutého modelu, konkrétně kontrola stacionarita odhadnutého modelu a testy nekorelovanosti v odhadnuté reziduální složce. Diagnostika modelu je stejná pro modely jednorovnicových i vícerovnicových časových řad.

OTÁZKY

1. Boxova-Jenkinsonova metodologie patří mezi adaptivní modely. ANO x NE
2. V Boxově-Jenkinsonově metodologii můžeme modelovat časové řady, které jsou:
   a) stacionární
   b) nestacionární
   c) invertibilní
3. V modelu ARIMA parametr modelu *d* představuje:
   a) logaritmování
   b) diferencování
   c) korelování
4. VAR modely lze využít k analýze funkcí odezvy. ANO x NE
5. VAR model je vhodné použít, pokud zjistíme, že existuje kointegrační vztah mezi proměnnými. ANO x NE

ODPOVĚDI

1. ANO
2. a)
3. b)
4. ANO
5. NE

4 KAUZALITA VE FINANČNÍCH A ČASOVÝCH ŘADÁCH

RYCHLÝ NÁHLED KAPITOLY

Čtvrtá kapitola studijní opory se zabývá kauzalitou ve finančních časových řadách. Zkoumání kauzality je častým jevem, kterým se analytik zabývá. Proto bude v této kapitole zaměřena pozornost na hledání a hodnocení závislostí mezi veličinami. Nejprve bude v práci představena korelační analýza, její výhody a nedostatky. V další části bude představena Grangerova kauzality. Praktická ukázka korelační analýzy či Grangerovy kauzality v EViews je součástí kapitoly.

CÍLE KAPITOLY

- Definovat kauzalitu časových řad.
- Popsat korelační analýzu a korelační koeficient.
- Definovat Grangerovu kauzalitu.

KLÍČOVÁ SLOVA KAPITOLY

Kauzalita, Grangerova kauzalita, korelační analýza, korelační matice.

4.1 Kauzalita

Žádný ekonomický jev nevzniká a neprobíhá libovolně, ale je ve vztahu k jiným jevům a nemůže být správně pochopen, je-li z podmínek, které determinují jeho existenci, vytržen.

Nyní se tedy zaměříme na hledání, zkoumání a hodnocení souvislostí (závislostí) mezi dvěma a více veličinami. Příčinnou (kauzální) závislostí mezi např. dvěma proměnnými rozumíme situaci, kdy existence určitého jevu souvisí s existencí jevu jiného.

Z hlediska metody zkoumání je dobré rozlišit pevné a volné závislosti. Pevná závislost znamená, že výskyt jednoho jevu nutně odpovídá výskytu druhého jevu (a často naopak). Každé hodnotě jedné proměnné odpovídá jedna hodnota jiných proměnných (a podobně naopak). Volná závislost znamená, že výskyt jednoho jevu ovlivňuje výskyt druhého jevu v tom smyslu, že se zvýšila pravděpodobnost nastoupení druhého jevu při nastoupení jevu prvního. Tedy vztah, kdy hodnotám např. jedné proměnné odpovídají různé hodnoty jiné proměnné, lze hovořit o jakési obecné tendenci, která se projevuje při změnách hodnot těchto proměnných. V reálných empirických finančních situacích se setkáváme prakticky výhradně s volnými závislosti. Za obecnými tendencemi projevujícími se v souboru statistických údajů se mohou skrývat hlubší zákonitosti vztahů mezi veličinami.

Pro potřeby těchto metod je vhodné rozlišit jednostranné a vzájemné závislosti. Jednostrannými závislostmi se zabývá regresní analýza. Tématu regresní analýzy je věnovaná pozornost ve 2. kapitole. Jedná se o situaci, kdy proti sobě stojí vysvětlující proměnná a vysvětlovaná proměnná. Vzájemnými (většinou lineárními) závislostmi se zabývá korelační analýza. V korelační analýze se klade důraz více na intenzitu (sílu) vzájemného vztahu než, není zde zkoumání veličin ve směru příčina-následek. Avšak často dochází k prolínání obou přístupů. (Hindls – Statistika pro ekonomy). Korelační analýza se tedy zabývá oboustrannými závislostmi mezi veličinami X a Y.

HLAVNÍ ÚKOLY REGRESNÍ A KORELAČNÍ ANALÝZY

Hlavním úkolem regresní a korelační analýzy je pomoci k poznání příčinných vztahů mezi statistickými znaky. Východiskem k popisu statistických závislostí jsou pochopitelně statistické údaje. Statistický soubor n pozorování sledovaných statistických znaků můžeme získat různým způsobem.

Úkolem regresní a korelační analýzy je matematický popis systematických aspektů statistických závislostí. Snahou je nalézt takovou matematickou funkci, aby co nejlépe vyjadřovala charakter závislosti a co nejpřesněji zobrazovala průběh změn podmíněných průměrů závisle proměnné. Tato (svou podstatou hypotetická) matematická funkce se nazývá regresní funkce. Cílem regresní analýzy je co nejlepší přiblížení empirické (vypočítané) regresní funkce k hypotetické regresní funkci. S uvedeným hlavním úkolem regresní analýzy souvisí řada dílčích úkolů. Jmenujme některé z nich:

- shromáždit a matematicky formulovat představy o charakteru regresní funkce,
- formulovat představy (předpoklady) o souhrnném působení neuvažovaných statistických znaků,
- odhadnout empirickou regresní funkci na základě statistických pozorovaní,
- posoudit kvalitu empirické regresní funkce z hlediska důvodů a cílů statistického zjišťovaní.

Pokud se podaří vystihnout průběh závislosti "relativně nejlepší" regresní funkcí, je otázka kvality regresní funkce souběžná s problémem síly (intenzity, těsnosti) závislosti.

4.2 Korelační analýza

Korelace v tom nejobecnějším smyslu označuje míru stupně asociace dvou proměnných. Dvě proměnné jsou tedy korelované, pokud určité hodnoty jedné proměnné mají tendenci se vyskytovat společně s určitými hodnotami druhé proměnné.

K ZAPAMATOVÁNÍ

Avšak je podstatné zdůraznit, že korelace není totéž jako příčinná závislost. Pokud spolu dvě veličiny korelují, neznamená to nutně, že jedna ovlivňuje druhou.

Vzájemný vztah mezi veličinami je zjišťován na základě korelační analýzy. Pokud je mezi dvěma veličinami potvrzena statisticky významná korelace, je zároveň potvrzena jejich vzájemná závislost. Korelace je tedy lineární závislost mezi dvěma veličinami.

Z hlediska kauzality je podstatné zmínit, že právě korelační koeficient byl v minulosti používán jako argument pro kauzální závěry, jeho vysoká hodnota podle některých autorů jasně dokazuje, že X zapříčiňuje Y, kde X a Y jsou statistické znaky, charakterizující určitou veličinu a je dokonce určitým kvantifikátorem takového zdánlivého kauzálního vztahu. Avšak je nutné zdůraznit, že jsou to mylné teze. Dva hlavní důvody pro mylnost korelace poukazující na kauzalitu jsou, že:

- Vysoký korelační koeficient (v absolutní hodnotě) značí pouze statistickou závislost mez proměnnými (navíc pouze lineární), která může poukazovat např. jen na podobný trend obou proměnných, a nikoliv na vztah příčina–následek.
- Korelační koeficient, jak již bylo řečeno, je symetrický, a pokud by prokazoval kauzalitu, nebylo by jasné, zda X zapříčiňuje Y nebo Y zapříčiňuje X. Situace, kdy platí obě tvrzení (obousměrná kauzalita), se nevyskytuje příliš často (kromě triviálního případu X= Y).

Přestože kauzální vztah mezi dvěma proměnnými může znamenat vysokou hodnotu jejich korelačního koeficientu, ale tato hodnota není rozhodně postačujícím kritériem kauzality. Při zkoumání korelačních vztahů má rozhodující význam kvalitativní rozbor příslušného materiálu. Nemá smysl měřit závislost tam, kde na základě logické úvahy nemůže existovat.

Míru korelace vyjadřuje tzv. korelační koeficient. Korelační koeficient pro jednoduchou korelaci, tedy párový korelační koeficient, měří míru lineární závislosti mezi dvěma veličinami.

PEARSONŮV KORELAČNÍ KOEFICIENT

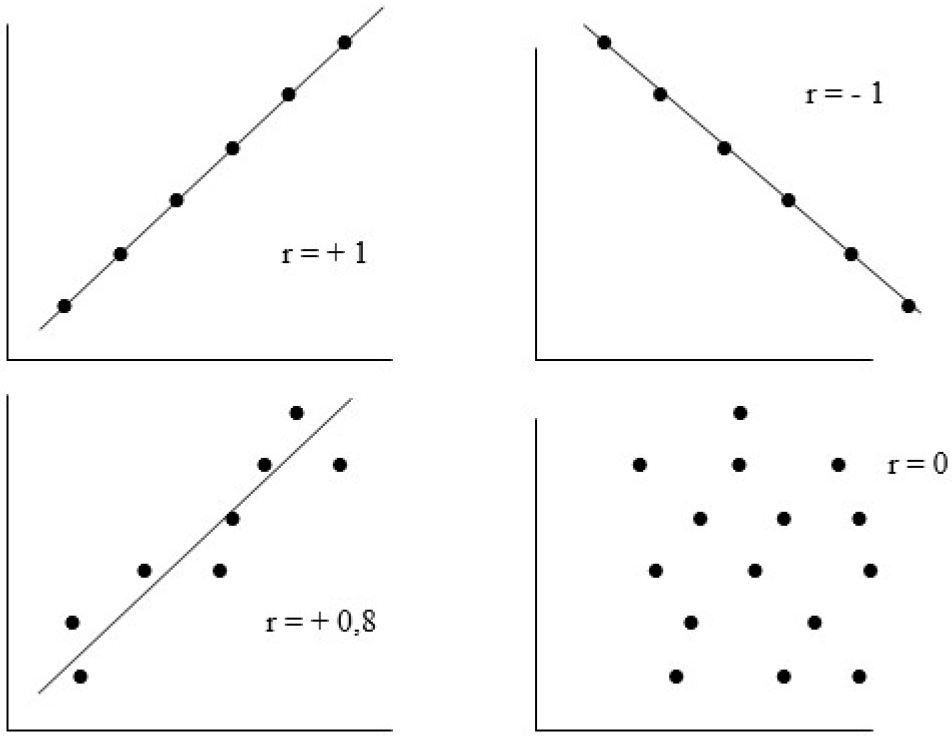
Tuto metodu lze použít za předpokladu, že rozdělení obou proměnných je normální a jejich vztah je přibližně lineární.

Korelační koeficient (r) nabývá hodnot od –1 do 1, kdy:

- $r = 0$, pak mezi proměnnými není lineární vztah nebo tento vztah zůstal na základě dat, které máme k dispozici, neprokázán;
- $r > 0$ značí přímou (pozitivní, kladná) korelaci, což znamená, že se veličiny vyvíjejí stejným směrem;
- $r = 1$ indikuje funkční pozitivní závislost, tedy mezi proměnnými existuje přesně lineární vztah;
- $r < 0$ nepřímá (negativní, záporná) korelace, tedy vzájemný inverzní vztah dvou veličin;
- $r = -1$ indikuje funkční negativní závislost, tedy mezi proměnnými existuje přesně lineární vztah.

Správná interpretace Pearsonova korelačního koeficientu předpokládá, že obě proměnné jsou náhodné veličiny a mají společné dvourozměrné normální rozdělení. Potom nulový korelační koeficient znamená, že veličiny jsou nezávislé. Pokud není splněn předpoklad dvourozměrné normality, z nulové hodnoty korelačního koeficientu nelze usuzovat na nic víc, než že veličiny jsou nekorelované.

Lineární závislost dvou statistických lze popsat vynesením proměnných do grafu. V případě korelace nestanovujeme rovnici přímky závislosti (to je úlohou lineární regrese), ale můžeme si přímku představit jako vyjádření lineárního vztahu a z odchylek bodů od přímky pak odhadnout míru tohoto vztahu.



Obrázek 4-1 Korelační koeficient

Interval	Síla vzájemné vazby
0,00 - 0,19	Velmi slabá
0,20 - 0,39	Slabá
0,40 - 0,59	Střední
0,60 - 0,79	Silná
0,80 - 1,00	Velmi silná (multikolinearita)

Tabulka 4-1 Kategorie síly vzájemné vazby Pearsonova korelačního koeficientu

Zdroj: Evans (1996, s. 130)

KORELAČNÍ MATICE

Základním prvkem korelační analýzy je korelační matice, kde je možné sledovat statistickou významnost vzájemného vztahu dvou proměnných na základě t-statistiky a pravděpodobnosti. Na průsečíku *i*-tého řádku a *j*-tého sloupce je tedy uveden korelační koeficient r_{ij} i-té a j-té proměnné. Korelační matice je čtvercová a na diagonále obsahuje jedničky.

SPEARMANŮV KORELAČNÍ KOEFICIENT POŘADÍ

Tento koeficient zachystuje monotónní vztahy, ne pouze lineární, ale obecně rostoucí či klesajcící a je rezistentní vůči odlehlým hodnotám. Spearmanovým korelačním koeficientem tedy měříme sílu vztahu X a Y, když nemůžeme předpokládat linearitu očekávaného vztahu nebo normální rozdělení proměnných X a Y. Závislost proměnných může mít obecně vzestupný nebo sestupný charakter.

ŘEŠENÁ ÚLOHA

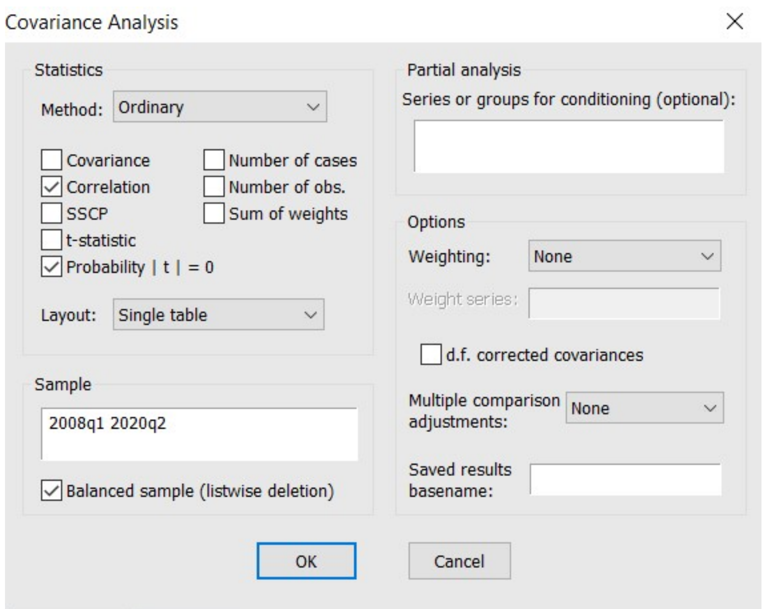
V příkladu použijeme data využitá v regresní analýze, konkrétně:

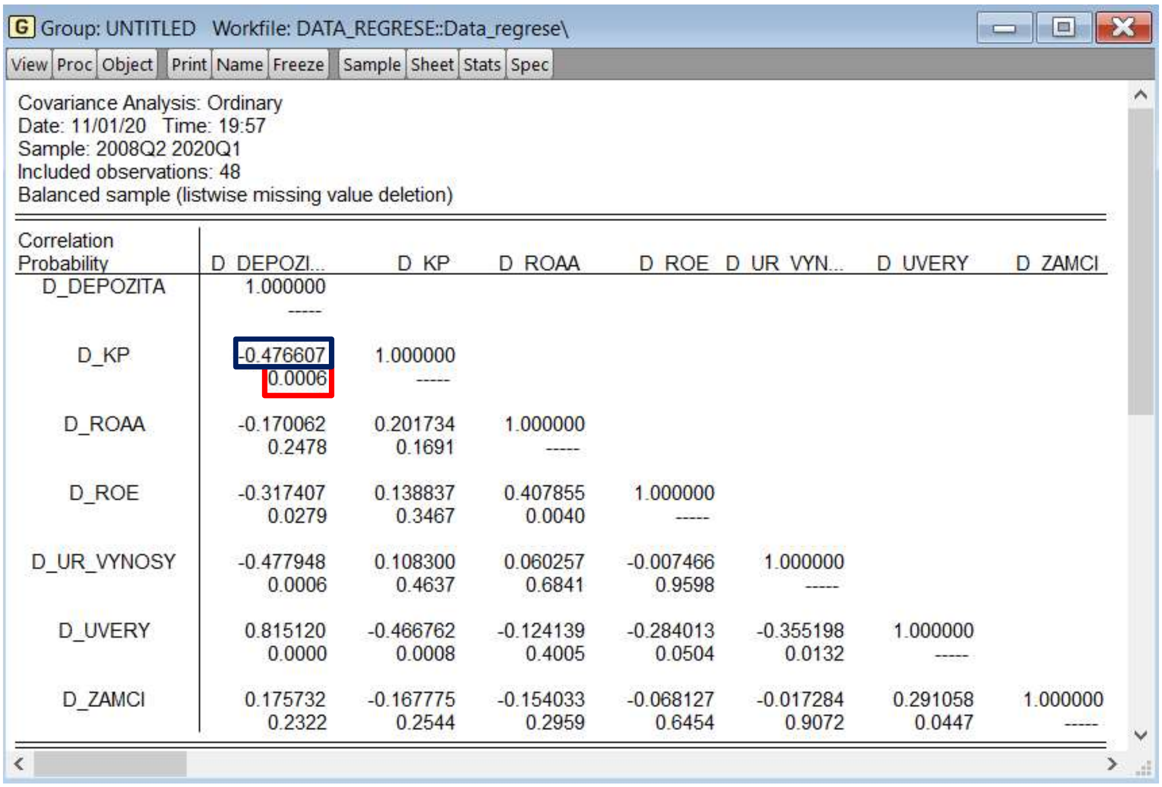
- klientská depozita,
- kapitálová přiměřenost,
- ROAA,
- ROE,
- úrokové výnosy,
- celkové poskytnuté úvěry,
- počet zaměstnanců.

Veškerá data jsou za český bankovní sektor z databáze ARAD v období 1. čtvrtletí 2008 – 2. čtvrtletí 2020. Jedná se o čtvrtletní data.

Nejprve je znázorněn korelační koeficient a korelační matice pro určení lineární závislosti mezi proměnnými.

Pracujeme se stacionárními daty. Otevřeme vybraná data jako skupinu a vybereme korelační koeficient (viz Řešená úloha v 2. kapitole).





Group: UNTITLED Workfile: DATA_REGRESE::Data_regrese\

View | Proc | Object | Print | Name | Freeze | Sample | Sheet | Stats | Spec

Covariance Analysis: Ordinary
Date: 11/01/20 Time: 19:57
Sample: 2008Q2 2020Q1
Included observations: 48
Balanced sample (listwise missing value deletion)

Correlation Probability	D_DEPOZI...	D_KP	D_ROAA	D_ROE	D_UR_VYN...	D_UVERY	D_ZAMCI
D_DEPOZITA	1.000000 -----						
D_KP	-0.476607 0.0006	1.000000 -----					
D_ROAA	-0.170062 0.2478	0.201734 0.1691	1.000000 -----				
D_ROE	-0.317407 0.0279	0.138837 0.3467	0.407855 0.0040	1.000000 -----			
D_UR_VYNOSY	-0.477948 0.0006	0.108300 0.4637	0.060257 0.6841	-0.007466 0.9598	1.000000 -----		
D_UVERY	0.815120 0.0000	-0.466762 0.0008	-0.124139 0.4005	-0.284013 0.0504	-0.355198 0.0132	1.000000 -----	
D_ZAMCI	0.175732 0.2322	-0.167775 0.2544	-0.154033 0.2959	-0.068127 0.6454	-0.017284 0.9072	0.291058 0.0447	1.000000 -----

V tomto případě testujeme nulovou hypotézu: H_0: ρ =0. Nulová hypotéza zní, že korelační koeficient je nulový. Pomocí *p*-hodnoty (hodnota pravděpodobností je pro každý znázorněna pár v červeném rámečku). Hodnota korelačního koeficientu je zvýrazněna v modrém rámečku.

4.3 Grangerova kauzalita

Jak bylo popsáno v předchozí podkapitole, korelační analýza se určuje lineární závislosti mezi dvěma proměnnými. Avšak jak bylo uvedeno, ne každá závislost musí být kauzální, tzn. vývoj jedné proměnné nemusí ovlivňovat druhou proměnnou. Existují dva způsoby kauzálních vztahů – bottom up a top down. První strategie bottom up předpokládá, že procesy, které generují hodnoty různých časových řad, nejsou na sobě závislé, ale také bere v úvahu, zda nejsou vzájemně závislé pouze některé konkrétní časové řady. Tato strategie je často využívána při testování kauzality a jako první ji navrhl Granger (1969). Druhá strategie top down předpokládá, že procesy, které generují hodnoty různých časových řad, nejsou na sobě závislé a bere v úvahu, jestli hodnoty některých časových řad jsou generovány nezávisle na ostatních zkoumaných proměnných. Tato strategie je využívána v případě aplikace modelů vektorové autoregrese. Obě zmíněné strategie lze využít při vyšetřování kauzálních vztahů mezi časovými řadami.

Jednou z možností aplikace VAR modelů je jejich využití k testování směru kauzální závislosti. V ekonometrii je chápana kauzalita jako schopnost určité proměnné predikovat jinou proměnnou. Tento koncept kauzality zavedli Granger (1969) a Sims (1972). V jejich pojetí je testování kauzality ověřením toho, zda změny určité proměnné předcházejí změně jiné proměnné a některá proměnná je příčinou a která následkem. (Hušek, 2009)

Granger tedy zkoumal kauzální vztahy mezi ekonomickými časovými řadami a definoval pojetí kauzality, při jehož ověřování lze použít modely VAR. Grangerova (1969) základní myšlenka kauzality zní: působí-li řada Z na řadu Y, pak by řada Z měla pomoci zlepšit předpovědi řady Y. K posouzení kauzálních vztahů v rámci modelu VAR je také určena analýza „Impuls-Reakce“, která informuje o reakci v jedné časové řadě vyvolané impulsem v jiné časové řadě.

Můžeme tedy říct, že Grangerův přístup spočívá v tom, že kauzalitu interpretuje jako schopnost vylepšit odhad budoucí hodnoty jedné složky vícerozměrné časové řady na základě minulých a současných hodnot jiné složky. Hovoříme-li tedy v této souvislosti o příčinnosti, míníme tím právě tolik, že se v budoucnosti první složky odráží minulost druhé. Nejedná se tedy nutně o přímé působení v klasickém slova smyslu.

V Grangerově testu kauzality se vychází z existence dvou stacionárních proměnných xta yt, kde proměnná yt je zkorelována se zpožděnými hodnotami proměnné xt. Z tohoto vztahu lze tvrdit, že změna proměnné xt předchází změnám proměnné yt. Vpřípadě Grangerovy kauzality jsou splněny následující podmínky:

- na větší přesnosti předpovědi yt se podílí zpožděná hodnota proměnné xt, a tak v regresi proměnné yt na jejích zpožděných hodnotách dochází k podstatnému zlepšení vypovídací schopnosti regresní závislosti při rozšíření vysvětlujících proměnných o běžná a minulá pozorování proměnné xt,
- proměnná yt nezvyšuje přesnost předpovědi proměnné xt, pokud by tomu tak bylo, znamenalo by to existenci další proměnné, která podmiňuje jak proměnnou yt, tak xt.

Z výše uvedených podmínek vyplývá, že nelze zaměňovat pojem Grangerova kauzalita s pojmem příčinná závislost, neboť podstatou Grangerovy kauzality je testování, zda změny určité proměnné předcházejí změně jiné proměnné a ne zjišťování, která veličina je příčinou a která následkem (Hušek, 2007).

Testování Grangerovy kauzality

U časových řad se předpokládá stacionarita. Pokud nejsou stacionární, používá se první diference proměnných. Prvním krokem testu Grangerovy kauzality je pro stacionární proměnné xt a yt odhad dvourovnicového VAR(p) modelu bez úrovňových konstant, který má tvar dle vzorců:

$$y_t = \sum_{i=1}^{p} \alpha_i y_{t-1} + \sum_{i=1}^{p} B_i x_{t-1} + u_{t1}\,, \tag{22}$$

$$x_t = \sum_{i=1}^{p} \gamma_i y_{t-1} + \sum_{i=1}^{p} \delta_i x_{t-1} + u_{t2,} \tag{23}$$

kde u_{t1} a u_{t2} představují náhodné složky, které nejsou korelované, α_i, B_i, γ_i a δ_i představují regresní koeficienty pro zpoždění i-tého řádu, p je maximální délka zpoždění, která není nijak omezena.

Dalším krokem je ověření statistické významnosti odhadnutých koeficientů, přičemž jako první se testuje důsledek vynechání zpožděných hodnot proměnné *x* dle rovnice (22), poté je proveden ten samý test pro proměnnou y podle rovnice (23). Dále uvedené vzorce pro testování se vztahují k proměnné *y*. Testovat Grangerovu kauzalitu lze pomocí t-testu či F-testu, a to za určení hypotéz:

- H0: x_t nepodmiňuje proměnnou y_t ve smyslu Grangerovy kauzality,
- H1: x_t podmiňuje proměnnou y_t ve smyslu Grangerovy kauzality.

Totožný postup platí i pro proměnnou x_t za určení následujících hypotéz:

- H0: y_t nepodmiňuje proměnnou x_t ve smyslu Grangerovy kauzality,
- H1: y_t podmiňuje proměnnou x_t ve smyslu Grangerovy kauzality.

Dle Grangerovy kauzality lze rozlišit několik výstupů:

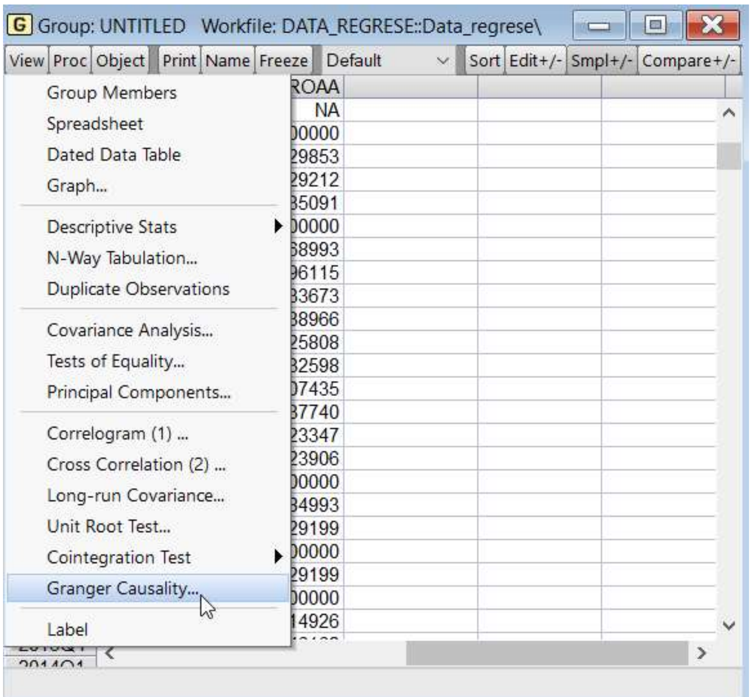
- Za prvé se jedná o jednosměrnou kauzalitu, kdy suma odhadnutých koeficientů zpožděných hodnot proměnné xt dle vzorce (22) a suma odhadnutých koeficientů proměnné yt dle vzorce (23) se rovná nule, tedy $\sum\beta i$≠0 a $\sum\delta i$=0. Z tohoto vztahu vyplývá, že proměnná x_t působí příčinně na proměnnou y_t, přičemž proměnná x_t není závislá na proměnné y_t.
- Druhým výstupem je také jednosměrná kauzalita, avšak v opačném tvaru, $\sum\beta i$=0 a $\sum\delta i$≠0. Proměnná y_t působí příčinně na proměnnou x_t, ale proměnná y_t není závislá na proměnné x_t.
- Třetí formou je oboustranná příčinná závislost, kde sumy koeficientů proměnných x_t a y_t se ani vjednom případě nerovnají nule. Proměnná y_t tedy závisí na proměnné x_t a také proměnná y_t působí příčinně na proměnnou x_t.
- Poslední formou je nezávislost mezi proměnnými, při které sumy koeficientů proměnných x_t a y_t se v obou případech rovnají nule. Poté platí, že proměnná x_t nepůsobí příčinně na proměnnou y_t a naopak.

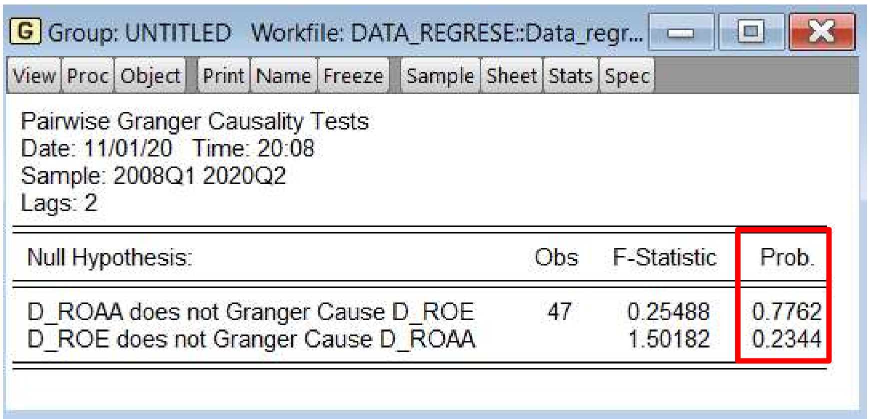
ENDOGENITA A EXOGENITA

Grangerova kauzalita je využívána nejen pro testování existence či neexistence příčinné souvislosti mezi stacionárními veličinami, ale také je vhodná pro testování exogenity. Existence nekauzality v Grangerově smyslu je totožná s exogenitou proměnných v různých ekonometrických modelech. S tímto tvrzením je třeba vysvětlit rozdíl mezi slabou, silnou a super exogenitou. Grangerova nekauzalita je jen postačující podmínkou exogenity, což odpovídá slabé exogenitě. V případě dvou proměnných x_t a y_t představuje x_t regresor y_t. Proměnná x_t je slabě exogenní v případě, že proměnná y_t ji současně nevysvětluje. Odhad parametrů modelu a testování se pak může provádět na základě hodnot proměnné x_t. O silně exogenní proměnnou se jedná v případě, že proměnná x_t není vysvětlena současnými a budoucími hodnotami proměnné y_t, a tudíž lze hovořit o neexistenci zpětné vazby. Pokud jsou regresní parametry proměnných x a y neměnné, a to ani v případě změny hodnoty proměnné x, pak proměnná x_t vystupuje jako super exogenní. Všechny tři formy exogenity lze využít v jiném kontextu. Slabá exogenita se využívá při odhadování a testování, silná pak při sestavování prognóz a předpovědí a super exogenita je využívána v rámci hospodářské politiky při provádění různých analýz (Hušek, 2007).

ŘEŠENÁ ÚLOHA

Data pro tuto řešenou úlohu jsou brána z předchozího příkladu. Pracujeme se stacionárními daty. Otevřeme zvolená data jako skupinu.





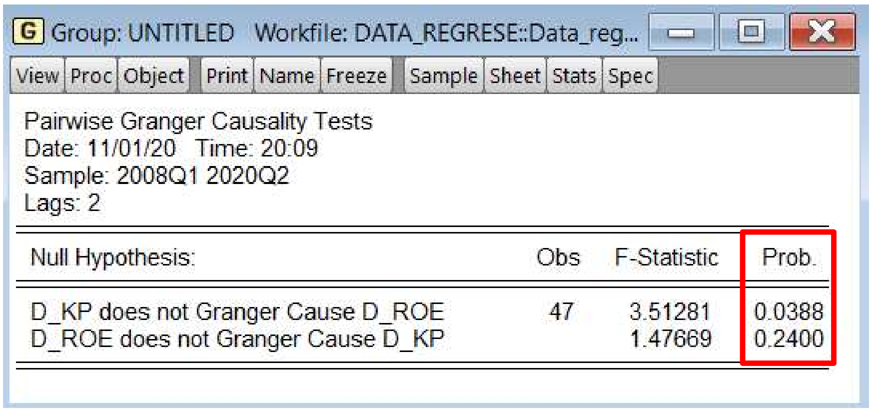
Group: UNTITLED Workfile: DATA_REGRESE::Data_regr...

View Proc Object Print Name Freeze Sample Sheet Stats Spec

Pairwise Granger Causality Tests
Date: 11/01/20 Time: 20:08
Sample: 2008Q1 2020Q2
Lags: 2

Null Hypothesis:	Obs	F-Statistic	Prob.
D_ROAA does not Granger Cause D_ROE	47	0.25488	0.7762
D_ROE does not Granger Cause D_ROAA		1.50182	0.2344

Testujeme zde nulovou hypotézu. Dle *p*-hodnoty nelze zamítnout nulovou hypotézu a můžeme říct, že hodnoty ROA nepřispívají k vývoji hodnot ROE. A naopak hodnoty ROE nepřispívají k vysvětlení vývoje hodnot ROA v českém bankovním sektoru.



Group: UNTITLED Workfile: DATA_REGRESE::Data_reg...

View Proc Object Print Name Freeze Sample Sheet Stats Spec

Pairwise Granger Causality Tests
Date: 11/01/20 Time: 20:09
Sample: 2008Q1 2020Q2
Lags: 2

Null Hypothesis:	Obs	F-Statistic	Prob.
D_KP does not Granger Cause D_ROE	47	3.51281	0.0388
D_ROE does not Granger Cause D_KP		1.47669	0.2400

Testujeme zde nulovou hypotézu. Dle *p*-hodnoty v zamítáme nulovou hypotézu a můžeme říci, že kapitálová přiměřenost přispívá k vysvětlení vývoje hodnot ROE v českém bankovním sektoru. Naopak v druhém případě nelze zamítnout nulovou hypotézu a můžeme říct, že hodnoty ROE nepřispívají k vývoji hodnot kapitálové přiměřenosti v českém bankovním sektoru.

Σ SHRNUTÍ KAPITOLY

V kapitole se čtenář seznámil s problematikou kauzality finančních řad. Pro určení lineární závislosti mezi dvěma proměnnými lze použít korelační koeficient. Korelační koeficient se zpravidla používá jako první krok k prozkoumání, zda můžeme mezi proměnnými očekávat závislosti. Po potvrzení se zpravidla použije další model, např. lineární regresní analýza. Korelační analýza tedy není totéž, co příčinná závislost. Pokud totiž spolu dvě veličiny korelují, nemusí to nutně znamenat, že se jedná o kauzalitu, tedy že jedna ovlivňuje druhou. Korelační koeficient nabývá hodnot od -1 do 1.

V ekonometrii je kauzalita chápána jako schopnost určité proměnné predikovat jinou proměnnou. Grangerovův přístup spočívá v tom, že kauzalitu interpretuje jako schopnost vylepšit odhad budoucí hodnoty jedné proměnné vícerozměrné časové řady na základě minulých a současných hodnot jiné proměnné. Předpokladem použití Grangerovy kauzality jsou stacionární data.

OTÁZKY

1. Příčinná kauzalita mezi dvěma proměnnými znamená, že existence určitého jevu souvisí s existencí druhého jevu. ANO x NE
2. Hodnota korelačního koeficientu -0.95 znamená, že:
    a) mezi proměnnými není lineární vztah
    b) mezi proměnnými existuje přímá korelace
    c) mezi proměnným existuje negativní vztah
3. Hodnota korelačního koeficientu 0.2 ukazuje vysokou pozitivní závislost. ANO x NE
4. K posouzení kauzálních vztahů v rámci modelu VAR může být využita analýza impuls-reakce. ANO x NE
5. V rámci Grangerovy kauzality je možno určit, která proměnná je endogenní a která exogenní. ANO x NE

ODPOVĚDI

1. ANO
2. c)
3. NE
4. ANO
5. ANO

5 KOINTEGRACE A MODELY KOREKCE CHYB

RYCHLÝ NÁHLED KAPITOLY

Ekonomické a finanční proměnné často vykazují trend či sezónnost. Tyto časové řady nejsou zpravidla stacionární a pro stacionarita je nutné je diferencovat. To však může mít i svá úskalí, zejména může dojít ke zkreslení dat a výsledků modelu. Proto se v páté kapitole studijní oporu bude věnovat pozornost koinegraci a modelům korekce chyb, které s těmito daty dokážou pracovat. V kapitole tedy bude popsána koinegrace a testování kointegrace, dále bude pozornost věnována testování řádu kointegrace – metoda Johansena. V další části textu bude popsán model korekce chyb a vektorový model korekce chyb (VEC).

CÍLE KAPITOLY

- Definovat kointegraci.
- Popsat postup testování kointegrace.
- Definovat model korekce chyb.
- Popsat postup konstrukce modelu korekce chyb.

KLÍČOVÁ SLOVA KAPITOLY

Kointegrace, nestacionární časová řada, model korekce chyb, zdánlivá regrese, Johansenův kointegrační test, verifikace modelu.

5.1 Kointegrace

Makroekonomické proměnné často vykazují trend (stochastický nebo deterministický), takže makroekonomické modely popisují v řadě případů zdánlivou regresi. Zatímco deterministický trend eliminujeme zahrnutím vysvětlující proměnné čas do modelu, stochastický trend odstraníme pomocí diferencí proměnných. Diferencování ukončíme tehdy, dosáhneme-li na základě příslušných testů požadované stacionarity proměnných. Nejčastěji postačuje k stacionaritě přechod první diference. Tento postup má však některá úskalí. Je-li model přesně specifikován a diferencujeme-li jeho proměnné, pak implicitně diferencujeme i jeho náhodné složky. To se projeví i v náhodných složkách modelu MA a mohou tak vzniknout problémy při odhadu. Navíc model obsahující diferencované proměnné nedává jednoznačně dlouhodobá řešení pro vysvětlovanou proměnnou. Chceme-li, aby model umožňoval získat jak krátkodobé, tak dlouhodobé řešení a zároveň byla zachována stacionarita všech jeho proměnných, je třeba se vrátit k použití proměnných měřených v jejich původních úrovních (Hušek, 2009).

Jsou-li např. dvě proměnné x a y nestacionární, pak náhodná složka v jejich regresi je také obvykle nestacionární a lze ji vyjádřit jako výsledek kombinace dvou kumulovaných náhodných procesů. Tyto kumulované procesy generující náhodné složky se označují jako stochastické trendy.

Ve zvláštním případě se však vzájemné závislé proměnné y a x mohou vyvíjet stejným směrem, přičemž navíc jejich stochastické trendy jsou shodné nebo téměř stejné a jejich kombinací lze vyloučit nestacionárnost. V tomto případě označujeme proměnné za kointegrované. Nejsou-li proměnné kointegrované, vzniká problém zdánlivé regrese, takže výsledky odhadnutého modelu jsou nereálné. Nejsou-li proměnné y a x kointegorvané, mohou v čase divergovat, takže mezi nimi neexistuje žádný dlouhodobý vztah. Naopak při existenci dlouhodobé závislosti mezi nimi, která se projevuje zhruba společným trendem, dochází pouze krátkodobě k odchylkám obou kointegrovaných proměnných od jejich dlouhodobého trendu, resp. rovnovážného vztahu.

Koncept kointegrace zavedl Granger (1981) a později jej dále rozpracovali Engle a Granger (1987). Pro existenci rovnováhy nebo dlouhodobé závislosti mezi dvěma proměnnými y_ta x_t musí platit, že jejich lineární kombinace je stacionární proměnnou, tj. integrovanou I(0). O obou proměnných se předpokládá, že jsou integrované stejného řádu, tzn., že k dosažení jejich stacionarity je potřeba stanovit diference shodného řádu. Nejčastěji postačí k dosažení stacionárnosti proměnných použití jejich prvních diferencí, takže je označujeme jako I(1) neboli integrované řádu jedna (Hušek, 2009).

TESTOVÁNÍ KOINTEGRACE

Zjistíme-li, že statistická data obsahují jednotkové kořeny (tj. nejsou stacionární), můžeme před jejich diferencováním, které vede ke ztrátě dlouhodobé informace obsažené v datech testovat kointegraci aplikací vhodného testu jednotkového kořenu na spočtená rezidua odhadnutého modelu.

Testy kointegrace jsou obvykle založeny na DF nebo ADF testech jednotkových kořenů. Nejčastěji se používají Engleův-Grangerův test kointegrace nebo Johansenův test kointegrace.

EViews umožňuje aplikaci Johansenova testu kointegrace. Avšak tato metoda je citlivá na délku zpoždění a dodržení předpokladu nezávislého normálního rozdělení náhodných složek. K dosažení statistické spolehlivosti odhadů je nutný velký výběr pozorování. Výsledek tohoto testu kointegrace se také může mírně měnit v závislosti na tom, zda a jakým způsobem je zahrnutý v modelu deterministický trend (Hušek, 2009).

Engle-Grangerova metoda

Tuto jednoduchou metodu navrhli Engle a Granger (1987). Autoři vycházeli z toho, že pokud sestavíme model z proměnných, které jsou kointegrované, rezidua tohoto modelu budou stacionární. Stačí tedy otestovat tyto rezidua na stacionaritu.

Podle definice, aby byly proměnné kointegrované, musí být I(d), kde d>2 a rovněž se *d* rovná pro všechny proměnné. Tedy omezíme na případ I(1), tedy časové řady s jednotkovým kořenem, které jsou stacionární na první diferenci. Pokud výsledky testů ukazují na stacionaritu, proměnné nemůžou být kointegrované. Jestliže podle výsledků testů jsou proměnné nestacionární, můžeme pokračovat v testování samotné kointegrace.

Použitím metody nejmenších čtverců odhadneme kointegrační rovnici ve tvaru:

$$y_t = \beta_0 + \beta_1 x_t + \epsilon_t \tag{24}$$

kde y_t a x_t jsou zkoumané časové řady. Z výsledků této regrese nás nezajímají hodnoty koeficientů, ale pouze odhadnutá rezidua ϵ_t. Použijeme opět test jednotkového kořene na tyto odhady reziduí.

Pokud nemůžeme zamítnout nulovou hypotézu o přítomnosti jednotkového kořene, musíme konstatovat, že zde neexistuje kointegrační vztah. Pokud ale nulovou hypotézu zamítáme, znamená to, že rezidua z kointegrační rovnice jsou stacionární a můžeme tak vyvodit závěr, že časové řady yt a xt jsou kointegrované. Hypotézy jsou tedy definované jako:

H0: $\epsilon t \sim I(1)$ Nepřítomnost kointegrace

H1: $\epsilon t \sim I(0)$ Kointegrace

Pokud jsou proměnné kointegrované, dalším krokem Engle-Grangerovy metody je odhad modelu korekce chyb. Odhadnutá rezidua z kointegrační regrese použijeme jako jednu z proměnných. Model může mít následující tvar:

$$\Delta y_t = \gamma_1 \Delta y_t + \gamma_2 \epsilon_{t-1} + e_t \tag{25}$$

kde $\epsilon_{t-1} = y_{t-1} - \beta_1 x_{t-1}$ a nazývá se člen korekce chyby (někdy taky korekční člen). β_1 popisuje dlouhodobý vztah mezi proměnnými x_t a y_t, rovněž se nazývá dlouhodobý multiplikátor. Oproti tomu γ_1 popisuje krátkodobý vztah mezi změnami v x_t a změnami v y_t. Parametr γ_2 popisuje rychlost, s jakou se vychýlení z rovnováhy přizpůsobí zpět k rovnováze. Pomocí tohoto modelu jsou k dispozici informace jak o krátkodobém chování proměnných, tak o dlouhodobé rovnováze.

Engle-Grangerova metoda má ovšem i své nevýhody. Pokud jsou proměnné v modelu nestacionární, odhad tohoto modelu pomocí MNČ je značně nespolehlivý. Dále jelikož Engle-Grangerova metoda používá Dickey-Fullerův test, přebírá i jeho problém malé síly. Rovněž může být problém se specifikací kointegrační rovnice. Když neznáme směr závislosti proměnných, nebo je tato závislost obousměrná, musíme i tak jednu z proměnných označit za závisle a druhou nezávisle proměnnou.

Další problém je přítomnost více kointegračních vztahů. Pokud máme kointegrační regresi o více proměnných, nevíme, zda je zde přítomno více kointegračních vztahů, či který z nich případně odhadujeme, protože Engle-Grangerova metoda nám odhalí pouze jeden kointegrační vztah, i když je jich třeba přítomno více. Toto můžeme vyřešit provedením více Engle-Grangerových testů na rovnicích, které budou různými kombinacemi zkoumaných proměnných. V tomto smyslu je Engle-Granger metoda převážně testem jednotkového kořene. Další a lepší možností je využití Johansenova testu.

Johansenův kointegrační test

Johansenův kointegrační test je především vhodnější k testování kointegrace v případě, kdy je možné, že je v odhadované rovnici přítomno více kointegračních vztahů.

Prvním krokem, než bude proveden Johansenův test, je nutnost ověřit, zda veškeré proměnné jsou nestacionární a jsou I(1). Obecně, musí být proměnné I(d), kde *d* je stejné pro všechny proměnné. Pokud je *d* pro proměnné různé, můžeme uvažovat ještě případ multikointegrace.

Pokud je výše uvedená podmínka splněná, pak můžeme provádět Johansenův test. Johansenův přístup testuje omezení vyplývající z kointegrace na VAR modelu:

$$\Delta\overrightarrow{Y_t} = \delta + \Gamma_1\Delta\overrightarrow{Y_{t-1}} + \cdots + \Gamma_{\mathrm{p}-1}\Delta\overrightarrow{Y_{t-p+1}} + \Pi\overrightarrow{Y_{t-1}} + \overrightarrow{\varepsilon_t}, \tag{26}$$

kde $\overrightarrow{\varepsilon_t}$ je vektor reziduálů $NID(0, \Sigma)$, tedy normální rozdělení (NID). Za předpokladu, že $\overrightarrow{Y_t}$ je *k*-rozměrný vektor proměnných, kde se předpokládá, že řady jsou integrovány řádu *I*(1), zatímco *r* lineární kombinace $\overrightarrow{Y_t}$ jsou stacionární, lze zapsat:

$$\Pi = \gamma\beta', \tag{27}$$

kde γ a β jsou rozměry $k \times r$. β označuje matici kointegračních vektorů, přičemž γ představuje matici vah, se kterými každý kointegrační vektor vstupuje do každé $\Delta\overrightarrow{Y_t}$ rovnice. Johansenův přístup je založen na maximálně věrohodném odhadu rovnice (26) s omezením (rovnice 27) pro danou hodnotu *r*.

Z rovnice (27) lze pozorovat, že pouze jeden člen v rovnici, $\Pi\overrightarrow{Y_{t-1}}$, je na stupni, kointegrační vztahy závisí na vlastnostech matice Π. Ale $\Pi\overrightarrow{Y_{t-1}}$ musí být *I*(0) nebo nula a $\overrightarrow{Y_t}$ je už stacionární. Tady mohou nastat tři situace:

a) $\Pi = \gamma\beta'$ má rozsah $0 < r < k$,

b) $\Pi = \gamma\beta'$ má nulový rozsah,

c) $\Pi = \gamma\beta'$ má plný rozsah.

Podle situace (a), γ a β jsou obě matice $k \times r$ a mají rozsah r. Existuje r kointegračních vektorů $\beta'\overrightarrow{Y_t}$, které jsou stacionární na první diferenci $I(0)$. To je ekvivalentem k r společným trendům mezi $\overrightarrow{Y_t}$. Stacionarita $\beta' y_t$ značící dlouhodobý vztah mezi y_t nebo náhradním souborem $\overrightarrow{Y_t}$ proměnných v kointegračních vektorech se od sebe nesmí v průběhu času odchýlit. $\beta'\overrightarrow{Y_t}$ jsou korekční členy, kde se odklon z rovnováhy jednotlivých proměnných v kointegračních vektorech následně vrací zpět k rovnováze, tedy dynamický proces úprav, nebo-li model korekce chyb. Rovnice (26) se proto nazývá VAR s modelem korekce chyb. Podle situace (b) neexistuje kointegrační vztah mezi y_t proměnnými na stupni nevstupující do rovnice (26) a ta se tak stává jednoduchým VAR bez modelu korekce chyb. Proměnné na stupni jsou již stacionární za situace (c).

V závislosti na tom, zda y_t a/nebo kointegrační vektory mají konstantu a/nebo deterministický trend, existuje v praxi pět modelů: (i) žádný deterministický trend v $\overrightarrow{Y_t}$ a žádná konstanta v kointegračních vektorech, (ii) žádný deterministický trend v $\overrightarrow{Y_t}$, ale jsou konstanty v kointegračních vektorech; (iii) existují deterministické trendy v $\overrightarrow{Y_t}$ a jsou konstanty v kointegračních vektorech, (iv) existují deterministické trendy v $\overrightarrow{Y_t}$ a v kointegračních vektorech a (v) existují kvadratické trendy v y_t a deterministické trendy v kointegračních vektorech.

V Johansenově testu je Π počítáno pomocí maximalizace. Výše uvedených pět specifikací je pak testováno pomocí dvou pravděpodobnostních testovacích kritérií k určení počtu kointegračních vazeb, tedy Maximum eigenvalue statistic a Trace statistic. K testování hypotézy $H_0: r \leq r_0$, že existuje r kointegračních vektorů a alternativní hypotéze $H_1: r = r_0 + 1$, slouží Maximum eigenvalue statistic:

$$\lambda_{max}(r_0) = -T\log\left(1 - \hat{\lambda}_{r_{0+1}}\right). \tag{28}$$

Maximum eigenvalue test je založen na odhadu $(r_0 + 1)$ největší charakteristické hodnoty matice (eigenvalue). Nulová hypotéza $H_0: r \leq r_0$ a alternativní hypotéza $H_1: r_0 < r \leq k$, může být testována použitím Trace testu, který kontroluje, zda nejmenší $k - r_0$ charakteristické hodnoty matice jsou významně odlišné od nuly:

$$\lambda_{trace}(r_0) = -T \sum_{j=r_0+1}^{k} \log\left(1 - \hat{\lambda}_j\right). \tag{29}$$

5.2 Modely korekce chyby

Nejsou-li pozorování proměnných v modelu rozdělených zpoždění stacionární, ani kointegrovaná, použijeme obvykle jejich diferencované hodnoty, abychom odstranili v nich obsažené jednotkové kořeny a neodhadovali zdánlivou regresi. Použití diferencí proměnných sice umožňuje reálnou specifikaci dynamických modelů rozdělených či autoregresních zpoždění i aplikaci standardních postupů statistické indukce, avšak současně se zbavujeme možnosti kvantifikovat vztahy mezi původními úrovněmi proměnných, které jsou v ekonomické analýze důležité. Alternativní způsob specifikace umožňují modely korekce chyby (VEC), resp. korekce rovnováhy, nevyžadující stacionaritu proměnných, pokud jsou kointegrovanné.

Skupinu kointegrovaných časových řad je možné popsat modelem korekce chyb, prostřednictvím kterého lze odlišit ve zkoumaných procesech dlouhodobé tendence od krátkodobých.

Modely korekce chyb jsou modely s kointegrovanými proměnnými, které obsahují jak krátkodobé, tak dlouhodobé vztahy mezi procesy. Tyto modely umožňují oddělit tyto druhy vztahů a zkoumat je každý zvlášť. Modely korekce chyb nevyžadují stacionaritu proměnných, pokud se mezi nimi vyskytuje kointegrace. Pokud je model VAR sestaven ze stacionárních řad upravených diferencemi a jsou navíc kointegrované, vzniká v modelu specifikační chyba při odhadu. Je tedy žádoucí použití VEC modelu. VEC model vychází z vícerovnicového systému pro jednotlivé proměnné, stejně jako modely VAR.

VEC model pro dvě proměnné y_t a x_t:

$$\Delta y_t = a_1 + a_{11}\Delta y_{t-1} + \gamma_{11}\Delta \mathrm{x}_{t-1} + \eta_1\big(y_{t-1} - -\hat{\beta}x_{t-1}\big) + u_{t1} \tag{30}$$

$$\Delta \mathrm{x}_t = a_2 + a_{21}\Delta y_{t-1} + \gamma_{21}\Delta \mathrm{x}_{t-1} + \eta_2\big(y_{t-1} - -\hat{\beta}x_{t-1}\big) + u_{t2} \tag{31}$$

kde $y_{t-1} - \hat{\beta}x_{t-1} = e_{t-1}$, což je korekční chybový člen. modelu se může vyskytovat více než jeden kointegrační vztah. V takovém případě je v každé rovnici více korekčních chybových členů. Model může mít i více zpoždění než jedno a m > 2proměnných.

Model korekce chyby (EC - error correction) patří do kategorie několika modelů časových řad nejčastěji používaných v případě, kdy mezi podkladovými pro-měnnými existuje dlouhodobý kointegrační vztah, také známý jako kointegrace. EC model je teoreticky řízený přístup vhodný pro odhadování jak krátkodobých, tak i dlouhodobých účinků jedné časové řady na jiné. Nyní uvažujme dvě řady y_t a x_t, které jsou nestacionární typu I(1). Podezření, že jedná řada ovlivňuje jinou, budeme vzhledem k její stacionaritě vyšetřovat pomocí modelu

$$\Delta y_t = \delta \Delta x_t + \epsilon_t.$$

Tento vztah udává pouze krátkodobý vztah mezi y_t a x_t, protože nás ale zajímá vztah mezi proměnnými x a y až po jeho dlouhodobém vyvažení do rovnovážného stavu, nemá z tohoto pohledu tento vztah žádnou vypovídací schopnost. Pokud je reálné považovat řady y_t a x_t za kointegrované v dlouhodobém horizontu, pak můžeme model korigovat a uvažovat opravený model

$$\Delta y_t = \Delta x_t + \alpha\left(y_{t-1} - \beta x_{t-1}\right) + \epsilon_t,$$

kam jsme dodali korekční člen, který je vytvořen z úrovňových hodnot daných veličin v předchozím čase t−1.

Vzhledem k tomu, že jsme uvažovali y_t a x_t integrované řádu 1, jejich diference jsou stacionárním procesem. Existuje-li dlouhodobý rovnovážný vztah mezi *y* a *x* (jsou kointegrované), pak i kointegrační vztah $y_{t-1} - \beta x_{t-1}$ je stacionárním procesem. Tedy všechny členy v modelu jsou stacionární. Model sice popisuje krátkodobý vztah mezi přírůstky Δxt a Δyt, ale zároveň provádí korekci pro případ, že krátkodobé změny obou veličin odchylují úrovně těchto veličin od jejich dlouhodobého rovnovážného vztahu. Model se nazývá EC model (error correction), výrazy typu $y_{t-1} - \beta x_{t-1}$ se nazývají korekční členy, parametry typu β popisují dlouhodobé kointegrační vztahy mezi proměnnými, parametry typu δ krátkodobé vztahy mezi proměnnými a parametry typu určují rychlost přizpůsobení rovnovážnému stavu.

KONSTRUKCE MODELU KOREKCE CHYB

Konstrukci modelu korekce chyb lze popsat v několika krocích:

1. Provedeme testy jednotkového kořene pro jednotlivé časové řady. Pokud jsou nulové hypotézy o jednotkových kořenech zamítnuty, pak jsou časové řady stacionární a lze na tyto časové řady zkonstruovat např. VAR model. V opačném případě obsahují dané řady vzhledem k jednotkovým kořenům stochastické trendy, přejdeme k následujícímu kroku (2).
2. Provedeme Johansenův kointegrační test. Jestliže je kointegrace zamítnut (*r* = 0), přejdeme ke kroku (3). Pokud je potvrzena existence *r* kointegračních vztahů, přejdeme ke kroku (4).
3. Protože kointegrace byla zamítnuta, převedou se data na stacionární a zkonstruuje se jiný model, např. VAR model.
4. Protože existuje *r* koinegračních vztahů, odhadneme pro $y_1, \dots, y_n$, příslušný model korekce chyb.

VEKTOROVÝ MODEL KOREKCE CHYB

Za předpokladu, že jsou časové řady nestacionární a zároveň kointegrované, je možné zkoumat vzájemný vztah proměnných pomocí vektorového modelu korekce chyb (Vector

Error Correction Model - VECM). Model korekce chyb zohledňuje díky kointegračnímu vektoru reziduí i dlouhodobé vztahy mezi proměnnými.

Ilustračně se dá model popsat na dvou referenčních časových řadách, které jsou obě řádu I(1). Pokud lze tyto časové řady považovat za kointegrované, jsou dlouhodobě v rovnováze. Jak je zřejmé z předchozí teorie, mohou se tyto řady krátkodobě vychylovat, a proto je nutné zavést korekční člen. Model korekce chyb pak lze zapsat dle následující rovnice:

$$\Delta y_t = \alpha \cdot (x_{t-1} - \beta \cdot y_{t-1}) + \varepsilon_{1,t},$$
$$\Delta x_t = \alpha \cdot (x_{t-1} - \beta \cdot y_{t-1}) + \varepsilon_{2,t},$$

kde výraz v závorce představuje korekční člen, který je vytvořen z úrovňových hodnot v čase t-1. Krátkodobé nerovnovážné vztahy v modelu jsou popsány přírůstky Δyt a Δxt, jsou však korigovány, aby se dlouhodobě neodchylovaly od rovnovážného stavu. Dle alternativního maticového zápisu VECM (3.25), je stejně možné rozdělit Π na dva členy, pak Π=αβ‘, pak lze rovnici zapsat podle (3.29):

$$\Delta y_t = \alpha(\beta' y_{t-1}) + \Gamma_1 \Delta y_{t-1} + \Gamma_2 \Delta y_{t-2} + \ldots + \Gamma_{p-1} \Delta y_{t(p-1)} + \varepsilon_t,$$

kde α udávají hodnotu koeficientů adjustace, pro každou rovnici v maticovém modelu. Tyto koeficienty určí, jak rychle se systém dostane po vychýlení zpět do rovnovážného stavu. Jednotlivé členy β‘ matice představují kointegrační vztahy, které v modelu popisují dlouhodobé vztahy mezi proměnnými, sestavují se z nich tzv. kointegrační vektory. Koeficienty ΓiΔyi pak představují členy korekce chyb matice stacionárních hodnot zpožděných proměnných.

Pokud se v rámci kointegrace zjistí dlouhodobá rovnovážná vazba mezi proměnnými a možnosti vzniku krátkodobých výkyvů mezi danými veličinami, k odhalení těchto výkyvů slouží v rámci kointegrace model korekce chyby. Cipra (2008) uvádí, že ECM je adekvátním nástrojem pro zkoumání krátkodobých odchylek nutných k dosažení dlouhodobé rovnováhy mezi zkoumanými proměnnými. Pokud lze považovat řady $\{x_t\}$ a $\{y_t\}$ právě v dlouhodobém horizontu za kointegrované, pak lze model korekce chyby zapsat:

$$\Delta y_t = \gamma_1 + \gamma_2 \Delta x_t + \alpha(y_{t-1} - \beta_1 - \beta_2 x_{t-1}) + \varepsilon_t,$$

kde výrazy y_{t-1} a x_{t-1} jsou korekční členy, parametry β popisují dlouhodobé kointegrační vztahy mezi proměnnými a zapisují se do kointegračních vektorů typu $(1, -\beta)'$, parametry γ popisují krátkodobé vztahy mezi proměnnými a parametry α určují rychlost přizpůsobení rovnovážnému stavu. Konstanta γ_1 z hlediska úrovně vysvětlované proměnné y_t značí deterministický lineární trend. Korekční člen je vytvořen z úrovňových (a ne rozdílových) hodnot daných veličin v předchozím čase $t - 1$. Tento model sice popisuje krátkodobý vztah mezi přírůstky Δx_t a Δy_t, ale zároveň provádí korekci pro případ, že krátkodobé změny obou veličin odchylují úrovně těchto veličin od jejich dlouhodobého rovnovážného vztahu. Pokud jsou řady $\{x_t\}$ a $\{y_t\}$ opravdu kointegrované a korekční člen v rovnici je právě kointegračním vztahem (a představuje tedy stacionární časovou řadu), pak všechny členy v modelu jsou zřejmě stacionární. Nevzniká tedy situace, že by se v jednom modelu porovnávaly stacionární a nestacionární členy, což obvykle působí hlavní problémy při konstrukci takového modelu. Uvedený model se označuje jako EC model, tedy model korekce chyby a pro zdůraznění VAR kontextu se používá označení VEC (vector error correction). Řepková (2012).

STATISTICKÁ VERIFIKACE MODELU

Statistická verifikace je zásadní pro význam a použití odhadovaného modelu, stanoví na kolik je model přesný a spolehlivý. Mezi základní metody ověření lze řadit např. výpočet a interpretaci koeficientu determinace a upraveného koeficientu determinace, dále také testování statistické významnosti modelu.

a) Koeficient determinace

Koeficient determinace vyjadřuje poměr, ve kterém jsou změny výstupu modelu způsobeny vysvětlujícími proměnnými. Zbytkový efekt je zapříčiněn vlivem residuální složky. Determinační koeficient se značí R^2.

Za předpokladu, že je hodnota determinačního koeficientu 1, vysvětlují proměnné perfektně odhadovaný model. Pokud je hodnota R^2 rovna nule, vysvětluje celý model residuální část, odhadovaný model pak nemá smysl. Nedostatky modelu jsou jeho nízká reakce na změny počtu pozorování v modelu a nezohlednění vyššího počtu vysvětlujících proměnných, jak uvádí Hančlová (2012). Z toho důvodu se používá upravený ukazatel - korigovaný koeficient determinace.

Z důvodů problémů, které představuje používání koeficientu determinace, se častěji používá upravený, tj. korigovaný koeficient determinace. Tato statistika má stejnou interpretaci, jako R^2 a označuje se R^2_{adj}. Korigovaný koeficient determinace může nabývat hodnot v rozmezí <0,1>.

b) F-test

Podstatou F-testu je zkoumání statistické významnosti modelu jako celku. Testování statistické významnosti lze provést pomocí zkoumání kritického oboru. Pro tento účel musí být nejprve stanoveny hypotézy, které budou testovány porovnáním testové a kritické statistiky F-testu. Hypotézy jsou stanoveny následovně podle vztahů

$$H_0: \beta_1 = \beta_2 = \cdots = \beta_k = 0,$$

kde βi představují parametry jednotlivých vysvětlujících proměnných. Nulovou hypotézu lze slovně formulovat jako: H0: Model není statisticky významný. Pokud nedojde k zamítnutí nulové hypotézy, je variabilita modelu způsobena náhodnou složkou, jak uvádí Hančlová (2012).

$$H_1: \beta_1 \neq 0 \vee \beta_2 \neq 0 \vee \ldots \beta_k \neq 0,$$

kde βi představují parametry jednotlivých vysvětlujících proměnných. Alternativní hypotézu lze slovně formulovat jako: H1: Model je statisticky významný. Pokud dojde k zamítnutí nulové hypotézy, znamená to, že v modelu existuje alespoň jedna proměnná s nenulovým významným koeficientem.

DALŠÍ ZDROJE

Řešený vzorový příklad na VAR model a Johansenův kointegrační test a model korekce chyb najdete ve videonahrávce v LMS Moodle. Videonahrávky slouží jako přímý doplněk ke studijní opoře.

SHRNUTÍ KAPITOLY

Pátá kapitola studijní opory byla zaměřena na kointegrace a model korekce chyb. V případě, že máme nestacionární časové řady a všechny jsou integrovány na stejném stupni, ideálně $I(1)$, můžeme testovat kointegraci. Jinými slovy, ve zvláštním případě se vzájemné závislé proměnné y a x mohou vyvíjet stejným směrem, přičemž navíc jejich stochastické trendy jsou shodné nebo téměř stejné a jejich kombinací lze vyloučit nestacionárnost. V tomto případě označujeme proměnné za kointegrované. Nejsou-li proměnné kointegrované, vzniká problém zdánlivé regrese, takže výsledky odhadnutého modelu jsou nereálné.

Testy kointegrace jsou obvykle založeny na DF nebo ADF testech jednotkových kořenů. Nejčastěji se používají Engleův-Grangerů test kointegrace nebo Johansenův test kointegrace. Ve studijní opoře a v řešeném příkladu uvedeném ve videonahrávce, která slouží jako doplněn ke studijní opoře, je aplikován Johansenův kointegrační test. V případě, že zjistíme, že data mají kointegrační vazbu, aplikujeme model korekce chyb. Model korekce chyb slouží k určení krátkodobých odchylek od dlouhodobé rovnováhy.

OTÁZKY

1. Engleův-Grangerův test je nejčastěji používaným testem koinegrace, která podává nejvhodnější výsledky a závěry. ANO x NE
2. Výhodou Johansenova kointegračního testu je, že jej můžeme použít v případě, kdy závislá proměnná je stacionární a nezávislá proměnná je nestacionární. ANO x NE
3. Vektorový model korekce chyb slouží k určení krátkodobých výkyvů veličin od dlouhodobé rovnováhy. ANO x NE
4. Při statistické verifikaci modelu zjišťujeme výši koeficientu determinace. Pokud je koeficient determinace roven 0,10, znamená to, že:
   a) proměnné perfektně vysvětlují odhadovaný model
   b) proměnné velmi omezeně vysvětlují odhadovaný model
   c) model je sestavený správně a jsou použité vhodné proměnné
5. Statistickou významnost modelu jako celku lze testovat pomocí:
   a) koeficientu determinace
   b) F-testu
   c) autotokorelace

ODPOVĚDI

1. NE
2. NE
3. ANO
4. b)
5. b)

6 ANALÝZA PANELOVÝCH DAT

RYCHLÝ NÁHLED KAPITOLY

V šesté kapitole studijní opory bude popsána analýza panelových dat. V předešlých kapitolách studijní opory se pracovalo s časovými řadami, nyní bude představeno zpracování panelových dat. Nejprve budou pospána panelová data a jejich výhody a nevýhody. V další části bude popsán model panelové regrese, její výhody a nevýhody. Také u analýzy panelových dat je předpoklad použití stacionárních dat, proto bude pozornost věnována testům jednotkových kořenů panelových dat. Budou popsány taky individuální efekty, konkrétně fixní a náhodné efekty.

CÍLE KAPITOLY

- Definovat panelová data.
- Popsat panelovou regresi.
- Popsat testy jednotkových kořenů panelových dat.
- Definovat individuální efekty.

KLÍČOVÁ SLOVA KAPITOLY

Panelová data, fixní efekty, náhodné efekty, metoda nejmenších čtverců, zobecněná metoda momentů, GMM.

6.1 Panelová data

Panelová data byla popsána v první kapitole. Panelová data lze využít v případě, pokud je k dispozici soubor jednotek, které si jsou příbuzné nebo velmi blízké určitou charakteristikou, a u nichž se pozorování v čase opakuje.

Panelová data jsou data, kde jsou charakteristiky za jednotlivá pozorování zjišťovány za více časových období. Soubor panelových dat tvoří vzorek obsahující N průřezových jednotek, které jsou pozorovány v různých T časových obdobích (Asteriou a Hall, 2011). Verbeek (2004) uvádí, že panelový soubor dat (panel data set) obsahuje opakované zjišťování stejných jednotek (jednotlivci, domácnosti, firmy), shromážděné v průběhu několika období. Dostupnost opakovaných pozorování na stejných jednotkách umožňuje ekonomům specifikovat a odhadnout složitější a realističtější modely, než lze udělat v jedné průřezové nebo časové řadě.

6.2 Panelová regrese

Jak bylo uvedeno v 2. kapitole, regresní analýza slouží pro kvantitativní popis vztahu mezi ekonomickými a finančními veličinami (proměnnými). Úkolem regresní analýzy je vysvětlit změny hodnot vysvětlované změnami hodnot vysvětlujících proměnných.

K ZAPAMATOVÁNÍ

Panelová regrese (někdy se také nazývá analýza panelových dat) je statisticko-ekonometrická metoda, která se zabývá analýzou vztahů a souvislostí dat ve dvourozměrném prostoru. První rozměr tvoří časová veličina, druhým rozměrem jsou průřezová data jednotlivých objektů pozorování (Novák, 2007).

Jednoduchý regresní model panelových dat lze zapsat:

$$y_{it} = \alpha_1 z_{i1} + \alpha_2 z_{i2} + \cdots + \alpha_q z_{iq} + \beta_1 x_{it1} + \beta_2 x_{it2} + \cdots + \beta_k x_{itk} + u_{it}, \tag{32}$$

kde y_{it} je vysvětlovaná proměnná, x_{it1} až x_{itk} jsou vysvětlující proměnné a proměnné z_{i1}, z_{i2} představují individuální efekty a zde se zařazuje případný vektor jednotek. Individuální efekty se v čase nemění. Index i označuje průřezový rozměr $i = 1, \ldots, N$, index t označuje časový rozměr $t = 1, \ldots, T$. Jak uvádí Asteriou a Hall (2011), pokud se vzorek dat skládá z konstantního T pro všechny průřezové jednotky, pokud je získaný vzorek kompletní u všech časových i průřezových řad, pak se datový soubor nazývá balancovaný (vyvážený). Naopak, pokud pro nějaký časový úsek nebo některá průřezová data pozorování schází, pak se panel nazývá nebalancovaný (či nevyvážený).

Mezi výhody používání panelových dat lze např. zařadit:

- Kontrola individuální heterogenity, tedy panelová data naznačují, že jednotlivci, firmy, státy nebo země jsou různorodé. U časových a průřezových řad, které nekontrolují tuto heterogenitu, hrozí získání zkreslených výsledků.
- Panelová data poskytují více informativní údaje (větší vzorek dat), větší variabilitu, menší kolinearitu mezi proměnnými, také více stupňů volnosti a větší efektivnost.
- Panelová data jsou schopna lépe studovat dynamiku změn.

- Panelová data jsou schopna lépe identifikovat a měřit dopady, které nelze jednoduše detekovat v čistě průřezových nebo čistě časových řadách.
- Panelové modely umožňují sestavit a testovat složitější modely chování, než čistě průřezové nebo časové řady.
- Shromážděná mikro panelová data o jednotlivcích, firmách a domácnostech mohou být přesněji měřená než podobné proměnné na makro úrovni. Zkreslení vyplývající z agregace u firem či jednotlivců, může být sníženа nebo odstraněna.
- Na druhé straně, makro panelová data mají delší časovou řadu a na rozdíl od problému nestandardního rozdělení typických testů jednotkových kořenů v analýze časových řad, panelové testy jednotkových kořenů mají standardní asymptotické rozdělení (Řepková, 2012).

Samozřejmě práce s panelovými daty má i určité nevýhody, které často souvisejí s odhady či interpretací:

- Problém heteroskedasticity či autokorelace, problém průřezových korelací mezi jednotlivými subjekty v daných časových obdobích apod.
- Problémy se sběrem dat důsledkem problémů s neúplnými daty respondentů, nezaslání odpovědí (kvůli nedostatku spolupráce respondentů nebo z důvodu chyby tazatele), frekvence dotazování, mezery v rozhovoru, referenční období, ohraničení vzorku apod.
- Zkreslení chyb měření, které mohou vzniknout v důsledku chybných odpovědí kvůli nejasným otázkám, záměrnému zkreslení odpovědí, výběrem nevhodných respondentů nebo chybějícímu záznamu odpovědí.
- Problémy výběru vzorku, krátká časová řada, nebo průřezové závislosti jsou dalšími omezeními panelových dat.

6.3 Testy jednotkových kořenů panelových dat

Baltagi (2005) uvádí, že testy stacionarity lze rozdělit dle vstupních předpokladů na dvě skupiny. První, která obsahuje naprostou většinu standardně používaných testů, předpokládá průřezovou nezávislost dat – časové řady jednotlivých pozorovaných jednotek tedy nejsou korelovány. Druhá skupina tento často spíše restriktivní požadavek odmítá a v testovacích procedurách je s touto možností počítáno.

Práce současných autorů navrhují tedy testy jednotkových kořenů panelových dat, které mají větší sílu než testy jednotkových kořenů používaných pro ověřování stacionarity jednorozměrných časových řad. Testy předpokládající průřezovou nezávislost dat zahrnují metody autorů:

- Levin, Lin a Chu (2002) – test LLC,
- Breitung (2000),
- Im, Pesaran a Shin (2003) – test IPS,

- Maddala a Wu (1999), Choi (2001) – Fisher-ADF test a Fisher-PP test,
- Hadri (2000).

Ačkoli se tyto testy nazývají jako panel unit root testy, jedná se v podstatě o testy paralelního zapojení jednoduchých testů jednotkových kořenů jednotlivých časových řad, které jsou aplikovány na panelovou strukturu dat.

V praxi se velmi často používá LLC test. LLC (2002) tvrdí, že testování jednotlivých časových řad panelu, pokud jsou časové řady krátké, vede často k zamítnutí alternativní hypotézy – tj. testy indikují přítomnost jednotkového kořene, přestože se v datech nevyskytuje. Navrhli tedy třístupňovou testovací proceduru, kdy nulovou hypotézou je tvrzení, že každá jednotlivá časová řada obsahuje jednotkový kořen. Alternativní hypotézou je pak tvrzení, že žádná jednotlivá časová řada jednotkový kořen neobsahuje. Metoda LLC vykazuje dva závažné nedostatky. Za prvé implicitně předpokládá, že časové řady jednotek panelu nejsou korelovány. Za druhé je jako velmi restriktivní vnímáno samotné stanovení hypotéz – buď je jednotkový kořen přítomen ve všech časových řadách panelu, nebo v žádné.

Přístup IPS (2003) se od předchozího odlišuje v tom, že zmírňuje restrikci požadavku na (ne)stacionaritu všech časových řad všech jednotek v panelu. IPS umožňuje testovat nulovou hypotézu, že každá časová řada v panelu obsahuje jednotkový kořen (tj. H0: ρi = 0), oproti alternativní hypotéze, kdy pouze některé (ale ne všechny) časové řady mají jednotkový kořen.

Další metoda je odvozená z prací Maddala, Vu (1999) a Choi (2001) a využívá testu Fisherova typu. Testuje se tedy kombinace všech p-hodnot testů jednotkových kořenůčasové řady každé individuální jednotky panelu, které byly provedeny zvlášť.

Další přístup k testování stacionarita panelových dat navrhl Hadri (2000), který oproti předchozím metodám testuje nulovou hypotézu, že jednotkový kořen není přítomen, oproti alternativní hypotéze, kdy jednotkový kořen přítomen je.

6.4 Fixní a náhodné efekty

K ZAPAMATOVÁNÍ

Problém heterogenity v panelových modelech řeší dva typy modelů, a to modely fixních efektů a modely náhodných efektů. Efekty slouží k odstranění určité veličiny z panelových dat, která se liší mezi jednotlivými skupinami nebo v čase.

Model s fixními efekty (fixed effects model) je lineární regresní model, ve kterém se konstanta mění v průběhu jednotlivých jednotek *i*, tj.:

$$y_{it} = \alpha_i + \beta x_{it} + \varepsilon_{it}, \tag{33}$$

$$\varepsilon_{it} \sim IID(0, \sigma_\varepsilon^2), \tag{34}$$

kde β je vektor konstant rozměru 1xK a α je konstanta reprezentující efekty těch proměnných, které jsou charakteristické i-tému pozorování. Chybová složka ε_{it} reprezentuje efekty nejvýznamnějších proměnných příznačných i-tým pozorováním a danému časovému intervalu. O této složce se předpokládá, že pochází z nezávisle identického rozdělení s nulovou střední hodnotou a konstantním rozptylem a dále se předpokládá, že všechny x_{it} jsou nezávislé na jakékoliv ε_{it}. Což lze zapsat obvyklým regresním vztahem včetně dummy proměnné pro každou jednotku i v modelu:

$$y_{it} = \sum_{j=1}^{N} \alpha_j d_{ij} + \beta x_{it} + \varepsilon_{it}, \tag{35}$$

kde d_{ij} = 1, pokud $i = j$ a všude jinde je 0. Parametry $\alpha_{1,\dots,}\alpha_N$ a β lze odhadnout metodou nejmenších čtverců. Všeobecně se v regresní analýze předpokládá, že faktory, které ovlivňují závislou proměnnou, ale přitom nejsou součástí regresorů, jsou zahrnuty do faktoru vyjadřujícího náhodné chyby.

Model s náhodnými efekty (random effects model) lze zapsat jako:

$$y_{it} = \mu + x'_{it}\beta + \alpha_i + \varepsilon_{it}, \tag{36}$$

$$\varepsilon_{it} \sim IID(0, \sigma_\varepsilon^2), \tag{37}$$

$$\alpha_i \sim IID(0, \sigma_\alpha^2), \tag{38}$$

kde $\alpha_i + \varepsilon_{it}$ je považováno za chybovou složku skládající se ze dvou složek, tedy jednotlivý specifický komponent, který se v průběhu času nemění a zbytková složka, kde je předpoklad nekorelovanosti v průběhu času. To znamená, že všechny korelace chybových podmínek v čase jsou přisuzovány jednotlivým efektům α_i. Předpokládá se, že α_i a ε_{it} jsou vzájemně nezávislé a nezávislé na x_{js} (pro všechna j a s). To znamená, že metoda nejmenších čtverců pro odhad μ a β je objektivní a konzistentní.

Pro panelová data vhodnou volbou mezi fixními a náhodnými efekty zahrnuje zkoumání, zda regresory jsou korelovány s individuálním (nepozorovaným ve většině případů) efektem. K rozhodování mezi metodou fixních a náhodných efektů se používá např. Hausmanův test (Hausman, 1978). Hausmanův test testuje, zda odhad fixními a náhodnými efekty je výrazně odlišný. Nulová hypotéza H_0 zní, že náhodné efekty jsou konzistentní a efektivní. Alternativní hypotéza H_1 říká, že náhodné efekty jsou nekonzistentní (a fixní efekty budou vždy konzistentní). Pokud je nulová hypotéza zamítnuta, není vhodné použití metody náhodných efektů z důvodu, že náhodná proměnná, která je zahrnutá do chybové složky, je pravděpodobně korelována s jednou nebo více nezávisle proměnnými. V tomto případě je tedy vhodnější použití metody fixních efektů. V případě, že nulová hypotéza zamítnuta není, což značí, že mezi odhadnutými koeficienty neexistují podstatné rozdíly, lze použít obě metody.

Základní pravidla pro rozhodování mezi fixními a náhodnými efekty ukazuje Gujarati (2002):

- (i) Je-li T (počet časových řad, počet období) velký a N (počet průřezových jednotek) malý, je pravděpodobné, že existuje malý rozdíl v hodnotách mezi hodnotami koeficientů odhadnutých metodou fixních a náhodných efektů. V tomto případě je však častěji vhodnější model fixních efektů.
- Pokud je N velký a T malý, odhady získané těmito dvěma metodami se mohou výrazně lišit. V náhodných efektech $\beta_{1i} = \beta_i + \varepsilon_i$, je ε_i průřezová náhodná složka, zatímco ve fixních efektech je β_{1i} považována za fixní. Ve druhém případě je statistický závěr podmíněn sledovanými průřezovými jednotkami ve výběrovém souboru. To je vhodné pouze v případě, že jednotlivé průřezové jednotky ve vzorku nejsou náhodné rysy vybrané z většího vzorku. V takovém případě jsou fixní efekty vhodné. Pokud však jsou průřezové jednotky ve vzorku považovány za náhodné rysy, pak náhodné efekty jsou vhodné a v tomto případě je statistický závěr bez výhrad.
- Pokud N je velký a T je malý, a drží-li předpoklady náhodných efektů, náhodné efekty jsou efektivnější než fixní efekty.
- Je-li individuální chybová složka ε_i korelována s jedním či více regresory, jsou odhady náhodnými efekty zkreslené a je vhodné použít metodu fixních efektů.

Je zde důležité poznamenat, že uvedená pravidla jsou pro rozhodování mezi fixními a náhodnými efekty pouze pomocnými pravidly a vždy záleží na konkrétním typu dat a dalších faktorech vstupujících do analýzy, včetně přístupu ekonometra (Řepková, 2012).

6.4.1 Statický a dynamický lineární model

Většina ekonomických veličin jsou ve své podstatě dynamickými procesy, proto by bylo vhodné je i jako dynamické modelovat. V případě použití panelu pro modelaci jako datové základny nám časový rozměr umožňuje popsat v čase probíhající proces přizpůsobování.

Panelová regrese může být plnohodnotným nástrojem pro odhad funkční závislosti velkého množství ekonomických proměnných. Prucha (2014) však uvádí, že mnohá panelová data trpí problémem kratší časové řady a z hlediska panelové regrese užitím metody nejmenších čtverců na tempech růstu jsou tak zcela nevhodná k jejímu použití. Metoda náhodných momentů (Generalized Method of Moments, GMM, v češtině také Zevšeobecněná metoda momentů) představuje způsob, jak zkoumat funkční vztahy právě mezi takovými panelovými daty. Finanční data na roční frekvenci, získána ze základních účetních výkazů, uspořádána v panelech, jsou z tohoto důvodu vhodným kandidátem pro výzkum užitím této regresní metody. Velkou výhodou užití GMM oproti metodě nejmenších čtverců je rovněž fakt, že mezi regresory také figuruje zpožděná endogenní, vysvětlovaná, závislá proměnná

(Hall, 2005). Nevýhodou tohoto modelu je nemožnost testování heteroskedasticity a autokorelace parametrů. Aby však bylo možné považovat výsledky testování za relevantní a byla podložena vypovídací schopnost výsledků testů v jednotlivých modelech, budou všechny modely v této metodě testovány jak z hlediska statistické významnosti jednotlivých položek modelu, tak z hlediska robustnosti modelu pomocí Sargan/Hansen testu. U Sargan/Hansen testu jde především o to, do jaké míry je metoda schopna v daném modelu poskytnout prakticky stejné výsledky i při zatížení mírnými změnami parametrů. Model je v tomto ohledu robustní, jestliže jsou výsledky Sargan/Hansen testu vyšší než 0,05 (Růčková, 2016).

Jak uvádí Pražák (2020), jedním z nejdůležitějších úkolů v ekonometrii je najít takové techniky, které umožní odhadnout pro daný soubor dat neznámé parametry daného modelu. Postupy odhadu založené na minimalizaci nebo maximalizaci hodnotového kritéria funkce byly úspěšně použity pro mnoho různých typů modelů. Hlavní rozdíl mezi těmito odhady spočívá v tom, co musí být specifikováno v modelu. Nejpoužívanější techniky odhadu vyžadují úplnou specifikaci modelu a jeho rozdělení pravděpodobnosti. Metoda GMM tento druh znalostí nevyžaduje. Požaduje pouze specifikaci momentových podmínek, které by měl model uspokojit.

Většina ekonomických veličin vytváří svou podstatou dynamické procesy, a proto je vhodné je i jako dynamické modelovat. V případě použití modelu pro modelaci, časový rozměr datové základny slouží k popsání v čase probíhajících procesů přizpůsobování (Novák, 2007).

I přesto při správném užití konečný model GMM splňuje podmínky konzistentnosti, asymptotické normálnosti a asymptotické efektivnosti ve třídě všech odhadů, které kromě informací obsažených v momentálních podmínkách nepoužívají žádné další informace. Dostupná data v rámci modelu GMM vystupují zpozorování T {Yt} t=1, ..., T, ve kterém každé pozorování Yt je n-dimenzionální multivariační náhodnou proměnnou. Obecným předpokladem metody GMM je, že data Yt lze generovat slabě stacionárním ergodickým stochastickým procesem.

DALŠÍ ZDROJE

Řešený vzorový příklad na analýzu panelových dat pomocí metody nejmenších čtverců najdete ve videonahrávce v LMS Moodle.

SHRNUTÍ KAPITOLY

Kapitola popisuje práci s panelovými daty a zabývá se analýzou panelových dat, konkrétně panelovou regresní analýzu. V posledních letech se lze v oblasti finanční ekonometrie setkávat právě s problematikou analýzy panelových dat. Panelová data jsou kombinací průřezových a časových řad. Panelová regresní analýzy (nebo také analýza panelových dat) je statisticko-ekonometrická metoda, která se zabývá analýzou vztahů a souvislostí dat ve dvourozměrném prostoru. První rozměr tvoří časová veličina, druhým rozměrem jsou průřezová data jednotlivých objektů pozorování. I u panelové regresní analýzy je nutnost mít stacionární data, proto je nutné nejprve provést test jednotkového kořene. K odstranění heterogenity v panelových modelech lze použít dva druhy individuálních efektů, a to modely fixních efektů a modely náhodných efektů. Individuální efekty tedy slouží k odstranění určité veličiny z panelových dat, která se liší mezi jednotlivými skupinami nebo v čase.

Při použití metody nejmenších čtverců je nutné nejprve mít splněny předpoklady, které jsou popsány v druhé kapitole. U panelových dat bývá jako estimační technika využívána kromě metody nejmenších čtverců, také zobecněná metoda momentů. Tato metoda je vhodná zejména v případě, že máme malé množství pozorování dat s malou frekvencí.

OTÁZKY

1. Pokud máme roční data za jednotlivé banky v českém bankovním sektoru za posledních 10 let a chceme zkoumat determinanty růstu poskytnutých úvěrů, jako individuální efekty použijeme zpravidla:

   a) fixní efekty

   b) náhodné efekty

   c) žádné

2. Pokud máme roční data za jednotlivé banky v českém bankovním sektoru za posledních 10 let a chceme zkoumat determinanty růstu poskytnutých úvěrů, jako vhodnou estimační metodu použijeme zpravidla:

   a) Metodu nejmenších čtverců

   b) Zobecněnou metodu náhodných momentů

   c) ARMA model

3. Hausmanův test testuje, zda odhad fixními a náhodnými efekty je výrazně odlišný. ANO x NE

4. Pokud zamítneme nulovou hypotézu u Hasumanova testu, použijeme náhodné efekty v panelové regresní rovnici. ANO x NE

5. Pro analýzu panelových dat je nutné použít stacionární data. ANO x NE

ODPOVĚDI

1. a)
2. b)
3. ANO
4. NE
5. ANO

7 NELINEARITA FINANČNÍCH ČASOVÝCH ŘAD A MODELY VOLATILITY

RYCHLÝ NÁHLED KAPITOLY

Modelování časových řad se ve velké míře využívá například ve finanční oblasti i v ekonomii obecně. Volatilita neboli kolísavost časových řad se nejčastěji studuje právě v oboru financí, protože určitým způsobem odráží například rizikovost cenných papírů na trhu. Aplikaci modelů volatility však lze nalézt například při oceňování opcí nebo řízení rizika. V této kapitole budou popsány charakteristické rysy volatility finančních časových řad. Bude věnována pozornost testování nelinearity časových řad. Budou přestaveny ARCH a GARCH modely. Pouze okrajově budou popsány asymetrické modely typu EGARCH a TARCH.

CÍLE KAPITOLY

- Definovat volatilitu.
- Definovat ARCH a GARCH modely.
- Popsat charakteristické rysy volatility časových řad.

KLÍČOVÁ SLOVA KAPITOLY

Volatilita, pákový efekt, ARCH, GARCH, podmíněný rozptyl, nelinearita časových řad.

7.1 Modelování a analýza volatility časových řad

Rozvoj finanční ekonometrie v posledních třiceti letech si vyžádal aplikací modelů a metod, které umožňují analyzovat a predikovat nejen očekávané zisky (výnosy) investorů, ale i velikost jejich rizika. K tomuto účelu jsou vhodné statistické postupy a ekonometrické modely, které jsou schopné popsat podmíněnou heteroskedasticitu, resp. volatilitu časových řad, znamenající období ve vývoji časových řad spojené s vysokou variabilitou či rostoucím rizikem (Hušek, 2009). V praxi, především ve finanční oblasti, se objevují řady, které nemají konstantní rozptyl (variabilitu). Tuto vlastnost dokážou zohlednit modely volatility.

Pokud bychom chtěli definovat volatilitu, najdeme řadu různých vysvětlení. Např. Kohout (2010) uvádí, že volatilita je číslo, které udává míru kolísavosti kursů akcií, měn, komodit nebo obligací. Z matematického hlediska je volatilita chápána jako jakákoliv míra rozptylu dané časové řady. Nemusí být přímo pozorovatelná, příkladem může být volatilita u dat s denní frekvencí, protože v jeden obchodní den je pouze jedno pozorování.

CHARAKTERISTICKÉ RYSY VOLATILITY

Pro volatilitu je typické, že není přímo pozorovatelná. Finanční ekonometrie rozlišuje mezi historickou a implikovanou volatilitou. Historická volatilita je reprezentována výběrovou směrodatnou odchylkou za určitý časový interval, která se používala i jako krátkodobá předpověď očekávané volatility finanční aktiv. Při vyjádření např. ceny opce pomocí modelu Blacka-Scholese lze využít cenu ke stanovení implikované volatility. Tak např. z pozorované ceny opce lze pomocí Blacka-Scholese formule odvodit podmíněnou směrodatnou odchylku, která představuje implikovanou volatilitu sledovaného aktiva, stejně jako index volatility. Implikovaná, resp. predikovaná volatilita se může lišit od aktuální volatility, protože v praxi nemusí být splněny teoretické předpoklady geometrického Brownova pohybu, takže např. změna logaritmu ceny není normálně rozdělena (Hušek, 2009).

Důležitou vlastností je shlukování volatility v čase (tzv. persistentní charakter volatility). V časových řadách výnosů finančních aktiv lze pozorovat shluky volatility, které se projevují tím, že volatilita může být vysoká v určitých obdobích a nízká v jiných obdobích. Je to způsobeno vlastností finančních dat, že velká rezidua (šoky) jsou následována dalšími velkými rezidui a obdobně malá rezidua (šoky) následují opět malá rezidua, takže velká, resp. malá rezidua vytvářejí shluky. To znamená, že období vysokého rizika se koncentrují a jsou střídána s nízkorizikovými obdobími nižší volatility, která se také koncentrují. Je zřejmé, že v těchto případech není splněn předpoklad homoskedasticity, takže není adekvátní vyjadřovat variabilitu nepodmíněným rozptylem, který predikuje úroveň rozptylu dlouhodobě, ale pomocí podmíněného rozptylu, založeného na modelech podmíněné heteroskedasticity.

Další vlastností volatility je to, že se vyvíjí plynuje, tj. spojitým způsobem, takže zřídka kdy je o vývoj skokem. Volatilita má zpravidla stacionární charakter neboli její změny nedivergují k nekonečně a ohybují se v určitém pevném rozmezí.

Volatilita výnosů finančních aktiv obvykle reaguje rozdílným způsobem na velký cenový růst nebo značný cenový pokles. Tento pákový efekt, kdy kladné a záporné šoky se v podmíněném rozptylu nepromítají symetricky, vyžaduje aplikovat asymetrické modely podmíněné heteroskedasticity.

K ZAPAMATOVÁNÍ

Typické vlastnosti volatility můžeme tedy shrnout takto:

- Shlukování volatility: volatilita může být v některých obdobích s vysoká a v jiných nízká.
- Pákový efekt: volatilita reaguje odlišně na cenový vzestup a na cenový pokles.
- Kontinuita: Volatilita se vyvíjí spíše spojitě bez nějakých výrazných skoků.
- Omezenost: Volatilita nesměřuje k extrémně vysokým (neomezeným) hodnotám, ale její průběh bývá spíše stacionární v určitém rozmezí.

Modely volatility

Volatilita se často používá k měření rizika. Pokud se volatilita zvětšuje, zvětšuje se i pravděpodobnost výskytu příslušného rizika, s nímž je úzce spojen i výnos z finančního instrumentu. Čím vyšší riziko, tím větší výnos. Nesmí být ale opomenuto, že místo vysokých výnosů může být dosaženo i vysoké ztráty. Výstupy modelů volatility jsou využívány například k analýzám na finančních trzích nebo měření míry tržního rizika investic metodou Value at Risk (VaR). Metoda VaR je v současné době standardizovaným přístupem k měření finančních rizik. Z hlediska statistiky představuje VaR kvantil uvažovaného pravděpodobnostního rozdělení. Zjednodušeně se jedná o odhad nejhorší ztráty plynoucí z daného typu rizik, která nebude překročena s předepsanou pravděpodobností ve stanoveném budoucím období. Další široké uplatnění nacházejí modely volatility při odhadech budoucích cen, kdy je nezbytná znalost podmíněného rozptylu pro konstrukci předpovědních intervalů.

Existuje celá řada modelů volatility, které se od sebe odlišují přístupem k definování podmíněného rozptylu. Nejstarším přístupem k volatilitě je tzv. historická volatilita, která se odhadovala jako výběrový rozptyl (resp. směrodatná odchylka) přes určité historické období. V současnosti význam tohoto modelu upadá a používá se už jen prakticky na stanovení srovnávacích hodnot (benchmarks) pro posouzení efektivnosti komplexnějších modelů volatility. Dalšími modely jsou model EWMA, který navazuje na historickou volatilitu, a implikovaná volatilita.

Jako přímá aplikace Boxovy-Jenkinsovy metodologie pro volatilitu byly zavedeny autoregresní modely volatility, které se využívali ve finanční praxi např. pro časové řady denních zlogaritmovaných zisků či cen finančních aktiv. Průlomem v modelování volatility byl model ARCH a jeho vylepšená forma GARCH. Jedná se o lineární modely volatility. Modifikace ARCH a GARCH modelů jsou modely IGARCH, GJR GARCH, EGARCH a další (Arlt a Arltová, 2003), které jsou považovány za modely nelineární. Tyto modely zachycují tzv. podmíněnou heteroskedasticitu, což znamená, že časové řady mají proměnlivou a většinou i nelineární funkci rozptylu, který je závislý na minulých hodnotách pozorování. Jejich význam spočívá v zachycení měnící se nejistoty na trhu, pomocí nichž lze empiricky ověřovat ekonomické a finanční teorie týkající se finančního trhu. Jedná se o modely, které dnes představují nejúspěšnější nástroj pro modelování finančních časových řad. Tyto modely mají však i své odpůrce.

Cílem všech modelů volatility je zachytit co nejlépe vývoj volatility časové řady finančních instrumentů v minulosti a co nejpřesněji předpovědět jejich budoucí vývoj.

7.1.1 MODELY ARCH

Základy modelu ARCH (Autoregressive Conditional Heteroskedasticity) uvedl Engle (1982). Model předpokládá tzv. podmíněný rozptyl, tedy závislost rozptylu časové řady na minulosti. Poprvé model použil na modelování inflace ve Velké Británii. Jedná se o model, ve kterém je volatilita jednoduchou kvadratickou funkcí minulých předpovědních chyb, odchylek od podmíněné úrovně (resp. střední hodnoty). Výhodou modelu ARCH je schopnost zachytit shluky volatility v časové řadě výnosů a vyšší špičatost rozdělení výnosů než má normální rozdělení. Na druhou stranu je jednou z nevýhod modelu předpoklad, že kladné a záporné šoky mají stejný vliv na volatilitu, protože závisí na druhé mocnině minulých šoků. V praxi je velmi dobře známo, že cena finančního aktiva reaguje jinak na pozitivní a negativní šoky.

Model ARCH(m) má tvar:

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 \,,$$

kde $\omega > 0, \alpha_i \geq 0$ pro $i > 0$.

Identifikace řádu modelu ARCH se provádí stejně jako u modelu AR pomocí parciálního korelogramu. Pro odhadování modelu se používají tzv. modely maximální věrohodnosti, což je softwarová záležitost.

Model ARCH má několik vlastností. V ARCH modelu je patrné, že velké absolutní hodnoty šoků v modelu způsobují velké hodnoty volatility v čase t, což ukazuje shlukování volatility. Při velké volatilitě v čase t je větší šance, že šok bude také velký, a díky tomu i volatilita v čase $t + 1$. Avšak velký rozptyl náhodné veličiny nemusí znamenat, že její hodnota bude daleko od střední hodnoty, pouze ukazuje zvýšenou pravděpodobnost, že se tak stane. Díky tomu nejsou shluky zvýšené volatility nekonečně dlouhé a po obdobích s vysokou volatilitou, přicházejí zase období s průměrnou či nízkou volatilitou. Další vlastností je, že volatilita v modelu ARCH reaguje stejně na dobré i na špatné zprávy.

Nevýhodu ARCH modelu je, že nevysvětluje, co a proč způsobuje volatilitu, pouze ji matematicky popisuje. Model ARCH (1) ukazuje, že šoky mají rozdělení s těžšími chvosty. Proto se častěji vyskytují odlehlá pozorování.

7.1.2 MODELY GARCH

Bollerslev (1986) navrhnul užitečné rozšíření modelu ARCH o zpožděný podmíněný rozptyl. Model GARCH (Generalized Autoregressive Conditional Heteroskedasticity) nahrazuje ARCH v případech, kdy je nutné odhadovat velké množství parametrů αi (ARCH(q) model má vysoký stupeň q).

Model GARCH(p,q)

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2,$$

kde $\omega > 0$, $\alpha_i > 0$ pro $i = 1, 2, ..., q$ a $\beta_j \geq 0$ pro $j = 1, 2, ..., p$.

Model ARCH(m) často vyžaduje vysoké řády m a nezohledňuje pákový efekt finančních dat. Model GARCH odstraňuje problém s vysokými řády. Jedná se o model ARCH rozšířený o zpožděný podmíněný rozptyl. Model ARCH nebyl ovlivněn minulými hodnotami podmíněného rozptylu. Model GARCH je již o minulé hodnoty podmíněného rozptylu rozšířen. Nejjednodušším a nejpoužívanějším modelem při zkoumání volatility finančních časových řad je GARCH (1,1), který představuje rozšířený model ARCH (1) o podmíněný rozptyl v prvním zpoždění. Rovnice modelu GARCH (1,1) má jednoduchý tvar

$$y_t = \mu_t + e_t, \quad e_t = \sigma_t \varepsilon_t, \qquad \sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{24}$$

za podmínek $\alpha_0 > 0$, $\alpha_1, \beta_1 \geq 0$, $\alpha_1 + \beta_1 < 1$, které zaručují kladný podmíněný rozptyl a stacionaritu v kovariancích. Model GARCH(1,1) lze prakticky použít tam, kde by bylo vhodné použít model ARCH s mnoha zpožděními. (Arlt & Arltová, 2003).

Výhodou modelu GARCH je, že ho lze použít tam, kde by bylo dobré zvolit model ARCH(m) s mnoha zpožděními m. Odstraňuje tento problém pomocí minulých hodnot zahrnutých v modelu. I přes vylepšení modelu ARCH o minulé hodnoty nebere ani GARCH model ohled na pákový efekt.

MODELY EGARCH

Model EGARCH (Exponential Generalized Autoregressive Conditional Heteroskedasticity) je jedním z nejvýznamnějších nelineárních modelů volatility. Model byl publikován v roce 1991 D.B. Nelsonem. Cílem modelu EGARCH je zachytit tzv. asymetrické efekty ve finančních časových řadách. Nejznámějším je pákový efekt, při kterém je vliv záporných šoků na hodnotu podmíněného rozptylu výrazně vyšší nežli vliv šoků kladných. Pákový efekt pojmenoval F. Black (1976).

Model EGARCH popisuje vztah mezi logaritmem podmíněného rozptylu a minulým šoky. Zápis pomocí logaritmických volatilit má výhodu v tom, že podmínky nezápornosti parametrů jsou nyní irelevantní a pákový efekt je exponenciální nikoli kvadratický. Navíc absence podmínek nezápornosti parametru usnadňuje jejich odhad. Model podmíněného rozptylu má následující tvar:

EGARCH (1, 1):

$$\log \sigma_t^2 = \varpi + \beta \log \sigma_{t-1}^2 + \gamma \frac{\varepsilon_{t-1}}{\sigma_{t-1}} + \alpha \left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right|$$

EGARCH (p, q):

$$\log \sigma_t^2 = \varpi + \sum_{j=1}^{q} \beta_j \log \sigma_{t-1}^2 + \sum_{j=1}^{p} \left(\alpha_j \left| \frac{\varepsilon_{t-j}}{\sigma_{t-j}} \right| + \gamma_j \frac{\varepsilon_{t-j}}{\sigma_{t-j}} \right)$$

Platí, že pro popis případné asymetrie je důležitá hodnota parametrů γi. Je-li různá od nuly, asymetrie se v modelu vyskytuje. Je-li hodnota parametru záporná, existuje v časové řadě pákový efekt, tedy vyšší vliv záporných šoků než šoků kladných. Je-li hodnota γi kladná, je asymetrický efekt opačný. Kladné šoky v takovém případě zvyšují volatilitu časové řady více než šoky záporné.

Model TGARCH

Model TGARCH (Threshold Generalized Autoregressive Conditional Heteroskedasticity) je vytvořen Zakoian (1990). Model autor dále dopracoval spolu Rabemananjarem (1993) do tvaru:

$$\sigma_t^2 = \varpi + \sum_{j=1}^{p} \alpha_j \varepsilon_{t-j}^2 + \sum_{k=1}^{r} \gamma_k \varepsilon_{t-k}^2 d_{t-k}$$

TGARCH (1, 1):

$$\sigma_t^2 = \varpi + \alpha \varepsilon_{t-1}^2 + \gamma \varepsilon_{t-1}^2 d_{t-1}$$

Vliv kladných šoků je popsán parametry αi a vliv šoků záporných parametry γk. Pákový efekt se v řadě vyskytuje pokud platí nerovnost $\alpha i < \gamma k$. Jednoduše lze říci, že v tomto modelu mají špatné a dobré zprávy odlišný vliv na podmíněný rozptyl.

Rozdíl mezi GARCH a TGARCH modelem je v tom, že GARCH pracuje pouze v jednom režimu a všechna pozorování mají vliv na parametry modelu. Prahové modely pracují ve dvou a více režimech, které rozlišují různé podmínky při vzniku volatility. Data ze strukturálně podobných režimů pak ovlivňují parametry jen z příslušného režimu, přičemž přepínání mezi jednotlivými režimy pak probíhá na základě nějaké známé veličiny y_t. Tato specifikace dává možnost zakomponovat do modelu asymetrii dat a model dovoluje volatilitě reagovat odlišně na minulé šoky v závislosti na jejich znaménku. V porovnání s modelem EGARCH, který také umí zachytit asymetrii volatility, výhoda modelu TGARCH spočívá vtom, že poměr dopadu pozitivního šoku a dopadu negativního šoku se v průběhu času může měnit, je flexibilnější.

VÝSTAVBA MODELŮ VOLATILITY

Při konstrukci modelů volatility lze postupovat pomocí těchto kroků:

1. Pro časovou řadu se určí vhodný úrovňový model. Např. vhodným ARMA modelem se z dané časové řady odstraní případné lineární závislosti a dále se pracuje s rezidui tohoto modelu.

2. Otestuje se, zda je v řadě reziduí přítomná podmíněná heteroskedasticita. Testuje se nulová hypotéza podmíněné homoskedasticity proti alternativní hypotéze podmíněné heteroskedasticity.

3. V případě podmíněné heteroskedasticity je třeba určit řád modelu GARCH a odhadnout parametry tohoto modelu.

4. Ověření vhodnosti daného modelu diagnostickými testy.

5. Použití modelu pro predikční účely.

Pro diagnostiku modelu lze použít testy diagnostické kontroly úrovňového lineárního modelu, a to testy autokorelace a normality nesystematické složky. Dále lze využít Box-Ljungovu statistiku nebo test podmíněné heteroskedasticity reziduí modelu.

DALŠÍ ZDROJE

Řešený vzorový příklad na modely volatility v EViews 11 najdete ve videonahrávce v LMS Moodle.

SHRNUTÍ KAPITOLY

Poslední kapitola studijního textu se zabývá nelinearitou časových řad a modely volatility. V kapitole bylo uvedeno, že typickým rysem volatility je, že není přímo pozorovatelná. Mezi typické vlastnosti volatility patří shlukování volatility, pákový efekt, kontinuita (což znamená, že se volatilita vyvíjí spíše spojitě bez nějakých výrazných skoků) a omezenost. Ve financích se volatilita často využívá k měření rizika. Protože pokus se zvyšuje volatilita, zvyšuje se i pravděpodobnost výskytu příslušného rizika. K měření volatility lze využít řadu modelů. Cílem všech modelů volatility je zachytit co nejlépe vývoj volatility časové řady finančních instrumentů v minulosti a co nejpřesněji předpovědět jejich budoucí vývoj. V kapitole byl představen model ARCH, GARCH a okrajově i model EGARCH a TGARCH.

OTÁZKY

1. Výstupy modelů volatility lze využít ve financích např. k měření míry tržního rizika investic metodou Value at Risk. ANO x NE

2. Pákový efekt znamená, že volatilita může být v některých obdobích vysoká a v jiných nízká. ANO x NE

3. Vlastnost, že se volatilita vyvíjí spíše spojitě bez výrazných šoků se nazývá:
   a) shlukování volatility
   b) pákový efekt
   c) kontinuita

4. TGARCH model je vhodný použít, pokud chceme do modelu zakomponovat asymetrii dat a zjistit, zda mají dobré a špatné zprávy odlišný vliv na volatilitu. ANO x NE

5. Model ARCH modeluje vliv minulých hodnot podmíněného rozptylu. ANO x NE

ODPOVĚDI

1. ANO
2. NE
3. c)
4. ANO
5. NE

LITERATURA

[1] ARLT, J. a M. ARLTOVÁ, 2009. Ekonomické časové řady. Praha: Professional Publishing. ISBN 978-80-86946-85-6.

[2] ASTERIOU, D. a S.G. HALL, 2011. *Applied Econometrics*. 2nd ed. Houldmills, Palgrave Macmillan. ISBN 978-0-230-27182-1.

[3] BREITUNG, J., 2000. The Local Power of Some Unit Root Tests for Panel Data. In: B. Baltagi, B., ed. *Advances in Econometrics*, **15**: Nonstationary Panels, Panel Cointegration, and Dynamic Panels. Amsterdam: JAI Press, s. 161–178. ISBN 978-0-76230-688-6.

[4] CIPRA, T., 2014. Finanční ekonometrie. Praha: Ekopress. ISBN 978-80-86929-93-4.

[5] CROARKIN, P.E., C. A. WALL, S.M. MCCLINTOCK, F.A. KOZEL, M.M. HUSAIN a S.M. SAMPSON, 2010. The emerging role for repeti-tive transcranial magnetic stimulation in optimizing the treat-ment of adolescent depression. *Journal of ECT*, **26**, 323-329.

[6] ENGLE, R.F. a C.W.J. GRANGER, 1987. Co-Integration and Error Correction: Representation, Estimation and Testing. *Econometrica*, **55**, 251–276. ISSN 1468-0262.

[7] EVANS, J. D., 1996. *Straightforward statistics for thebehavioral sciences*. Pacific Grove, CA: Brooks Cole Publishing.

[8] GRANGER, C., 1969. Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica*, **37** (3), 424-438.

[9] HADRI, K., 2000. Testing for Stationarity in Heterogeneous Panel Data. *Econometric Journal*, **3**, 148–161. ISSN 1368-423X.

[10] HANČLOVÁ, J., 2012. *Ekonometrické modelování: klasické přístupy s aplikacemi*. Praha: Professional Publishing. ISBN 978-80-7431-088-1.

[11] HUŠEK, R. a J. PELIKÁN, 2003. *Aplikovaná ekonometrie: teorie a praxe*. Praha: Professional Publishing. ISBN 80-86419-29-0.

[12] HUŠEK, R. a J. WALTER, 1976. Ekonometrie. Praha: Nakladatelství technické literatury.

[13] HUŠEK, R., 2009. Aplikovaná ekonometrie: teorie a praxe. Praha: Oeconomica. ISBN 978-80-245-1623-3.

[14] CHOI, I., 2001. Unit Root Tests for Panel Data. *Journal of International Money and Finance*, **20**, 249–272. ISSN 0261-5606.

[15] IM, K.S., M.H. PESARAN a Y. SHIN, 2003. Testing for Unit Roots in Heterogeneous Panels. Journal of Econometrics, 115, 53–74. ISSN 0304-4076.Maddala a Wu (1999),

[16] KLÍMEK, P., 2006. Ekonometrie. Studijní pomůcka pro distanční studium. Zlín: Univerzita Tomáše Bati. ISBN 80-7318-428-1.

[17] LEVIN, A., C.F. LIN a C. CHU, 2002. Unit root test in panel data: Asymptotic and finite sample properties. Journal of Econometrics, 108, 1–25. ISSN 0304-4076.

[18] NOVÁK, P., 2007. Analýza panelových dat. *Acta Oeconomica Pragensia*, **15**(1), 71–78. ISSN 0572-3043.

[19] POLOUČEK, S., 2009. *Peníze, banky a finanční trhy*. Praha: C.H. Beck. ISBN 978-80-7400-152-9.

[20] PRAŽÁK, T., 2020. *Vliv makroekonomického prostředí na finanční situaci podniků.* Karviná: SU OPF.

[21] RŮČKOVÁ, P., 2016. *Vliv vybraných determinantů teorií kapitálové struktury na rozložení použitých zdrojů financování v podnicích v zemích Visegradské čtyřky*. Karviná: SU OPF.

[22] ŘEPKOVÁ, I., 2012. *Konkurence, koncentrace a efektivnost v českém bankovním sektoru*. Karviná: SU OPF.

[23] VERBEEK, M., 2004. *A Guide to Modern Econometrics.* 2nd ed. Chichester: John Wiley and Sons Ltd. ISBN 0-470-85773-0.

SHRNUTÍ STUDIJNÍ OPORY

Řada subjektu z finanční oblasti vyžaduje hlubší znalosti vztahů mezi ekonomickými, či konkrétně finančními veličinami. Znalost vztahů mezi ekonomickými veličinami je velmi důležitá pro jednotlivé oblasti. Disciplína, které tyto vztahy dokáže popsat či měřit je právě ekonometrie. Studijní opora Finanční ekonometrie se tedy věnuje teoretickému vymezení základních oblastí a metod využitelných v oblasti financí či ekonomie obecně.

Nejprve studijní opora vysvětluje základní pojmy, jako je ekonometrický model a jednotlivé typy dat. Stěžejní je popis základních kroků formulace ekonometrického modelu. Bez znalosti této problematiky není možné zkonstruovat a pochopit jednotlivé modely popisované ve studijní opoře. Dále byly ve studijním textu popsány základní metody lineárního regresního modelu, která je jednou z nejdůležitějších metod finanční ekonometrie. Pozornost byla podrobně věnována odhadové metodě parametrů, tedy metodě nejmenších čtverců. Aplikace metody nejmenších čtverců vyžaduje splnění jednotlivých předpokladů. Proto byly jednotlivé předpoklady konkrétně definovány a popsány možnosti jejich testování. Pro použití metody nejmenších čtverců je tedy nutné testovat multikolinearitu, autokorelace reziduální složky, heteroskedasticitu, normalitu a autokorelaci reziduální složky.

Modely jednorozměrných a vícerozměrných časových řad byly popsány v třetí části studijního textu. V další části byla pozornost věnována kauzalitě finančních časových řad, kde byla podrobně popsána a testována korelační analýza a Grangerova kauzalita. V další části byla popsána práce s nestacionárními časovými řadami a definována kointegrace. V případě, že časové řady mají kointegrační vazbu, je nutné aplikovat model korekce chyby k určení krátkodobých odchylek od dlouhodobé rovnováhy.

V další části byla věnována pozornost práci s panelovými daty. Byly definována panelová data, která jsou zejména ve finanční ekonometrii často předmětem zkoumání a znalost analýzy panelových dat přináší ekonometrovi širší možnosti zkoumání a práce s daty. V poslední kapitole byly popsány modely volatility. Definovány byly základní modely volatility ARCH a GARCH modely a jejich využití.

Věřím, že po přečtení studijního textu jste zde našli důležité informace, které využijete nejen u zkoušky předmětu Finanční ekonometrie, ale budete schopni propojit získané informace i s jinými studovanými kurzy a zejména uplatníte tyto informace i v budoucí praxi. Protože studijní opora není sbírka řešených příkladů, je studijní text doplněn videonahrávkami, kde jsou konkrétně aplikovány jednotlivé modely a představeno jejich řešení v ekonometrickém software EViews 11. Ekonometrický program EViews byl zvolen zejména z důvodu snadného využití a grafického rozhraní, což je pro studenty snadněji použitelné. Lze však využít i dalších ekonometrických programů, jako je STATA, SPSS nebo R-project, což je open source program.

Iveta Palečková

PŘEHLED DOSTUPNÝCH IKON

Čas potřebný ke studiu	Cíle kapitoly
Klíčová slova	Nezapomeňte na odpočinek
Průvodce studiem	Průvodce textem
Rychlý náhled	Shrnutí
Tutoriály	Definice
K zapamatování	Případová studie
Řešená úloha	Věta
Kontrolní otázka	Korespondenční úkol
Odpovědi	Otázky
Samostatný úkol	Další zdroje
Pro zájemce	Úkol k zamyšlení

Název: **Finanční ekonometrie**

Autor: **Doc. Ing. Iveta Palečková, Ph.D.**

Vydavatel: Slezská univerzita v Opavě

Obchodně podnikatelská fakulta v Karviné

Určeno: studentům SU OPF Karviná

Počet stran: 122

Tato publikace neprošla jazykovou úpravou.