Aplikovaná statistika 2.

Veronika Svobodová
email: schneck@email.cz

Organizace výuky

- **11. 3. 2007** 9.20 – 11.50 **tutoriál 1**
- **22. 4. 2007** 9.20 – 11.50 **tutoriál 2**
- **6. 5. 2007** odevzdat POT1 (str. 42)
- **20. 5. 2007** 12.15 – 17.05 **tutoriál 3**
- **3. 6. 2007** odevzdat POT2 (str. 76)

Organizace zkoušky

- vypracování a zaslání POT na adresu: schneck@email.cz
  - POT 1 do dvou týdnů po 2. tutoriálu
  - POT 2 do dvou týdnů po 3. tutoriálu
- test s použitím programu Excel
  - zadání na webových stránkách
  - příklady (nutné komentáře výsledků, správné spočtení Excelem bez interpretace je nedostačující)
  - teoretické otázky

Literatura

- Špalek, J.: *Aplikovaná statistika II* (Distanční studijní opora), MU Brno 2004
- Seger, J., Hindls, R., Hronová, S.: *Statistika v hospodářství*, Praha ETC Publishing, 1998

Obsah předmětu

- Regresní a korelační počet
  - tvorba regresní funkce (1. tutoriál)
  - kvalita regresní funkce (2. tutoriál)
- Časové řady (3. tutoriál)
- Souhrnné cenové indexy (3. tutoriál)
  - užití v praxi (samostudium)

Ekonomické časové řady

- studium ekonomických jevů a veličin v čase
  - možnost posouzení minulého vývoje a odhadu do budoucnosti
  - úzká souvislost s regresní analýzou (nezávisle proměnnou je čas)
- časová řada je posloupnost vzájemně srovnatelných dat uspořádaných chronologicky
  - př.: hodnota HDP v jednotlivých letech, nezaměstnanost v jednotlivých měsících, ...

Ekonomické časové řady

- dělení časových řad (správné vymezení nutné pro lepší pochopení jevů):
  - dle časového hlediska:
    - intervalové (potřeba stejně dlouhých intervalů!)
    - okamžikové
  - dle periodicity:
    - dlouhodobé (roční a delší cykly)
    - krátkodobé (odstupy kratší než rok)
  - dle druhu údajů:
    - absolutní (zjištěné, empirické hodnoty)
    - odvozené (lépe čitelná změna)

Ekonomické časové řady

- dělení časových řad:
  - dle vyjádření údajů:
    - naturální (jevy v jednotkách jiných než peněžních – kusy, ...)
    - peněžní (ať už přímé nebo přepočítané z naturálních)

- požadavky na srovnatelnost časových řad:
  - <u>věcnou</u> – řada musí zachycovat jedinou skutečnost, resp. musí být stále stejně tvořena – nemůžeme použít různé metody výpočtu apod.

Ekonomické časové řady

- požadavky na srovnatelnost časových řad:
  - <u>prostorovou</u> – řada by měla čerpat data ze stále stejného „ekonomického prostoru“ (problémy při změně státních hranic, reorganizaci podniku apod.)
  - <u>časovou</u> – intervalové řady by měly být vztaženy vždy ke stejnému časovému intervalu (počtu pracovních dní apod.)
    - časové nesrovnatelnosti z důvodů kalendářních variací se zbavíme tzv. kalendářním očištěním – na kalendářní nebo na pracovní dny
    - příklad 3.1

$$y_t^{(0)} = y_t * \frac{\overline{k_t}}{k_t}$$

Ekonomické časové řady

- požadavky na srovnatelnost časových řad:
  - cenovou – jen u řad vyjádřených v peněžních jednotkách
    - běžné ceny (empirické) bývají převáděny na stálé ceny (ceny určitého roku) očištěné od vlivu inflace
- elementární charakteristiky časových řad:
  - diference
  - tempo růstu
  - průměrné tempo růstu

Ekonomické časové řady

- elementární charakteristiky časových řad:
  - diference – rozdíly hodnot v časové řadě
    - diference prvního řádu – rozdíl sousedních pozorování v časové řadě

$$D_t^1 = y_t - y_{t-1}$$

    - diference *n*-tého řádu je počítána jako rozdíl sousedních diferencí *(n-1)* řádu

$$D_t^n = D_t^{n-1} - D_{t-1}^{n-1}$$

    - diference vyšších řádů podávají informaci o vývoji změn v časové řadě

Ekonomické časové řady

- elementární charakteristiky časových řad:
  - <u>tempo růstu</u> udává okamžitý relativní přírůstek hodnoty v časové řadě (diference v procentním vyjádření)

$$T_t = \frac{y_t}{y_{t-1}}$$

  - <u>průměrné tempo růstu</u> udává jaká byla průměrná relativní změna proměnné, počítáme ji jako geometrický průměr jednotlivých temp růstu.
  - příklad 3.2

$$\overline{T} = \sqrt[n]{T_1 * T_2 * \dots * T_n}$$

Ekonomické časové řady

- modelování časových řad:
  - základní úlohou je nalezení obecné tendence ve vývoji časové řady (trendu)
  - omezíme se na jednorozměrný (v roli vysvětlující proměnné je pouze čas) klasický (náhodná složka nehraje velkou roli) model
  - klasický model vychází z dekompozice na složky:
    - trendové – T
    - sezónní – S
    - cyklické – C (nebudeme v podstatě uvažovat)
    - náhodné – ε

Ekonomické časové řady

- klasický model je popsán pomocí aditivního tvaru:

$$y_t = T_t + S_t + C_t + \varepsilon_t$$

- východiskem nalezení vhodné trendové funkce je věcná ekonomická analýza (existují reálné důvody pro růst ukazatele v čase?)
- kritéria pro volbu trendové funkce
  - metody vycházející z regresního počtu (min. čtverce)
  - index korelace (hledáme co nejvyšší index korelace)

Ekonomické časové řady

- kritéria pro volbu trendové funkce
  - statistické testy a intervaly spolehlivosti
  - analýza diferencí a tempa růstu
  - klouzavý průměr (v praxi velmi oblíbená metoda spočívající v nahrazení empirických hodnot klouzavými průměry počítanými z okolních hodnot)
  - extrapolační kritéria
    - vybereme z časové řady časový úsek a na jeho základě se snažíme simulovat následující (již známá) pozorování – kritériem je potom co největší shoda empirických a simulovaných hodnot

Ekonomické časové řady

- ad analýza diferencí
  - v momentu, kdy jsou diference některého řádu v podstatě konstantní, lze řadu popsat trendovou funkcí řádu stejného jako diference
  - příklad
  - analogie u indexů (je-li tempo růstu konstantní, časovou řadu popisujeme exponenciálním trendem)
  - příklad 4.1

Ekonomické časové řady

- ad klouzavý průměr
  - postupné nahrazování empirických hodnot klouzavými průměry (průměry, které získáme z hodnoty, jež nahrazujeme a určitého počtu hodnot předcházejících a následujících)
  - „kloužeme“ tedy po řadě a počítáme postupně příslušné aritmetické průměry
  - obvyklý je lichý počet členů (3, 5, 7, ...)

$$\overline{y}_t = \frac{y_{t-\frac{m-1}{2}} + y_{t-\frac{m-1}{2}+1} + \ldots + y_t + \ldots + y_{t+\frac{m-1}{2}-1} + y_{t+\frac{m-1}{2}}}{m}$$

Ekonomické časové řady

- ad klouzavý průměr
  - klouzavá část má délku rovnu sudému číslu – centrované klouzavé průměry
    - spočítáme je jako aritmetický průměr dvou klouzavých průměrů pro dané období – jeden má o jednu hodnotu více předcházejících hodnot než následujících a druhý naopak
  - příklad 4.2

Ekonomické časové řady

- významné trendové funkce
  - lineární trend
    - používán k získání výchozí informace o časové řadě
    - aproximace složitější trendové funkce (pomocí kratších časových intervalů)
    - parametry *a* odhadujeme např. pomocí metody nejmenších čtverců
    - t je časová proměnná nabývající přirozených hodnot

$$T_t = a_0 + a_1 * t$$

  - parabolický trend

$$T_t = a_0 + a_1 * t + a_2 * t^2$$

Ekonomické časové řady

- významné trendové funkce
  - exponenciální trend

$$T_t = a_0 + a_1^t$$

  - logistický trend
    - tvar písmene *S*
    - možno rozlišit 5 fází
  - Gompertzova křivka (také omezená)

Ekonomické časové řady

- analýza sezónní složky
  - tvoří pravidelně opakující se odchylku od trendové složky
  - odraz klimatických cyklů (jen u časových řad kratších než rok)
  - postup: identifikace a vyloučení sezónní složky
  - vyloučení = sezónní očišťování, např. takto:
    - nahrazením původní řady řadou vhodně zvolených klouzavých průměrů – odstraní sezónní složky, jejichž perioda nepřesahuje délku klouzavého průměru
    - původní hodnoty vydělíme příslušným sezónním faktorem

Ekonomické časové řady

- analýza sezónní složky
  - výpočet sezónního faktoru
    - pomocí sezónního indexu
    - nejprve určíme vhodné klouzavé průměry z časové řady
    - sezónní index je podílem empirické a vyrovnané hodnoty
    - spočítáme průměrné sezónní indexy pro sledované sezóny (např. pro čtvrtletí roku)
    - tyto hodnoty musíme znormovat, aby jejich součet byl roven počtu sledovaných sezón (každou z nich vynásobíme počtem sezón a vydělíme součtem průměrných sezónních indexů)
    - znormovaná hodnota se nazývá sezónní faktor

Ekonomické časové řady

- příklad 5.1
- analýza cyklické složky
  - kolísání kolem trendu s periodou delší než jeden rok
  - cykly nemusí být stejně dlouhé – problematické očištění
- analýza náhodné složky
  - po sezónním a cyklickém očištění jsou všechny odchylky od trendu zahrnovány do náhodné složky
  - nelze popsat funkcí času

Ekonomické časové řady

- analýza náhodné složky
  - jestliže má nulovou střední hodnotu a konstantní rozptyl, nazývá se *bílý šum*
  - autoregrese náhodných poruch (náhodné složky v časové řadě na sobě závisí)
    - Durbin-Watsonův test autokorelace
    - nulová hypotéza je nezávislost
    - testové kritérium

$$d = \frac{\sum_{i=1}^{n} \left(\varepsilon_t - \varepsilon_{t-1}\right)^2}{\sum_{i=1}^{n} \varepsilon_t^2}$$

Ekonomické časové řady

- analýza náhodné složky
  - hodnoty kritéria se pohybují od nuly do čtyř
  - ve prospěch zamítnutí nulové hypotézy hovoří krajní hodnoty kritéria (0 ... přímá závislost, 4 ... nepřímá závislost)
  - příklad 5.2
- příklad 5.3 (ukázka předdefinovaných funkcí v programu Excel)

Ekonomické časové řady

- POT2
  - body a) a b) se vztahují pouze k roku 2000, ostatní body k celé časové řadě
  - u bodu e) proveďte očištění oběma způsoby (za sezónu považujte čtvrtletí)
  - u bodu d) se posouzením kvality myslí provedení testu spolehlivosti (viz regresní počet)
  - při výpočtu náhodné složky předpokládejte nulovou cyklickou složku

Souhrnné cenové indexy

- popisují míru změny sledovaného ukazatele
- členění ukazatelů
  - dle zjišťování
    - primární (přímo zjištěné, neodvozené – např. počty výrobků)
    - sekundární (odvozené – průměry, rozdíly, podíly primárních ukazatelů)

Souhrnné cenové indexy

- popisují míru změny sledovaného ukazatele
- členění ukazatelů
  - dle zjišťování
  - dle vyjádření
    - absolutní – velikost bez vztahu (především primární)
    - relativní – vždy sekundární (ve formě podílu)

Souhrnné cenové indexy

- popisují míru změny sledovaného ukazatele
- členění ukazatelů
  - dle zjišťování
  - dle vyjádření
  - dle intenzity
    - extenzitní – ukazatele množství (počet kusů, ...)
    - intenzitní – ukazatele úrovně (většinou cena)

Souhrnné cenové indexy

- popisují míru změny sledovaného ukazatele
- členění ukazatelů
  - dle zjišťování
  - dle vyjádření
  - dle intenzity
  - dle způsobu shrnování v čase
    - okamžikové
    - intervalové

Souhrnné cenové indexy

- popisují míru změny sledovaného ukazatele
- členění ukazatelů
  - dle zjišťování
  - dle vyjádření
  - dle intenzity
  - dle způsobu shrnování v čase
- zajímá-li nás absolutní změna ukazatele, hovoříme o rozdílu
- zajímá-li nás relativní změna ukazatele, hovoříme o indexu

Souhrnné cenové indexy

- indexy
  - jednoduché indexy – srovnávají dvě hodnoty téhož ukazatele, jehož hodnoty nejsou nijak členěny ani shrnovány
    - způsob výpočtu: srovnáme hodnotu v čase 1 (běžné období) s hodnotou v čase 0 (základní období)
    - příklad 6.1
- sdružíme-li indexy do časových řad, můžeme z nich zjistit *bazické a řetězové indexy*

Souhrnné cenové indexy

- bazické indexy
  - indexy počítané ke stále stejnému základnímu období
  - udávají vývoj ukazatele vzhledem k pevně zvolenému období

$$I_0 = \frac{q_0}{q_0}, I_1 = \frac{q_1}{q_0}, \ldots, I_n = \frac{q_n}{q_0}$$

Souhrnné cenové indexy

- řetězové indexy
  - indexy počítané vzhledem k předchozímu období
  - udávají zvýšení/snížení ukazatele vzhledem k předchozímu období

$$I_1 = \frac{q_1}{q_0}, I_2 = \frac{q_2}{q_1}, \ldots, I_n = \frac{q_n}{q_{n-1}}$$

Souhrnné cenové indexy

- vztah mezi bazickými a řetězovými indexy
  - vynásobením dvou sousedních řetězových indexů získáme příslušný bazický index
  - podělením dvou sousedních bazických indexů získáme příslušný řetězový index
  - příklad 6.2

Souhrnné cenové indexy

- souhrnné indexy
  - charakterizují změnu nestejnorodého ukazatele (přičemž všechny položky nemusí mít stejnou váhu)
  - 1. generace – prosté průměry
    - průměry (aritmetický, harmonický, geometrický)
    - nevhodnost plyne z toho, že se implicitně předpokládá stejná váha u každé položky
  - 2. generace – vážené průměry (nejčastější použití)
    - Laspayresův index (úroveň spotřeby základního období)

$$I^{(L)} = \frac{\sum_{i=1}^{n} p_{1,i} \cdot q_{0,i}}{\sum_{i=1}^{n} p_{0,i} \cdot q_{0,i}}$$

Souhrnné cenové indexy

- souhrnné indexy
  - 2. generace – vážené průměry
    - Paascheho index (současná úroveň spotřeby)

$$I^{(P)} = \frac{\sum_{i=1}^{n} p_{1,i} \cdot q_{1,i}}{\sum_{i=1}^{n} p_{0,i} \cdot q_{1,i}}$$

  - 3. generace
    - Fischerův index (geometrický průměr indexů 2. generace
  - příklad 6.3

Děkuji za pozornost.