

III. Časové řady

Časová řada (*Time-Series*)

– posloupnost věcně a prostorově srovnatelných pozorování (dat) uspořádaných podle času

$$\{y_t : t \in \mathcal{T}\}$$

t – časová proměnná

\mathcal{T} – množina hodnot časové proměnné t

- *Analýza* časových řad – soubor metod sloužících k popisu č.ř.
- *Prognóza* časových řad – soubor metod sloužících k předvídání budoucího chování č.ř.

Chronologicky uspořádaná data:

- Fyzika, biologie, seismologie, meteorologii aj.
- **Ekonomie:**
 - **makroekonomické ukazatele** (tvorba a užití hrubého domácího produktu, inflace, vývoj nezaměstnanosti)
 - **dílčí údaje** (vývoj kurzů cizích měn, peněžní zásoby, vývoj cen akcií na kapitálovém trhu)

Třídění časových řad ekonom. ukazatelů:

a) podle rozhodného časového okamžiku

- ◇ *intervalové* (hodnota ukazatele zjišťována za určitý časový interval, např. zisk firmy za čtvrtletí)
- ◇ *okamžikové* (hodnota ukazatele zjišťována k určitému časovému okamžiku, např. počet zaměstnanců k poslednímu dni v měsíci)

b) podle periodicity

- ◇ *dlouhodobé* (roční, víceletá pozorování)
- ◇ *krátkodobé* (perioda do 1 roku - týdenní, měsíční, čtvrtletní aj.)

c) podle druhu sledovaných ukazatelů

- ◇ *č.ř. primárních (prvotních) ukazatelů*
 - přímo sledované ukazatele, např. počet pracovníků k určitému datu, stav zásob apod.
- ◇ *č.ř. sekundárních (odvozených) charakteristik*

d) podle způsobu vyjádření údajů

- ◇ *č.ř. naturálních ukazatelů* (hodnoty ukazatele vyjadřovány v naturálních jednotkách)
- ◇ *č.ř. peněžních ukazatelů*

◇ Intervalové časové řady

Hodnoty y_t závisí na délce časového intervalu, ve kterém je znak sledován.

Kalendářní očišťování č.ř. od důsledků kalendářních variací

– přepočítávání všech období na jednotkový interval za účelem porovnávání hodnot ukazatelů:

- Údaje očištěné na kalendářní dny:

$$y_t^{(0)} = y_t \frac{\bar{k}_t}{k_t}$$

y_t – hodnota očišťovaného ukazatele v příslušném dílčím období roku

k_t – počet kalendářních dní v příslušném dílčím období roku

\bar{k}_t – průměrný počet kalendářních dní v příslušném dílčím období roku

- Údaje očištěné na pracovní dny:

$$y_t^{(0)} = y_t \frac{\bar{p}_t}{p_t}$$

p_t – počet pracovních dní v příslušném dílčím období roku

\bar{p}_t – průměrný počet pracovních dní v příslušném dílčím období roku

Příklad (TS1.):

Údaje o čtvrtletních tržbách nějaké firmy (v tisících) v roce 2003 přepočítáme tak, aby byl vyloučen vliv rozdílného počtu dní v jednotlivých čtvrtletích:

Čtvrtletí	Počet dní	Tržba (y_t)
I.	90	50
II.	91	60
III.	92	40
IV.	92	50

Průměrný počet kalendářních dní ve čtvrtletí:

$$\bar{k}_t = 365/4 = 91,25$$

Pro I. čtvrtletí provedeme výpočet:

$$y_I^{(0)} = 50000 \frac{91,25}{90} = 50694.443$$

- **Elementární charakteristiky čas. řad:**

y_1, y_2, \dots, y_n – hodnoty ukazatelů pro k -časových okamžiků t_1, t_2, \dots, t_n ($t_1 < t_2 < \dots < t_n$)

- Průměry

- ◊ Intervalové řady:

- *aritmetický* průměr:

$$\bar{y} = \frac{1}{n} \sum_{t=1}^n y_t$$

- ◊ Okamžikové řady:

- *prostý chronologický* průměr: stejná vzdálenost mezi okamžiky pozorování

$$\begin{aligned} \bar{y} &= \frac{1}{n-1} \left(\frac{y_1 + y_2}{2} + \frac{y_2 + y_3}{2} + \dots + \frac{y_{n-1} + y_n}{2} \right) \\ &= \frac{1}{n-1} \left(\frac{y_1}{2} + y_2 + \dots + y_{n-1} + \frac{y_n}{2} \right) \end{aligned}$$

- *vážený chronologický* průměr: nesterjná vzdálenost mezi okamžiky pozorování

$$\bar{y} = \frac{\frac{y_1+y_2}{2}(t_2 - t_1) + \frac{y_2+y_3}{2}(t_3 - t_2) + \dots + \frac{y_{n-1}+y_n}{2}(t_n - t_{n-1})}{(t_n - t_1)}$$

○ Diference (přírůstky) různého řádu

◇ *Absolutní diference*

$$\Delta^1 y_t = y_t - y_{t-1}, \quad t = 2, \dots, n$$

$$\Delta^2 y_t = \Delta^1 y_t - \Delta^1 y_{t-1} = y_t - 2y_{t-1} + y_{t-2}, \quad t = 3, 4, \dots, n$$

Průměrná absolutní diference

$$\bar{\Delta} = \frac{1}{n-1} \sum_{t=2}^n \Delta^1 y_t = \frac{y_n - y_1}{n-1}$$

◇ *Relativní diference*

$$\delta_t = \frac{\Delta^1 y_t}{y_{t-1}} = \frac{y_t - y_{t-1}}{y_{t-1}} = \frac{y_t}{y_{t-1}} - 1$$

○ Koeficient růstu (tempo růstu)

$$k_t = \frac{y_t}{y_{t-1}}, \quad t = 2, \dots, n$$

Průměrný koeficient růstu

$$\bar{k} = (k_2 k_3 \dots k_n)^{\frac{1}{n-1}} = \sqrt[n-1]{k_2 k_3 \dots k_n}$$

Modelování časových řad:

Jednorozměrný model

$$y_t = f(t, \epsilon_t), \quad t = 1, 2, \dots, n$$

t – časová proměnná

y_t – hodnota modelovaného ukazatele v čase t

ϵ_t – hodnota náhodné složky v čase t

● Přístupy k modelování časových řad:

a) Pomocí klasického modelu

Dekompozice (rozklad) č.ř.: na 4 složky časového pohybu (3 systematické, 1 náhodná)

- jádro spočívá v konstrukci systematické složky

- **Trend** T_t (hlavní tendence dlouhodobého vývoje hodnot ukazatele v čase)
- **Sezónnost** S_t (pravidelně se opakující odchylka od trendové složky - periodičita kratší nebo rovna 1 roku)
- **Cyklus** C_t (pravidelně se opakující kolísání okolo trendu v rámci několika let)
- **Náhodná složka** ϵ_t (výkyvy ukazatelů, které nemají systematický charakter)

◇ *Aditivní* model

$$y_t = T_t + S_t + C_t + \epsilon_t, \quad t = 1, 2, \dots, n$$

◇ *Multiplikativní* model

$$y_t = T_t \cdot S_t \cdot C_t \cdot \epsilon_t, \quad t = 1, 2, \dots, n$$

lze převést logaritmickou transformací na aditivní model

$$\ln y_t = \ln T_t + \ln S_t + \ln C_t + \ln \epsilon_t$$

b) Pomocí Boxovy-Jenkinsovy metodologie

Náhodná složka – základní prvek konstrukce modelu č.ř.

c) Pomocí spektrální analýzy

č.ř. – směs sinusovek a kosinusovek o rozličných amplitudách a frekvencích

Vícerozměrný model

$$y_t = f(t, x_1, x_2, \dots, x_p, \epsilon_t), \quad t = 1, 2, \dots, n$$

t – časová proměnná

y_t – hodnota modelovaného ukazatele v čase t

x_1, x_2, \dots, x_p – *faktorové (příčinné)* ukazatele ovlivňující analyzovaný ukazatel y_t

ϵ_t – hodnota náhodné složky v čase t

Popis trendu trendovými funkcemi

Model

$$y_t = T_t + \epsilon_t, \quad t = 1, 2, \dots, n$$

t – časová proměnná

y_t – hodnota modelovaného ukazatele v čase t

T_t – trendová funkce v čase t

ϵ_t – hodnota náhodné složky v čase t

V praxi nejvíce používané trendové funkce:

- **Jednoduché:**

- z hlediska průběhu a z hlediska odhadu parametrů

- **nemají asymptotu: jejich růst není omezen**

- **Lineární z hlediska parametrů:**

- Lineární, kvadratická (parabolická)*

- **Nelineární z hlediska parametrů:**

- exponenciální*

- linearizující transformací převeditelné na lineární

- **Složité:**

- z hlediska průběhu a z hlediska odhadu parametrů

- **lépe modelují ekonomickou realitu**

- **mají asymptotu: jejich růst je omezen**

- **Nelineární z hlediska parametrů**

- modifikovaná exponenciální, logistická, Gompertzova křivka*

- **Lineární trend (trendová přímka)**

$$T_t = \beta_0 + \beta_1 t, \quad t = 1, 2, \dots, n$$

t – časová proměnná

β_0, β_1 – neznámé parametry

- **Bodové odhady parametrů β_0, β_1 : b_0, b_1**

Kritérium: Minimalizace součtu čtverců:

$$S(\beta_0, \beta_1) = \sum_{t=1}^n [y_t - (\beta_0 + \beta_1 t)]^2$$

Soustava normálních rovnic

$$\begin{aligned} nb_0 + b_1 \sum_{t=1}^n t &= \sum_{t=1}^n y_t \\ b_0 \sum_{t=1}^n t + b_1 \sum_{t=1}^n t^2 &= \sum_{t=1}^n ty_t \end{aligned}$$

Řešení:

$$b_1 = \frac{\sum_{t=1}^n ty_t - n\bar{t}\bar{y}}{\sum_{t=1}^n t^2 - n\bar{t}^2}, \quad b_0 = \bar{y} - b_1 \bar{t}$$

Vlastnosti odhadů b_0, b_1 : – nezkreslené
– konsistentní

Rovnice odhadnuté trendové přímky:

$$\hat{T}_t = b_0 + b_1 t$$

- Praktický význam lineárního trendu:
 - slouží k orientačnímu určení základního směru vývoje analyzované č.ř.
 - v určitém omezeném časovém intervalu vhodná aproximace jiných trendových funkcí

• Kvadratický trend (parabolický)

$$T_t = \beta_0 + \beta_1 t + \beta_2 t^2, \quad t = 1, 2, \dots, n$$

t – časová proměnná

$\beta_0, \beta_1, \beta_2$ – neznámé parametry

Bodové odhady parametrů $\beta_i, i = 0, 1, 2$: $b_i, i = 0, 1, 2$

- Metoda nejmenších čtverců

$$\min_{\beta_0, \beta_1, \beta_2} \sum_{t=1}^n [y_t - (\beta_0 + \beta_1 t + \beta_2 t^2)]^2$$

- **Exponenciální trend**
(dvouparametrická exponenciála)

$$T_t = \beta_0 \beta_1^t, \quad \beta_1 > 0, \quad t = 1, 2, \dots, n$$

t – časová proměnná

β_0, β_1 – neznámé parametry

Bodové odhady parametrů $\beta_i, i = 0, 1$: b_i

- **Metoda nejmenších čtverců**

Funkce T_t – nelineární z hlediska parametrů \implies přímé použití MNČ vyloučeno

Linearizující logaritmická transformace:

$$\ln T_t = \ln \beta_0 + t \cdot \ln \beta_1$$

Položíme: $Z_t = \ln T_t, \quad \gamma_i = \ln \beta_i, \quad i = 0, 1$

Transformovaný trend: lineární

$$Z_t = \gamma_0 + \gamma_1 t$$

Metoda nejmenších čtverců:

$$\min_{\gamma_0, \gamma_1} \sum_{t=1}^n [\ln y_t - (\gamma_0 + \gamma_1 t)]^2$$

Odhady:

$$b_0 = e^{\hat{\gamma}_0}, \quad b_1 = e^{\hat{\gamma}_1}$$

Vlastnosti odhadů: – zkreslené (vychýlené)
– nejsou konsistentní

○ **Vážená metoda nejmenších čtverců**

$$\min_{\gamma_0, \gamma_1} \sum_{t=1}^n [\ln y_t - (\gamma_0 + \gamma_1 t)]^2 \cdot w_t$$

w_t – zvolené váhy (např. $w_t = y_t^2$)

Vlastnosti odhadů: – asymptoticky nezkreslené
– konzistentní
(zlepšení oproti MNČ)

- **Modifikovaný (posunutý) exponenciální trend (tříparametrická exponenciála)**

$$T_t = \kappa + \beta_0 \beta_1^t, \quad \beta_1 > 0, \quad t = 1, 2, \dots, n$$

t – časová proměnná

β_0, β_1 – neznámé parametry

κ – aditivní parametr - *parametr posunutí*

Trend T_t – nelineární z hlediska parametrů

Vzhledem k aditivnímu parametru κ nelze jednoduchou linearizující transformací převést na lineární funkci.

Bodové odhady parametrů β_0, β_1, κ : $b_0, b_1, \hat{\kappa}$

- **Metoda částečných součtů (MČS)**

Zvolíme počáteční t a 3 na sebe navazující disjunktní součty o délce m

Předpoklad: $n = 3m$

Je-li $n \neq 3m$ – vynecháme potřebný počet pozorování na začátku č.ř.

Částečné součty empirické (výběrové)

$$\hat{S}_1 = \sum_{t=1}^m y_t$$

$$\hat{S}_2 = \sum_{t=m+1}^{2m} y_t$$

$$\hat{S}_3 = \sum_{t=2m+1}^{3m} y_t$$

Částečné součty teoretické

$$S_1 = \sum_{t=1}^m T_t = \sum_{t=1}^m (\kappa + \beta_0 \beta_1^t) = m\kappa + \beta_0 \sum_{t=1}^m \beta_1^t$$

$$S_2 = \sum_{t=m+1}^{2m} T_t = \sum_{t=m+1}^{2m} (\kappa + \beta_0 \beta_1^t) = m\kappa + \beta_0 \sum_{t=m+1}^{2m} \beta_1^t$$

$$S_3 = \sum_{t=2m+1}^{3m} T_t = \sum_{t=2m+1}^{3m} (\kappa + \beta_0 \beta_1^t) = m\kappa + \beta_0 \sum_{t=2m+1}^{3m} \beta_1^t$$

Princip metody částečných součtů:

Položíme

$$\hat{S}_i = S_i, \quad i = 1, 2, 3$$

Platí: podle věty o součtu geometrické řady

$$(\sum_{i=1}^k q^i = q \frac{q^k - 1}{q - 1}, \quad q \neq 1)$$

$$\sum_{t=2m+1}^{3m} \beta_1^t = \beta_1^{2m+1} \frac{\beta_1^m - 1}{\beta_1 - 1}$$

Soustava 3 rovnic pro 3 neznámé:

$$\hat{S}_1 = m\hat{\kappa} + b_0 b_1 \frac{b_1^m - 1}{b_1 - 1}, \quad \hat{S}_2 = m\hat{\kappa} + b_0 b_1^{m+1} \frac{b_1^m - 1}{b_1 - 1}$$

$$\hat{S}_3 = m\hat{\kappa} + b_0 b_1^{2m+1} \frac{b_1^m - 1}{b_1 - 1}$$

Řešení:

$$b_1 = \left(\frac{\hat{S}_3 - \hat{S}_2}{\hat{S}_2 - \hat{S}_1} \right)^{\frac{1}{m}}, \quad b_0 = (\hat{S}_2 - \hat{S}_1) \frac{b_1 - 1}{b_1 (b_1^m - 1)^2}$$

$$\hat{\kappa} = \frac{1}{m} \left(\hat{S}_1 - b_0 b_1 \frac{b_1^m - 1}{b_1 - 1} \right)$$

○ Metoda vybranných bodů

Určení počátečních odhadů parametrů:

Zvolíme 3 hodnoty časové proměnné:

$$t, t + p, t + 2p$$

Soustava rovnic pro odhad parametrů:

$$y_t = \hat{\kappa} + b_0 b_1^t$$

$$y_{t+p} = \hat{\kappa} + b_0 b_1^{t+p}$$

$$y_{t+2p} = \hat{\kappa} + b_0 b_1^{t+2p}$$

Řešení:

$$b_1 = \left(\frac{y_{t+2p} - y_{t+p}}{y_{t+p} - y_t} \right)^{\frac{1}{p}}, \quad b_0 = \frac{y_{t+p} - y_t}{b_1^t (b_1^p - 1)}$$

$$\kappa = y_t - b_0 b_1^t$$

• Logistický trend

$$T_t = \frac{\kappa}{1 + \beta_0 \beta_1^t}, \quad \beta_0 > 0, \quad \beta_1 > 0, \quad \kappa > 0, \quad t = 1, 2, \dots, n$$

t – časová proměnná

β_0, β_1, κ – neznámé parametry

Logistická funkce – trendová funkce

- průběh ve tvaru písmene "S" (S-křivka)
- 1 inflexní bod

Použití v ekonomické oblasti:

- modely poptávky po předmětech dlouhodobé spotřeby
- modelování vývoje, výroby a prodeje některých druhů výrobků

Bodové odhady parametrů $\beta_i, i = 0, 1, \kappa$:

- **Metoda částečných součtů (MČS):**

Transformace logistického trendu:

$$\frac{1}{T_t} = \frac{1}{\frac{\kappa}{1+\beta_0\beta_1^t}} = \frac{1}{\kappa} + \frac{\beta_0}{\kappa} \cdot \beta_1^t$$

Položíme:

$$Z_t = \frac{1}{T_t}, \quad \xi = \frac{1}{\kappa}, \quad \beta'_0 = \frac{\beta_0}{\kappa}$$

Transformovaný trend: modifikovaný exponenciální

$$Z_t = \xi + \beta'_0\beta_1^t$$

ξ, β'_0, β_1 – parametry modif. exponenciálního modelu

Aplikace MČS na časovou řadu hodnot $1/y_t$:

Předpoklad:

$$\text{sign}(\hat{S}_3 - \hat{S}_2) = \text{sign}(\hat{S}_2 - \hat{S}_1) (\neq 0)$$

- **Metoda vybranných bodů:**

Postup analogický jako u modifikovaného exponenciálního trendu.

○ Metoda diferenčních odhadů:

$$\frac{dy_t}{dt} = \frac{\kappa}{(1 + \beta_0 \beta_1^t)^2} \beta_0 \beta_1^t \ln \beta_1 = -\frac{\ln \beta_1}{\kappa} y_t (\kappa - y_t)$$

Platí:

$$\frac{dy_t}{dt} \approx \frac{y_{t+1} - y_t}{(t+1) - t} = y_{t+1} - y_t = \Delta_t^1$$

Δ_t^1 – časová řada prvních diferencí

$$\frac{\Delta_t}{y_t} = -\ln \beta_1 + \frac{\ln \beta_1}{\kappa} y_t$$

Vztah proměnných $\frac{\Delta_t^1}{y_t}$, y_t :

– regresní funkce lineární v parametrech

– parametry: $-\ln \beta_1$, $\frac{\ln \beta_1}{\kappa}$

◇ Bodové odhady parametrů $\ln \beta_1$, κ :

Metoda nejmenších čtverců:

◇ Bodový odhad parametru β_0 :

Rhodesův vzorec:

$$y_t \approx \frac{\hat{\kappa}}{1 + b_0 b_1^t}, \quad b_0 b_1^t \approx \frac{\hat{\kappa}}{y_t - 1} \quad | \ln$$

$$\ln b_0 + t \ln b_1 = \ln \left(\frac{\hat{\kappa}}{y_t - 1} \right) \quad | \frac{1}{n} \sum_{t=1}^n$$

$$\ln b_0 = -\frac{n+1}{2} \ln b_1 + \frac{1}{n} \sum_{t=1}^n \ln \left(\frac{\hat{\kappa}}{y_t - 1} \right)$$

• Gompertzova křivka

$$T_t = \kappa \beta_0^{\beta_1^t}, \quad \beta_1 > 0, \quad \kappa > 0, \quad t = 1, 2, \dots, n$$

t – časová proměnná

β_0, β_1 – neznámé parametry

Křivka tvaru "S", asymetrická

Bodové odhady parametrů β_0, β_1, κ :

○ Metoda částečných součtů:

Linearizující logaritmická transformace:

$$\ln T_t = \ln \kappa + \beta_1^t \ln \beta_0$$

Položíme:

$$Z_t = \ln T_t, \quad \xi = \ln \kappa, \quad \beta'_0 = \ln \beta_0$$

Transformovaný trend: modifikovaný exponenciální

$$Z_t = \xi + \beta'_0 \beta_1^t$$

ξ, β'_0, β_1 – parametry modif. exponenciálního modelu

Aplikace MČS na časovou řadu hodnot $\ln y_t$:

Volba vhodného modelu trendu

Kriterium:

a) Věcně ekonomické kritérium

– základní tendence vývoje pouze v hrubých rysech: trendová fce rostoucí, klesající, inflexní bod, omezená, neomezená

b) Analýza grafu časové řady

– subjektivní hledisko; tvar grafu závisí na volbě měřítka

c) Analýza jednoduchých charakteristik č.ř.

Trend	Charakteristiky
Lineární	$\Delta_t^1 = y_t - y_{t-1} \approx \text{konst.}$
Kvadratický	$\Delta_t^2 = y_t - 2y_{t-1} + y_{t-2} \approx \text{konst.}$
Exponenciální	$k_t = \frac{y_t}{y_{t-1}} \approx \text{konst.}$
Logistický	$\frac{1/y_{t+2} - 1/y_{t+1}}{1/y_{t+1} - 1/y_t} \approx \text{konst.}$
Gompertzova k.	$\frac{\ln y_{t+2} - \ln y_{t+1}}{\ln y_{t+1} - \ln y_t} \approx \text{konst.}$

d) Interpolační kritéria

y_t – pozorovaná hodnota v časové řadě v čase t

\hat{T}_t – odhadnutá hodnota trendu v čase t

Míry "úspěšnosti" (přesnosti) zvolené trendové funkce

- ***Střední chyba odhadu*** (Mean Error):

$$ME = \frac{1}{n} \sum_{t=1}^n (y_t - \hat{T}_t)$$

- ◇ $ME = 0 \iff$

k odhadu parametrů použijeme klasickým způsobem MNČ (přímka, parabola, hyperbola a jiné trendové modely lineární v parametrech, konstruované včetně konstantního členu β_0)

- ◇ $ME \neq 0$

úprava postupu např. logaritmizací či inverzí hodnot č.ř. nebo k odhadu použita jiná metoda než MNČ

- ***Střední kvadratická chyba odhadu*** (Mean Squared Error):

$$MSE = \frac{1}{n} \sum_{t=1}^n (y_t - \hat{T}_t)^2$$

- ***Střední absolutní chyba odhadu*** (Mean Absolute Error):

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{T}_t|$$

- *Střední procentní chyba odhadu*
(Mean Percentage Error):

$$MPE = \frac{100}{n} \sum_{t=1}^n \left(\frac{y_t - \hat{T}_t}{y_t} \right)$$

- *Střední absolutní procentní chyba odhadu*
(Mean Absolute Percentage Error):

$$MAPE = \frac{100}{n} \sum_{t=1}^n \left(\frac{|y_t - \hat{T}_t|}{y_t} \right)$$

Nejvíce používaná míra: MSE

- zvýhodňuje víceparametrické trendové funkce

Korigovaná (upravená) MSE:

$$MSE = \frac{1}{(n - p)} \sum_{t=1}^n (y_t - \hat{T}_t)^2$$

p – počet parametrů

e) F-test o modelu

Porovnání různých modelů trendu:

Celkový F-test:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$$

H_1 : alespoň jeden parametr $\beta_j \neq 0$

◇ Hladina významnosti: α

◇ Testovací statistika:

$$F = \frac{\frac{S_T}{(p-1)}}{\frac{S_R}{(n-p)}} \sim F[p-1; n-2]$$

$$S_T = \sum_{t=1}^n (\hat{T}_t - \bar{y})^2, \quad S_R = \sum_{t=1}^n (y_t - \hat{T}_t)^2$$

◇ Kritický obor:

$$F > F_{1-\alpha}(p-1; n-2)$$

$F_{1-\alpha}(p-1; n-2) - 100(1-\alpha)\%$ kvantil Fisherova-Snedecorova rozdělení s $(p-1)$ a $(n-2)$ stupni volnosti

Klouzavé průměry (KP)

Analytické metody vyrovnávání č.ř. - vyrovnáme jednou trendovou funkcí všechna pozorování č.ř., která máme k dispozici.

Přístup k vyrovnání č.ř. pomocí klouzavých průměrů - rozsah období, v jehož rámci bude č.ř. vyrovnána, podstatně kratší, než období, za které máme č.ř. shromážděnou.

- Podstata vyrovnání pomocí KP:

- posloupnost pozorování y_t nahradíme posloupností průměrů vypočítaných z pozorování y_t - řadu vyrovnáváme po částech;
- každý průměr reprezentuje určitou skupinu pozorování.

- Důležité otázky:

- Stanovení počtu pozorování nutných k výpočtu KPů:
 - *klouzavá část* období vyrovnání - časový interval určité délky, který se posunuje po časové ose vždy o jednotku

$m = 2p + 1$, $m < n$, p délka klouzavé části,
 n celkový počet pozorování č.ř.

- Volba délky p klouzavé části - nelze stanovit exaktními statistickými metodami

- ◇ heuristický postup - na základě věcné analýzy zkoumaného ekonomického jevu

v praxi obvykle menší délky:

$$p = 2, 3, 4 \quad \text{t.j.} \quad m = 5, 7, 9$$

- Identifikace jednotlivých klouzavých částí - reprezentace použitím jejich středních bodů:

Předpoklad:

$$m \text{ liché} \implies p = (m - 1)/2 \quad \text{sudé} \implies$$

střední bod 1. klouzavé části je $(p + 1)$ ní bod;

střední bod poslední klouzavé části: bod s pořadovým číslem $(n - p)$.

Typy klouzavých průměrů

- **Prosté klouzavé průměry m -členné**
(m liché číslo)

Předpoklad: lineární trend definován na klouzavých částech o rozsahu

$$m = 2p + 1, \quad p = 1, 2, \dots,$$

Časová proměnná v celé č.ř.:

$$t = 1, 2, \dots, n$$

Střední body jednotlivých klouzavých částí:

$$t = p + 1, p + 2, \dots, n - p$$

Nová časová proměnná:

$$i = -(p - j), \quad j = 0, 1, \dots, 2p$$

- Lineární trend použitý k vyrovnaní jednotlivých klouzavých částí:

$$T_{t,i} = \beta_{0t} + \beta_{1t} i, \quad t = p + 1, p + 2, \dots, n - p$$

- *Prostý klouzavý průměr:*

$$\bar{y}_t = \frac{1}{m} \sum_{i=-p}^p y_{t,i} = \frac{y_{t-p} + y_{t-p+1} + \dots + y_{t+p}}{m}$$

$$t = p + 1, p + 2, \dots, n - p$$

Označení: $\frac{1}{m} \underbrace{[1, 1, \dots, 1]}_m$

- **Vážené klouzavé průměry**

Pro každý klouzavý úsek č.ř. počítáme vážený aritmetický průměr

$$\bar{y}_t = \sum_{i=-p}^p w_i y_{t+i}, \quad t = p + 1, p + 2, \dots, n - p$$

w_i - váhy splňující podmínku

$$\sum_{i=-p}^p w_i = 1$$

Váhy konstruovány různým způsobem:

Příklad: **Parabolický trend** definován na klouzavých částech o rozsahu $m = 2p + 1$ použitý k vyrovnání jednotlivých klouzavých částí:

$$T_{t,i} = \beta_{0t} + \beta_{1t} i + \beta_{2t} i^2, \quad t = p + 1, p + 2, \dots, n - p$$

- **Vážený klouzavý průměr:**

$$\bar{y}_t = \sum_{i=-p}^p W_i y_{t,i}, \quad t = p + 1, p + 2, \dots, n - p$$

$$W_i = \frac{3}{4m(m^2 - 1)}(3m^2 - 1 - 20i^2)$$

Váhy W_i splňují podmínky:

$$\sum_{i=-p}^p W_i = 1, \quad W_i = W_{-i} \quad \text{symetrické}$$

- **Centrované klouzavé průměry m -členné**
(m sudé číslo)

Předpoklad: m sudé \implies

střední body klouzavých částí nejsou celá čísla
- nelze přiřadit hodnoty klouzavých průměrů
přímo pozorováním dané č.ř.

Centrování prostých klouzavých průměrů –
speciální typ vážených klouzavých průměrů

- *Centrovaný klouzavý průměr:*

$$\bar{y}_t = \frac{1}{4p} (y_{t-p} + 2y_{t-p+1} + \dots + 2y_{t+p-1} + y_{t+p})$$

$$t = p + 1, p + 2, \dots, n - p$$

Označení: $\frac{1}{2m} \underbrace{[1, 2, \dots, 2, 1]}_{m+1}$

Identifikace a popis sezónní složky

Pomocí sezónní složky jsou v č.ř. reprezentovány sezónní vlivy.

Sezónní vlivy - soubor přímých či nepřímých příčin, které se rok co rok pravidelně opakují v důsledku existence pravidelného koloběhu Země kolem Slunce:

- **Vlivy klimatické** (zvýšená spotřeba a výroba nápojů v letních měsících opakující se po 12 měsících).
- **Vlivy zprostředkované** (společenské standardy a zvyklosti ve stereotypch chování lidí - školní prázdniny, dovolené, víkendy, Vánoce atd. a s tím související ekonomické, dopravní, kulturní a jiné důsledky).

Sezónní výkyvy - výsledky působení sezónních vlivů na analyzovanou č.ř. - pravidelné výkyvy zkoumané č.ř. nahoru a dolů vůči určitému ne-sezónnímu vývoji řady v průběhu let.

Základní úkoly:

- **Identifikace sezónní složky** (zda sezónní výkyvy jsou statisticky významné):
 - **Kvantifikace sezónní složky** (periodické kolísání zakrývá dynamiku ekonomických jevů)
 - ◇ **analýza sezónnosti**
 - **Výpočet sezónně očištěné č.ř.**

Popis sezónní složky

Pozorování v č.ř. se vztahují:

- k určité i -té periodě (např. rok)
- k j -tému dílčímu období v rámci určité periody (např. měsíc v roce)

Aditivní model:

$$y_{ij} = T_{ij} + S_{ij} + \epsilon_{ij}, \quad i = 1, 2, \dots, r, \quad j = 1, 2, \dots, s$$

s - počet dílčích období v rámci roku

Pro hodnoty časové proměnné

$$t = 1, 2, \dots, n, \quad n = r \cdot s$$

platí

$$t = s(i - 1) + j$$

Předpoklad v praktických aplikacích: s - sudé číslo
(4 čtvrtletí, 12 měsíců)

• Model konstantní sezónnosti

Předpoklad: sezónní kolísání je neměnné pro j -tou sezónu v letech i , $j = 1, 2, \dots, s$, $i = 1, 2, \dots, r$

$$S_{ij} = S_j, \quad i = 1, 2, \dots, r, \quad j = 1, 2, \dots, s$$

- **Normalizační podmínka** (za účelem vykompenzování vlivu sezónních faktorů v rámci periody):

$$\sum_{j=1}^s S_j = 0, \quad \forall i = 1, 2, \dots, r \quad (Nk)$$

- **Odhad sezónních parametrů**

Metoda empirických sezónních odchylek

- ◇ Odhadneme (vyrovnáme) trendové hodnoty v jednotlivých obdobích pomocí klouzavých průměrů délky s : \hat{T}_{ij}

- ◇ Určíme rozdíly pozorovaných hodnot č.ř. y_{ij} a hodnot vyrovnaných klouzavými průměry \hat{T}_{ij} :

$$y_{ij} - \hat{T}_{ij} - \text{empirické sezónní odchylky}$$

- ◇ Určíme *průměrnou empirickou sezónní odchylku*:

$$\hat{S}_j = \frac{1}{r} \sum_{i=1}^r (y_{ij} - \hat{T}_{ij})$$

- ◇ Transformace průměrných empirických sezónních odchylek:

$$\hat{S}'_j = \hat{S}_j - \frac{1}{s} \sum_{j=1}^s \hat{S}_j, \quad \sum_{j=1}^s \hat{S}'_j = 0$$

• Model proporcionální sezónnosti

Předpoklad, že pro danou sezónu $j = 1, 2, \dots, s$ se sezónní výkyvy pravidelně opakují ve stejné výši v letech $i = 1, 2, \dots, r$ není v některých situacích realistický.

Předpoklad: v dílčím období $j = 1, 2, \dots, s$ se sezónní výkyvy mění přímo úměrně dosažené úrovni trendové složky

$$S_{ij} = \gamma_j T_{ij}, \quad i = 1, 2, \dots, r, \quad j = 1, 2, \dots, s$$

$$y_{ij} = T_{ij} + \gamma_j T_{ij} + \epsilon_{ij} = (1 + \gamma_j) T_{ij} + \epsilon_{ij}$$

$$i = 1, 2, \dots, r, \quad j = 1, 2, \dots, s$$

$\gamma_j, \quad j = 1, 2, \dots, s$ – *sezónní parametry*

$(1 + \gamma_j)$ – *sezónní indexy*

Sezónní indexy – udávají relativní výkyvy od trendové hodnoty v důsledku sezónního kolísání (bezrozměrná čísla)

- If v j -té sezóně $\gamma_j > 0$ – *sezónní vzestup*
- If v j -té sezóně $\gamma_j < 0$ – *sezónní pokles*
- If v j -té sezóně $\gamma_j = 0$ – *nepůsobí sezónní vlivy*

- Normalizační podmínka :

$$\sum_{j=1}^s (1 + \gamma_j) = s, \quad i = 1, 2, \dots, r \quad (Np)$$

- Odhad $(1 + \hat{\gamma}_j)$ sezónních indexů ($\hat{\gamma}_j$ -odhad γ_j)

Metoda empirických sezónních indexů

- ◇ Odhad trendových hodnot v jednotlivých obdobích $t_{ij} = s(i-1) + j$ pomocí klouzavých průměrů délky s : \hat{T}_{ij}
- ◇ Určíme podíly pozorovaných hodnot č.ř. y_{ij} a hodnot vyrovnaných klouzavými průměry \hat{T}_{ij} :

$$\frac{y_{ij}}{\hat{T}_{ij}} - \text{empirické sezónní indexy}$$

- ◇ Určíme *průměrný empirický sezónní index* (odhad sezónního indexu):

$$(1 + \hat{\gamma}_j) = \frac{1}{r} \sum_{i=1}^r \frac{y_{ij}}{\hat{T}_{ij}}$$

- ◇ **Standardizace průměrných empirických indexů – sezónní faktory**
(průměrný empirický sezónní index přepočteme s ohledem na podmínku (Np)):

$$(1 + \hat{\gamma}'_j) = (1 + \hat{\gamma}_j) / \frac{1}{s} \sum_{i=1}^s (1 + \hat{\gamma}_j)$$

Metoda nejmenších čtverců:

$$(1 + \hat{\gamma}_j) = \frac{\sum_{i=1}^r y_{ij} \hat{T}_{ij}}{\sum_{i=1}^r \hat{T}_{ij}^2}$$

○ Test hypotézy o existenci sezónnosti

Test nulové hypotézy

$$H_0 : S_j = 0, j = 1, 2, \dots, s \quad \text{versus} \quad H_1 : \text{non}H_0$$

◇ **Testovací statistika:**

$$F = \frac{\frac{r}{s-1} \sum_{j=1}^s (\bar{y}_{\cdot j} - \bar{y})^2}{\frac{S_R}{(r-1)(s-1)}} \sim F[s-1, (r-1)(s-1)]$$

$$S_R = \sum_{i=1}^r \sum_{j=1}^s (y_{ij} - \bar{y})^2 - s \sum_{i=1}^r (\bar{y}_{i\cdot} - \bar{y})^2 - r \sum_{j=1}^s (\bar{y}_{\cdot j} - \bar{y})^2$$

$$\bar{y}_{\cdot j} = \frac{1}{r} \sum_{i=1}^r y_{ij}$$

$$\bar{y} = \frac{1}{r \cdot s} \sum_{i=1}^r \sum_{j=1}^s y_{ij}$$

$F[\nu_1; \nu_2]$ – Fisherovo-Snedecorovo rozdělení o ν_1 a ν_2 stupni volnosti

– **Kritický obor:**

$$F > F_{1-\alpha}[s-1, (r-1)(s-1)]$$

$F_{1-\alpha}[s-1, (r-1)(s-1)]$ – $100(1-\alpha)\%$ kvantil $F[s-1, (r-1)(s-1)]$

• Sezónní očišťování

Proč sezónní očišťování?

Možnost průběžně porovnávat po sobě jdoucí údaje v č. ř. uvnitř roku i tehdy, pokud jsou aktuálně ovlivněny sezónností.

Modelové rozdělení č.ř. na složku trendovou, sezónní a náhodnou – úkolem je zbavit č.ř. sezónní složky, ponechat složku trendovou (příp. cyklickou, pokud se vyskytuje).

Praktické metody očišťování používají:

- různé typy a varianty klouzavých průměrů;
- regresní metody;
- např. exponenciální vyrovnávání;
- poměrně komplikované počítačové programy.

Popis náhodné složky v časových řadách

Náhodná složka č.ř. – výsledek působení blíže nespecifikovaného souboru náhodných vlivů:

$$\epsilon_t = y_t - \hat{Y}_t = y_t - \hat{T}_t - \hat{S}_t - \hat{C}_t$$

Předpoklady:

- Střední hodnoty nulové

$$E(\epsilon_t) = 0, \quad t = 1, 2, \dots, n$$

- *Homoskedastické* náhodné veličiny

- ◊ Konstantní rozptyl v čase

$$D(\epsilon_t) = \sigma^2, \quad t = 1, 2, \dots, n$$

- ◊ Vzájemně nekorelované náhodné veličiny

$$\text{cov}(\epsilon_t, \epsilon_k) = E(\epsilon_t \epsilon_k) = 0, \quad t, k = 1, 2, \dots, n, \quad t \neq k$$

- *Heteroskedastické* náhodné poruchy

- ◊ Proměnlivé rozptyly typu

$$D(\epsilon_t) = s^2 w_t^{-1}, \quad t = 1, 2, \dots, n$$

w_t – váhy pozorování

$$\sum_{t=1}^n w_t^{-1} = 1$$

- ◊ Vzájemně nekorelované náhodné veličiny

$$\text{cov}(\epsilon_t, \epsilon_k) = E(\epsilon_t \epsilon_k) = 0, \quad t, k = 1, 2, \dots, n, \quad t \neq k$$

○ **Autoregrese náhodných poruch**

$$\epsilon_t = \rho\epsilon_{t-1} + u_t, \quad 0 < \rho < 1$$

ρ – autokorelační koeficient sousedních
(v čase) náhodných poruch - konstantní

$u_t, t = 1, 2, \dots, n$ – posloupnost náhodných
veličin vzájemně nezávislých

$$E(u_t) = 0, \quad D(u_t) = \text{konst.}$$

Reziduum e_t

$$e_t = y_t - \hat{Y}_t$$

\hat{Y}_t – odhad teoretické hodnoty č.ř. v čase t

Reziduum e_t – odhad náhodné veličiny ϵ_t
konkrétní hodnota pro daný model a daná data náhodné
veličiny ϵ_t .

Ověřování předpokladů o náhodné složce

K ověřování lze použít některé z následujících testů.

• Testování náhodného uspořádání reziduí

◦ Znaménkový test

- ◊ Určíme posloupnost: $\{e_t\}$, $t = 1, 2, \dots, n$
- ◊ Vypočteme difference: $e_t - e_{t-1}$
- ◊ Určíme: počet S_+ kladných diferencí
počet S_- záporných diferencí
Nulové difference z posloupnosti odstraníme.
- ◊ Určíme statistiku

$$S = \max(S_+, S_-)$$

- ◊ Posloupnost $\{e_t\}$ považujeme za náhodně uspořádanou, pokud se S málo liší od $E(S)$

$$E(S) = \frac{n-1}{2}$$

- ◊ Test nulové hypotézy

$$H_0 : S = \frac{n-1}{2} \quad \text{versus} \quad H_1 : S \neq \frac{n-1}{2}$$

– Testovací statistika:

$$U = \frac{S + \frac{1}{2} - \frac{n-1}{2}}{\sqrt{\frac{n-1}{4}}} = \frac{2S - n + 1}{\sqrt{n-1}} \approx N(0, 1)$$

– Kritický obor:

$$\left| \frac{2S - n + 1}{\sqrt{n-1}} \right| > u_{1-\alpha/2}$$

$u_{1-\alpha/2}$ – 100(1 - $\alpha/2$)% kvantil $N(0, 1)$

○ Test založený na bodech zvratu (bodech obratu)

- ◇ *Body zvratu (obratu)* – rezidua, která jsou větší nebo menší než obě sousední rezidua.
 - *Vrchol (dolík)* – reziduum s hodnotou vyšší (nižší) než jsou dvě sousední rezidua.
- ◇ P – počet úseků posloupnosti reziduí mezi body zvratu
- ◇ $(P - 1)$ – počet bodů zvratu
(Případná shodná rezidua z posl. odstraníme)
- ◇ **Je-li zkoumaná posloupnost $\{e_t\}$ náhodně uspořádaná, potom statistika P má střední hodnotu a rozptyl**

$$E(P) = \frac{2(n-2)}{3}, \quad D(P) = \frac{16n-29}{90}$$

- ◇ **Test nulové hypotézy**

$$H_0 : P = \frac{2(n-2)}{3} \quad \text{versus} \quad H_1 : P \neq \frac{2(n-2)}{3}$$

- **Testovací statistika:**

$$U = \frac{\sqrt{90} \left[P - \frac{2(n-2)}{3} \right]}{\sqrt{16n-29}} \approx N(0, 1)$$

- **Kritický obor:**

$$\left| \frac{2S - n + 1}{\sqrt{n-1}} \right| > u_{1-\alpha/2}$$

$u_{1-\alpha/2} = 100(1 - \alpha/2)\%$ kvantil $N(0, 1)$

- Testování autokorelace náhodných poruch

- *Durbinův-Watsonův test autokorelace*

Závislost náhod. poruch vyjádřena autoregresním modelem

$$\epsilon_t = \rho\epsilon_{t-1} + u_t$$

- ◊ Test nulové hypotézy

$$H_0 : \rho = 0 \quad \text{versus} \quad H_1 : \rho \neq 0$$

– Testovací statistika (Durbin-Watsonova):

$$DW = \frac{\sum_{i=2}^n (e_t - e_{t-1})^2}{\sum_{i=1}^n e_t^2} \in \langle 0, 4 \rangle$$

– **Kritické hodnoty:** d_L, d_U tabelovány

(speciální tabulky Durbin-Watsonova testu)

– pro danou hladinu významnosti

– pro daný počet n pozorování č.ř.

– pro daný počet strukturálních parametrů modelu (na př. u trendové funkce je počet parametrů roven 1)

H_0 zamítneme if

$$DW < d_L \text{ nebo } DW > 4 - d_L$$

H_0 nezamítneme if

$$d_U < DW < 2 \text{ nebo } 4 - d_U < DW < 4 - d_L$$

nerozhodneme if

$$d_L < DW < d_U \text{ nebo } 4 - d_U < DW < 4 - d_L$$

◇ **Test H_0 v praktických situacích:**

– **přibližné vyhodnocení**

H_0 zamítneme if

$$DW \text{ blízke } 0 \text{ nebo } DW \text{ blízke } 4$$

H_0 nezamítneme if

$$DW \doteq 2$$

Adaptivní přístupy k modelu časových řad

Adaptivní modely

(modely s měnlivými parametry)

Vlastnosti:

- společné s klasickými modely s konst. param.
 - ◇ neobjasňují kauzální mechanismus vývoje analyzované proměnné
pouze popisují její průběh v čase
- rozdílné od klasických modelů
 - ◇ nepředpokládají stabilitu
 - analytického tvaru v čase
 - strukturálních parametrů v čase
 - spojitost trendové funkce
 - ◇ rychle reagují na
 - strukturální změny v čase
 - vhodné pro předpovědi průběhu č.ř.
s nepravidelnostmi a zlomy v trendu

Předpoklad:

- nejcennější jsou nejnovější pozorování č.ř. \implies
 - největší váhy přiřazeny nejnovějším pozorováním
 - dřívější pozorování - nulové nebo malé váhy

Třída adaptivních modelů – poměrně široká

● Exponenciální vyrovnávání

Předpoklady:

- y_n – pozorování v přítomném čase
- v časové okamžiku n máme k dispozici č.ř. pozorovaných hodnot:
 $y_{n-k}, k = 0, 1, \dots, n - 1$
- k – stáří (věk) pozorování
- aditivní model

$$y_{n-k} = T_{n-k} + \epsilon_{n-k}$$

$$T_{n-k} = \beta_0 - \beta_1 k + \beta_2 k^2 + \dots + (-1)^k \beta_k k^k$$

Odhady parametrů:

- Metoda nejmenších čtverců

$$\min \left(\sum_{k=0}^{n-1} (y_{n-k} - \hat{T}_{n-k})^2 \right)$$

Stejná váha přiřazena všem pozorováním bez ohledu na "stáří".

- Vážená metoda nejmenších čtverců

$$\min \left(\sum_{k=0}^{n-1} (y_{n-k} - \hat{T}_{n-k})^2 w_k \right)$$

w_k - váhy nepřímo úměrné "stáří" pozorování

$$w_k = \alpha^k, \quad 0 < \alpha < 1, \quad k = 0, 1, \dots, n - 1$$

α - vyrovnávací konstanta

- Volba trendu T_{n-k}

- *Jednoduché exponenciální vyrovnávání*

Trend lze považovat v krátkých úsecích č.ř. za konstantní tj. $k = 0$.

Soustava normálních rovnic (1 rovnice)

$$b_0 \sum_{k=0}^{n-1} \alpha^k = \sum_{k=0}^{n-1} \alpha^k y_{n-k}$$

$\{\alpha^k\}$ - geometrická posloupnost

$$\sum_{k=0}^{n-1} \alpha^k \approx \sum_{k=0}^{\infty} \alpha^k = \frac{1}{1 - \alpha}$$

$$b_0 = y_{n-k} = (1 - \alpha) \sum_{k=0}^{n-1} \alpha^k y_{n-k}$$

Rekurentní výpočet vyrovnaných hodnot č.ř.

$$y_{n-k} = (1 - \alpha)y_{n-k} + \alpha y_{n-k-1}$$

- *Dvojitě exponenciální vyrovnávání*

Trend lze považovat v krátkých úsecích č.ř. za lineární.

- *Trojitě exponenciální vyrovnávání*

Trend lze považovat v krátkých úsecích č.ř. za kvadratický.

- **Konstrukce předpovědi na základě exponenciálního vyrovnávání**

Předpověď provedenou v čase n o jedno období dopředu ztotožníme s vyrovnanou hodnotou y_n

$$P_n(1) = y_n$$

Adaptace předpovědi pomocí rekurentního vztahu:

$$P_n(1) = P_{n-1}(1) + (1 - \alpha)[y_n - P_{n-1}(1)]$$

$e_t = y_n - P_{n-1}(1)$ – chyba předpovědi konstruovaná v čase $n - 1$

Konstrukce předpovědí časových řad

Statistické prognostické metody:

- Metody extrapolace časových řad.
- Metody modifikující různé metody regresní analýzy.
- Metody národního účetnictví.

V ekonomické a statistické praxi - metody extrapolace č.ř.

Podstata klasických extrapoláčnických metod:

- Studuje historie prognózovaného jevu.
- Zákonitosti vývoje v minulosti a přítomnosti se přenesou do budoucnosti.

Metody založené na extrapolaci klasických modelů trendu vycházejí z deterministického principu: budoucnost vyplývá z přítomnosti.

Předpoklad:

Tendence vývoje zkoumaného jevu jsou neměnné nebo alespoň relativně stabilní.

Přednosti extrapolace klasických modelů:

- Relativně jednoduchý matematicko-statistický aparát - snadná algoritmizace použitých metod
- Prognozovaná veličina je závisle proměnnou, čas je nezávisle proměnná \implies k analýze a prognóze stačí informace o vývoji analyzovaného jevu v minulosti.
- Konstrukce předpovědi je poměrně rychlá a jednoduchá.
- Není nutné uskutečňovat prognózy dalších jevů, které vysvětlují prognózovaný jev.

Nedostatky extrapolace klasických modelů:

- ◇ Neposkytují systémové prognózy - každý jev se posuzuje individuálně.
- ◇ Kvalita analýzy a prognózy ovlivněna zvoleným typem modelu.
 - ▷ Výběr modelu se provádí empiricky.
 - ▷ Hodnoty statistických kritérií jsou jen nutnou, ale ne postačující podmínkou správnosti modelu a jeho vhodnosti.

Vhodnost modelu posuzovat také na základě

- určitých ekonomických předpokladů o dynamice prognózovaných jevů;
- současně s kvantitativní analýzou respektovat výsledky kvalitativní analýzy;
- intuice.

Získaná předpověď by neměla být izolovaně základem pro rozhodování ale porovnána s předpověďmi získanými jinými prognostickými metodami

Metody extrapolace časových řad - význam především při konstrukci prognóz krátkodobých.

Korelace časových řad

- Zkoumání vztahu časových řad k jiným č.ř.

Předpoklad:

Každou č.ř. můžeme popsat určitým aditivním modelem

Zkoumání, zda mezi č.ř. existuje určitý vztah zahrnuje:

- sledování celkové vývojové tendence nebo sezónního kolísání;
- zkoumání, zda neexistuje vztah mezi náhodnými složkami analyzovaných č.ř.

Použití metod měření těsnosti závislosti řad náhodné složky.

Předpoklad: Časová ř. aditivního typu:

$$y_t = T_t + \epsilon_t, \quad t = 1, 2, \dots, n$$

Zkoumáme č.ř.: x_t, y_t

Odhadneme trendové hodnoty: $\hat{T}_{x,t}, \hat{T}_{y,t}$

Určíme rezidua: $e_{x,t} = x_t - \hat{T}_{x,t}, e_{y,t} = y_t - \hat{T}_{y,t}$

Určíme korelační koeficient hodnot x_t a y_t : r_{xy}

Určíme korelační koeficient reziduí: r_{e_y, e_x}

Zdánlivá korelace: Je možné sledovat vysokou hodnotu korelačního koeficientu mezi proměnnými, které nemusí být závislé.

- Zkoumání vzájemných vztahů mezi hodnotami jedné č.ř.
- Analýza závislostí mezi sousedními členy jedné č.ř.

Koeficient autokorelace prvního řádu:

Míra těsnosti závislosti mezi y_t a y_{t+1}

- Analýza závislostí dvou členů č.ř. mezi nimiž je $(k - 1)$ jiných pozorování

Koeficient autokorelace k -tého řádu:

Míra těsnosti závislosti mezi y_t a y_{t+k} , $t = 1, \dots, n - k$

$$r_k = \frac{\sum_{t=1}^{n-k} \left(y_t - \frac{1}{n-k} \sum_{t=1}^{n-k} y_t \right) \left(y_{t+k} - \frac{1}{n-k} \sum_{t=1}^{n-k} y_{t+k} \right)}{\sqrt{\sum_{t=1}^{n-k} \left(y_t - \frac{1}{n-k} \sum_{t=1}^{n-k} y_t \right)^2} \sqrt{\sum_{t=1}^{n-k} \left(y_{t+k} - \frac{1}{n-k} \sum_{t=1}^{n-k} y_{t+k} \right)^2}}$$

Příklad (prodej automobilů - klouzavý průměr
vyrovnává č.ř.)

Prodej automobilů ve stovkách:

Rok	Prodej aut
1995	710
	1 440
	980
	740
1996	770
	1 530
	1 020
	760
1997	920
	1 300
	940
	660
1998	890
	1 220
	820
	840
1999	1 000
	1 470
	930
	850

Model denní automobilové dopravy

Praktické využití jednoduchého lineárního modelu
- statistický přístup k modelování denní automobilové dopravy (použit v roce 1991 v Oslo).

Cíl: Odhadnout účinnost zavedení poplatků na vybraných frekventovaných silnicích v Oslo.

Poplatky za užívání některých silnic v Oslo - zavedeny 1.2.1990.

Rok před zavedením poplatků

- instalovány měřicí stanice na 16 místech

Rok po zavedení poplatků

- rozmístěno 50 měřících stanic na 30 silnicích
- zjišťován počet projíždějících automobilů v období od 1.1.1991 do 31.1.1992 (762 dní)

Průběh sledování počtu projíždějících automobilů:

- nepřetržitě na nejdůležitějších silnicích
- pouze několik týdnů každý rok na některých silnicích

Počet měření v každé měřicí stanici: 15 - 640.

Popis denní dopravy pomocí šesti komponent:

- Obecná úroveň
- Trend (pokles nebo růst)
- Sezonní efekt (variabilita opakující se každý rok)
- Vlivy způsobené dnem v týdnu
- Speciální dny (velikonoce, vánoce a ostatní volné dny)

Zvolený matematický model: *Multiplikativní*

Denní doprava =

Úroveň.trend.sezóna.den v týdnu.spec.dny.chyba

Model pro (přirozený) log denní dopravy: *Aditivní*

Model denní dopravy (pro každou stanici):

lineární regresní model

$$\log(\text{denní doprava v den } t) = y_t = \beta_0 + \sum_{i=1}^{46} \beta_i X_{it} + \epsilon_t,$$

- y_t - logaritmus denní dopravy v den t
 - X_{it} ($i = 1, 2, \dots, 46$) - vysvětlující veličina v den t (trend, sezonní vlivy, vliv dne v týdnu a speciální dny)
 - ϵ - chyba v den t
 - β_0 je konstantní člen reprezentující úroveň
 - β_i ($i = 1, 2, \dots, 46$) - určují vliv vysvětlujících veličin.
- 50 měřících stanic
 - 47 neznámých regresních parametrů pro každou stanici
 - Celkem 2 350 parametrů
 - Na některých stanicích k dispozici jen 15 měření

Použita metoda odhadu neznámých parametrů:

simultánní odhadování parametrů pro všechny měřící stanice

Výsledek zpracování údajů:

provoz se snížil na všech silnicích, kde byly zavedeny poplatky