Základní koncepce modelování časových řad

- cílem modelování ČŘ je jejich analýza a prognóza vývoje.
- snaha pochopit na základě abstrakce mechanismus chování ČŘ.

1. Jednorozměrný model

- nejjednodušší koncepce modelování ČŘ.
- reálná hodnota ČŘ (y_t) je funkcí času (t).

$$y_t = f(t;\varepsilon_t)\ ; \quad Y_t = f(t)\ ; \quad t = 1, 2, \ldots, n$$

$$y_t = Y_t + \varepsilon_t$$

$$\varepsilon_t = y_t - Y_t$$

t … časová proměnná

y_t … reálná hodnota ukazatele v čase *t*

Y_t … modelová (teoretická) hodnota ukazatele v čase *t*

ε_t … nepravidelná (náhodná) složka (porucha) v čase *t*.

Klasický (formální) model:

- zkoumá vliv časového faktoru na analyzovaný ukazatel.
- nezkoumá věcné příčiny dynamiky ČŘ.
- jeden z možných klasických přístupů je **dekompozice ČŘ**, tj. rozklad ČŘ na čtyři složky časového pohybu; popis jednotlivých forem pohybu.
- ČŘ může, ale nemusí všechny tyto složky obsahovat.

T_t … ***trendová složka (trend)*** = dlouhodobá tendence ve vývoji hodnot analyzovaného ukazatele v čase.

S_t … ***sezónní složka*** = pravidelně se opakující odchylka od trendu; periodicita maximálně 1 rok.

C_t … ***cyklická složka*** = kolísání okolo trendu v důsledku dlouhodobého vývoje; periodicita delší než 1 rok.

ε_t … ***náhodná složka*** = náhodné výkyvy, které nemají systematický charakter.

Podle způsobu rozkladu rozlišujeme dva základní typy modelu:

aditivní model: $y_t = T_t + S_t + C_t + \varepsilon_t$

multiplikativní model: $y_t = T_t \cdot S_t \cdot C_t \cdot \varepsilon_t$

2. Vícerozměrný model

- je založen na předpokladu, že vývoj analyzovaného ukazatele není ovlivněn pouze časovým faktorem, ale rovněž skupinou jiných, souvisejících ukazatelů.
- jedná se o tzv. *příčinné (faktorové) ukazatele*.
- pomocí těchto ukazatelů se snažíme vývoj analyzovaného ukazatele vysvětlit.

$$Y_t = f(t; x_1, x_2, \ldots, x_p)\ ; \quad t = 1, 2, \ldots, n$$

kde $x_1, x_2, \ldots, x_n$ jsou příčinné (faktorové) ukazatele.

Trendová analýza

- popis trendu (vyrovnání, vyhlazení) pomocí matematické (analytické, „hladké“) funkce = trendová funkce.
- tzv. modely s konstantními parametry.
- informace o charakteru hlavní tendence ve vývoji ukazatele.
- modelování vývoje trendu v budoucnu (prognóza = stanovení očekávaných hodnot ukazatele).

Předpoklad: $S_t = 0$ a $C_t = 0$ $\Rightarrow$ $Y_t = T_t$; $y_t = T_t + \varepsilon_t$.

<u>Lineární trend :</u>

- nejběžněji používaný, předností je jednoduchost.

$$T_t = \alpha_0 + \alpha_1 t \quad , \quad t = 1,2, ..., \ n .$$

- parametry α_0 a α_1 odhadneme pomocí metody nejmenších čtverců, čímž získáme výslednou rovnici odhadované trendové přímky:

$$\widehat{T}_t = a_0 + a_1 t \quad , \quad t = 1,2, ..., \ n .$$

$$a_0 = \bar{y} - a_1 \bar{t}$$

$$a_1 = \frac{\sum_{t=1}^{n} t\, y_t - \bar{t} \sum_{t=1}^{n} y_t}{\sum_{t=1}^{n} t^2 - n\overline{t^2}} = \frac{n\sum_{t=1}^{n} t\, y_t - \sum_{t=1}^{n} t \sum_{t=1}^{n} y_t}{n\sum_{t=1}^{n} t^2 - \left(\sum_{t=1}^{n} t\right)^2} = \frac{\overline{t\, y_t} - \bar{t} \cdot \overline{y_t}}{\overline{t^2} - \bar{t}^2}$$

a_0 …… počáteční hodnota trendu pro $t = 0$

a_1 …… průměrný absolutní přírůstek (úbytek) hodnoty $\widehat{T}_t$ při zvýšení časové proměnné t o jednotku.

Parabolický (kvadratický) trend : $T_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2$

Exponenciální trend : $T_t = \alpha_0 \cdot \alpha_1^t$

Modifikovaná (posunutá) exponenciála: $T_t = \kappa + \alpha_0 \cdot \alpha_1^t$

Logistický trend: $T_t = \dfrac{\kappa}{1 + \alpha_0 \cdot \alpha_1^t}$

Existuje řada dalších typů trendových funkcí, v konkrétní situaci je třeba vždy zvolit takový, který nejlépe vystihuje empirická data.

<u>Volba vhodného typu trendové funkce</u>

- věcná analýza zkoumaného jevu.
- posouzení empirické časové řady (ex ante).
- analýza grafu (pozor na subjektivitu).

Míry úspěšnosti zvolené trendové funkce:

- měří kvalitu zvoleného modelu (funkce), je jich celá řada.

Střední kvadratická (čtvercová) chyba odhadu: $M.S.E. = \dfrac{\sum_{t=1}^{n}\left(y_t - \widehat{T}_t\right)^2}{n}$

- nejčastější měřítko kvality modelu.
- přednost dáváme vždy tomu modelu, u něhož je hodnota M.S.E. nejnižší.
- prostřednictvím M.S.E. můžeme srovnávat jen funkce se stejným počtem parametrů.

Statistika F: $F = \dfrac{\dfrac{S_T}{p-1}}{\dfrac{S_R}{n-p}} = \dfrac{\dfrac{\sum\left(\widehat{T}_t - \overline{y}\right)^2}{p-1}}{\dfrac{\sum\left(y_t - \widehat{T}_t\right)^2}{n-p}}$, kde $\overline{y} = \dfrac{\sum \widehat{T}_t}{n}$

- za nejlepší považujeme model, pro který je hodnota statistiky F nejvyšší.

Index determinace: $I^2 = \dfrac{S_T}{S_y} = \dfrac{\sum\left(\widehat{T}_t - \overline{y}\right)^2}{\sum\left(y_t - \overline{y}\right)^2}$

- vhodnější je ta funkce, která vede k vyšší hodnotě I^2.
- u funkcí s vyšším počtem parametrů vychází I^2 nadhodnocené; v takovém případě je vhodné použít I^2_{adj} .