

Nové možnosti rozvoje vzdělávání na Technické univerzitě v Liberci

Specifický cíl A3: Tvorba nových profesně zaměřených studijních programů

NPO_TUL_MSMT-16598/2022



Časové řady – část 1

Ing. Vladimíra Hovorková Valentová, Ph.D.

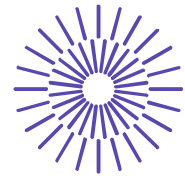


Financováno
Evropskou unií
NextGenerationEU



Národní
plán
obnovy

MSMT
MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



ČASOVÉ ŘADY

- posloupnost chronologicky uspořádaných pozorování.
- posloupnosti věcně a prostorově srovnatelných pozorování, která jsou jednoznačně uspořádána z hlediska času.
- existují různé typy časových řad.

A. Rozdělení ČŘ podle časového hlediska rozhodného pro zjišťování údajů:

1. ČŘ intervalové:

- řady intervalových ukazatelů (př. počet rozvodů za rok v ČR).
- intervalový ukazatel = ukazatel, jehož velikost závisí na délce intervalu, za který je sledován.
- lze vytvořit součty (resp. průměry), je nutná stejná délka intervalů.
- pokud jsou intervaly různě dlouhé, je třeba provést přepočítání na jednotkový interval (tzv. očišťování od důsledků kalendářních variací).

2. ČŘ okamžikové:

- řady okamžikových (stavových) ukazatelů (př. počet obyvatel ČR k 31.12.).
- okamžikový ukazatel = ukazatel, jehož hodnoty se vztahují ke konkrétnímu časovému okamžiku.
- součet nedává reálný smysl, průměr nelze stanovit běžným způsobem.
- k charakterizování úrovně hodnot těchto ČŘ používáme *průměr chronologický*:

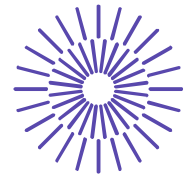
- a) Pokud jsou vzdálenosti mezi časovými okamžiky stejné, použijeme **prostý chronologický průměr**:

$$\bar{y} = \frac{\frac{y_1 + y_2}{2} + \frac{y_2 + y_3}{2} + \dots + \frac{y_{n-1} + y_n}{2}}{n-1} = \frac{y_1 + y_2 + \dots + y_{n-1} + \frac{y_n}{2}}{n-1}$$

- b) Pokud jsou mezi časovými okamžiky nesteré vzdálenosti, použijeme **vážený chronologický průměr**:

$$\bar{y} = \frac{\frac{y_1 + y_2}{2} \cdot d_1 + \frac{y_2 + y_3}{2} \cdot d_2 + \dots + \frac{y_{n-1} + y_n}{2} \cdot d_{n-1}}{d_1 + d_2 + \dots + d_{n-1}} = \frac{\sum_{i=1}^{n-1} \bar{y}_i d_i}{\sum_{i=1}^{n-1} d_i}$$

d_i vzdálenost mezi dvěma časovými okamžiky (= délka intervalu).

**B. Rozdělení ČŘ podle periodicity:****1. ČŘ roční (dlouhodobé):**

- jejich periodičita je jeden rok a více.
- aplikace jiných postupů než u ČŘ krátkodobých.

2. ČŘ krátkodobé:

- jejich periodičita je kratší než jeden rok.
- ČŘ čtvrtletní, měsíční, týdenní, atd.

C. Rozdělení ČŘ podle způsobu vyjádření ukazatelů:**1. ČŘ naturálních ukazatelů:**

- hodnoty jsou vyjádřeny v naturálních jednotkách (př. ukazatele produkce).

2. ČŘ peněžních ukazatelů:

- hodnoty jsou vyjádřeny v peněžní formě.
- je nutno zajistit souměřitelnost hodnot v čase (změny cenové hladiny!).

Základní charakteristiky časových řad

Absolutní diference (*přírůstky*) – charakterizují absolutní změnu (přírůstek nebo úbytek) hodnoty ukazatele v časovém okamžiku t oproti období předcházejícímu ($t-1$).

1. diference: $\Delta_t^{(1)} = y_t - y_{t-1}, \quad t = 2, 3, \dots, n$

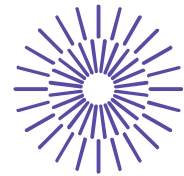
2. diference: $\Delta_t^{(2)} = \Delta_t^{(1)} - \Delta_{t-1}^{(1)}, \quad t = 3, 4, \dots, n$

3. diference: $\Delta_t^{(3)} = \Delta_t^{(2)} - \Delta_{t-1}^{(2)}, \quad t = 4, 5, \dots, n$

atd.

Průměrný absolutní přírůstek – aritmetický průměr jednotlivých prvních diferencí.

$$\bar{\Delta} = \frac{y_n - y_1}{n - 1}$$



Koeficient růstu – udává, kolikrát vzrostla hodnota ČŘ v časovém okamžiku t proti období předcházejícímu.

– $k_t \cdot 100 - 100$ udává, o kolik procent vzrostla (event. klesla) hodnota ukazatele v časovém okamžiku t oproti období předcházejícímu.

$$k_t = \frac{y_t}{y_{t-1}}, \quad t = 2, 3, \dots, n$$

Průměrný koeficient růstu – geometrický průměr jednotlivých koeficientů růstu.

$$\bar{k} = \sqrt[n-1]{k_2 \cdot k_3 \cdot \dots \cdot k_n} = \sqrt[n-1]{\frac{y_n}{y_1}}$$

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).

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

$$y_t = Y_t + \varepsilon_t$$

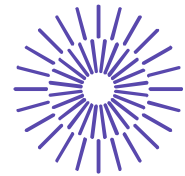
$$\varepsilon_t = y_t - Y_t$$

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; periodičita maximálně 1 rok.

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

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

Podle způsobu rozkladu rozlišujeme dva základní typy modelu:

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

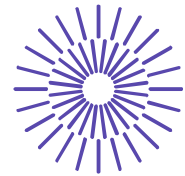
multiplikativní model: $y_t = T_t \cdot S_t \cdot C_t \cdot \varepsilon_t$

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.

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

kde x_1, x_2, \dots, 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 $C_t = 0 \Rightarrow Y_t = T_t ; y_t = T_t + \varepsilon_t$.

Lineární trend:

- nejběžněji používaný, předností je jednoduchost.

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

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

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

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

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

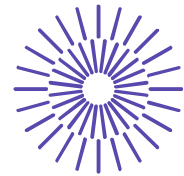
a_0 počáteční hodnota trendu pro $t = 0$

a_1 průměrný absolutní přírůstek (úbytek) hodnoty \hat{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: $T_t = \frac{\kappa}{1 + \alpha_0 \cdot \alpha_1^t}$

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

Volba vhodného typu trendové funkce

- 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: $M.S.E. = \frac{\sum_{t=1}^n (y_t - \hat{T}_t)^2}{n}$

- 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: $F = \frac{\frac{S_T}{p-1}}{\frac{S_R}{n-p}} = \frac{\sum (\hat{T}_t - \bar{y})^2}{\sum (y_t - \hat{T}_t)^2}$, kde $\bar{y} = \frac{\sum \hat{T}_t}{n}$

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

Index determinace: $I^2 = \frac{S_T}{S_y} = \frac{\sum (\hat{T}_t - \bar{y})^2}{\sum (y_t - \bar{y})^2}$

- 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_{adj}^2 .