Odpovídá to dnešní otázce

21 Vývojové trendy a časové řady - tři podmínky predikce, kinematický a dynamický přístup, cíle a úlohy *analýzy*, **koeficient korelace, autokorelace a autokorelační funkce,** tři základní modely predikce.

1. **Časová řada - trend a podmínky extrapolace (predikce)**

1.1 Časová řada

„Časovou řadou rozumíme řadu hodnot určitého ukazatele uspořádanou z hlediska přirozené časové posloupnosti, tj. od minulosti směrem k přítomnosti. Přitom je nutné, aby věcná náplň ukazatele a jeho prostorové vymezení byly shodné v celém sledovaném období." (Blatná, 2009, str. 40)

Obvyklý předpoklad o časové řadě:
Posloupnost údajů ekvidistantně rozložená v čase => JEDNODUCHOST ZPRACOVÁNÍ DAT.
Nesplnění podmínky ekvidistantnosti využití časové řady nevylučuje, ale zpracování dat to komplikuje.

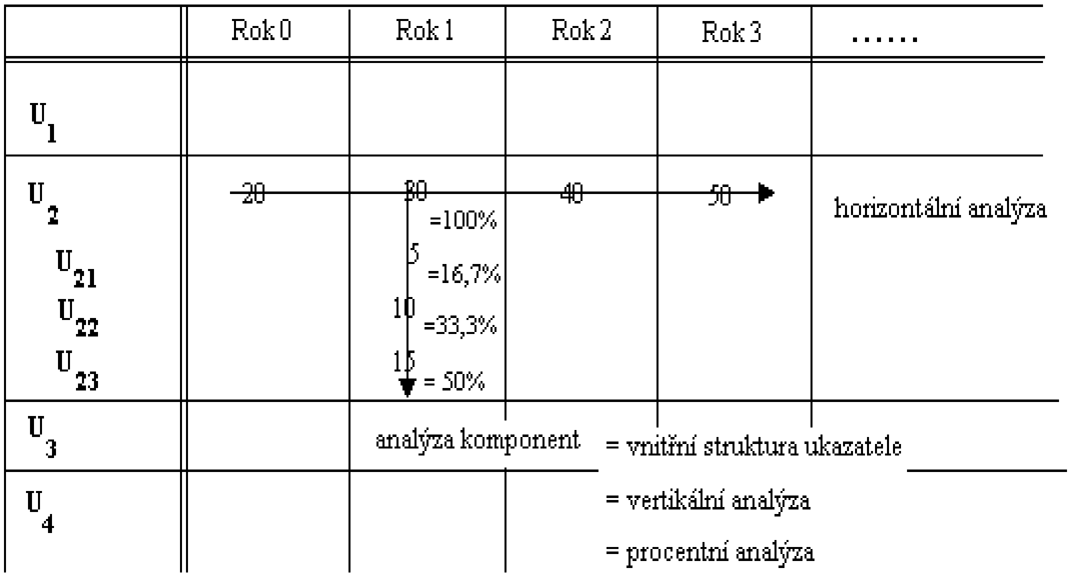
Dostatečný počet kvalitních - tj. srovnatelných - dat (průměr i rozptyl konstantní):

a) stovky (desítky) dat, či jejich dvojic
b) aplikace standardních statistických postupů předpokládá, že „rozsah výběru *n* překračuje cca 30-40 výběrových jednotek (položek, dokladů atd.)." (Hindls a kol., 2007, str. 109), přičemž by mělo zároveň platit, že
c) minimální nutná velikost vzorku (výběru)
Podle pramene (Hindls a kol., 2007, str. 112) existuje „ ….. velmi důležitý vztah, jímž můžeme stanovovat **nezbytně nutný minimální rozsah výběru** *n* za situace, kdy si předem zvolíme, s jak velkou přípustnou chybou jsme ochotni pracovat. …..

$$n \geq (u_{1-\alpha/2})^2 . (p(1-p)/\Delta^2 \qquad (2.31)“$$

Zde platí:
α …………………….. nespolehlivost výběru v %
$u_{1-\alpha/2}$ …………………. kvantil hodnoty α
p …………………….. pravděpodobnost chyby dat
Δ ……………………. přípustná chybovost základního souboru v %

Časové změny ukazatelů (nejen absolutních) = HORIZONTÁLNÍ ANALÝZA
(tzv. analýza po řádcích)



1.2 Využití trendu pro extrapolaci časové řady (predikci)

Minulé chování firmy je často dobrým indikátorem chování budoucího (PREDIKCE)

V každém případě je <u>současný stav</u> východiskem (<u>základnou</u>) budoucích aktivit, dokonce je může v jistém smyslu i <u>předurčovat.</u>

<u>Trend</u> může být dokonce i reprezentativnější než vlastní hodnoty ukazatele, zvláště v přechodových stavech (zahájení činnosti, fúze atd.)

Tři obecné podmínky využití trendu pro predikci (extrapolaci):

a) náhlé změny se objevují jen zřídka (chování, výstupy)
b) existují aspirace podniku do budoucna (cíle, záměry)
c) jsou známy <u>klíčové faktory</u> (vnitřní i podstatného okolí) a jejich změny.

Tyto obecné podmínky doplňují
„*<u>Podmínky použití klasických statistických metod k extrapolacím:</u>*
- časová řada musí být přiměřeně dlouhá,
- časová řada musí mít jednoznačný trend, který lze aproximovat co nejjednodušší analytickou funkcí,
- je třeba rozlišovat mezi krátkodobou a dlouhodobou prognózou (podle účelu),
- statistickou analýzu je třeba provádět současně s věcnou analýzou,
- kvalitu předpovědi posuzovat statistickými kritérii.“ (Blatná, 2009, str. 59)

Trendové (regresní) funkce – regresní model

Trendová funkce by měla co nejlépe vystihovat závislost prvků časové řady na čase.

„Vhodnou analytickou funkci volíme na základě:
- ♣ věcně-logického rozboru zkoumaných závislostí,
- ♣ grafického znázornění (bodového diagramu),
- ♣ pomocí matematicko- statistických kritérií“ (Blatná, 2009, str. 7),

přičemž „Platí zásada, že se snažíme k popisu závislosti použít pokud možno jednodušší funkci, která vyhovuje z hlediska uvedených kritérií (tzv. *„princip parsimonie“*), tamtéž.

Grafická analýza má podle pramene (Arlt, Arltová a Rublíková, 2002, str. 26,27) následující logiku. „Je známo, že pokud

1. řada prvních diferencí (y_t - y_{t-1}) pro t = 2, 3, ..., T kolísá okolo nuly, volíme konstantní trend,
2. řada prvních diferencí (y_t - y_{t-1}) pro t = 2, 3, ..., T kolísá okolo nenulové konstanty, volíme lineární trend,
3. řada prvních diferencí (y_t - y_{t-1}) pro t = 2, 3, ..., T má přibližně lineární trend a řada druhých diferencí (y_t - $2y_{t-1}$ + y_{t-2}) pro t = 3, 4, ..., T má přibližně konstantní trend, volíme kvadratický trend (parabolu),
4. řada koeficientů růstu y_t/y_{t-1} pro t = 2, 3, ..., T nebo řada prvních diferencí (ln y_t - ln y_{t-1}) kolísá okolo nenulové konstanty, volíme jednoduchý exponenciální trend,
5. řada ln y_t pro t = 1, 2, ..., T má přibližně hyperbolický průběh, volíme S-křivku,
6. řada podílů sousedních diferencí (y_t - y_{t-1})/(y_{t-1} - y_{t-2}) pro t = 3, 4, ..., T kolísá okolo nenulové konstanty, volíme modifikovaný exponenciální trend,
7. řada podílů sousedních diferencí (ln y_t - ln y_{t-1})/(ln y_{t-1} - ln y_{t-2}) pro t = 3, 4, ..., T kolísá okolo nenulové konstanty, volíme Gompertzovu křivku.

Výběr trendové funkce na základě grafu je subjektivní a v případě složitějších funkcí nebo mají-li časové řady velkou variabilitu, nevede k jednoznačným výsledkům.“

2. Analýza časové řady – přístupy, cíle a úlohy

2.1 Přístupy k analýze časové řady

a) Sledujeme se kinematika prvků časové řady
Jde o analýzu trendu ≐ technická analýza.
Dává odpověď na otázku „Jak se se řada/ukazatel vyvíjí v čase?“.

b) Hodnotíme dynamiku časového vývoje prvků časové řady
Zde analyzujeme vliv faktorů, ovlivňujících časovou řadu ≐ fundamentální analýza.
Získáváme odpověď na otázku „Proč se řada/ukazatel vyvíjí tak, jak se vyvíjí?“.

Převažuje kinematický pohled, místo dynamických analýz

Poplatné nerobustním metodám, tradiční statistické analýzy
tj. – aritmetické průměry
– nerobustní regresní modely (především)

2.2 Cíl a úlohy analýza (zpracování) časových řad

Cíl analýzy časových řad

a) od kinematických metod k dynamice => PRINCIPIÁLNÍ NÁSKOK FUNDAMENTÁLNÍCH ANALYTIKŮ
b) od lineárních postupů k nelineárním
c) od nerobustních modelů k robustním (vůči předpokladům i datům)

Úlohy analýzy časových řad

1. filtrace dat
2. odhad trendu (vyrovnání, regrese)
3. analýza složek
4. odhad korelace a autokorelace
5. modely časových řad (matematické) a jejich predikce

ad 1) Filtrace

Filtrace = takový odhad informační složky dat, který minimalizuje (potlačuje) vliv jejich rušivých složek.

Lineární filtr
- každý element chyby má stejnou váhu
- minimalizuje rozptyl chyb výsledku

Gnostický filtr
- malé odchylky mají plnou váhu, velké mají tím menší váhu, čím jsou odlehlejší
- necitlivý vůči krátkodobým výkyvům
- maximalizuje se informace obsažená ve výsledku (???)

ad 2) Odhad trendu (vyrovnání, regrese)

Princip:
Srovnání po sobě jdoucích hodnot ukazatele (starší U_1 a nová U_2) – růst, pokles

Změna (první diference) $= U_2 - U_1$ NE pro situace („svítí – nesvítí“)!!!
Rychlost změny $= (U_2 - U_1) / \Delta T$

Vyrovnání dat hladkou funkcí (metoda nejmenších čtverců).

ad 3) Analýza složek

Složky

a) konstanta
b) lineární složka
c) kvadratická složka atd.
d) periodická složka

Metoda určení složek – odečítání členů výchozí řady

Ad a) vytvoření rozdílové řady $U_i - U_{i-1}$ => odstraníme konstantní složku postupným opakováním tvorby rozdílové řady

Ad b) atd.

ad 4) Odhad korelace a autokorelace

Obvyklým východiskem v hodnocení závislostí mezi sledovanými proměnnými je bodový graf.

Pro prvotní orientaci tento způsob identifikace vzájemné závislosti mezi proměnnými může postačovat, nicméně "I když nám bodový diagram umožňuje subjektivně posoudit, jak silná vazba mezi proměnnými existuje, je užitečné mít k disposici nástroj, který by umožnil kvantifikovat těsnost závislosti. Statistický název pro závislost je *korelace* a míra její těsnosti se nazývá *korelační koeficient*." (Wisniewski, 1996, str. 215).

Koeficient korelace

Krajní hodnoty = 1 ⇔ přímá úměra
= -1 ⇔ nepřímá úměra

Podrobněji či přesněji hodnotí současná teorie statistiky význam číselných hodnot korelačního koeficientu podle následujícího Schématu 5.1:

Schéma 5.1 Slovní vyjádření závažnosti číselné hodnoty korelačního koeficientu

0,01-0,09 Trivial
0,10-0,29 Low to moderate
0,30–0,49 Moderate to substantial

0,50–0,69 Substantial to very strong
0,70–0,89 Very strong
0,90–0,99 Near perfect

Pramen: Upraveno podle Figure 35.3 Descriptors of various sized correlation coefficients. De VAUS, D.A. (2010): Analyzing Social Science Data. SAGE Publications, London 2010, p. 272.

Významnost korelačních koeficientů:
- jejich číselnou hodnotu lze podrobit testu, zda jde o skutečnost či náhodnou závislost (hypotézy)
- vyžaduje to vysoký počet (stovky) srovnatelných dat (a jako obvykle - mimo jiné – konstantní průměry a rozptyly)

Koeficient autokorelace (závislost U_i na U_{i-d})

Autokorelace = vzájemná korelace řady s řadou zpožděnou (U_{i-d})
Soubor hodnot **s** autokorelačních koeficientů pro postupně narůstající zpoždění **d** = autokorelační funkce

Pro stacionární řady (pouze!!!) - všechny statistické charakteristiky jsou konstantní

<u>ad 5) Modely časových řad (matematické) a jejich predikce</u>

Tři základní modely predikce:

a)predikce konstantou (triviální predikce)

$\hat{Y}(t+1) = Y(t)$ tzv. anglická predikce

Invarianty: konstanta i trend, ale s posunem řady o jeden krok dopředu

b)lineární predikce (prodloužení)

$$\hat{Y}(t+1) = Y(t) + (Y(t) - Y(t-1)) = 2 * Y(t) - Y(t-1)$$

Invarianty: konstanta i trend

c)**nerobustní (?) matematické predikce**

Box – Jenkinsův lineární model (ARMA = AutoRegresive Moving Average = AUTOREGRESNÍ MODEL S POHYBLIVÝM OKNEM)

$$X_i = \sum \qquad + \sum$$

Pozn.: Průměrný absolutní podíl predikovaných a skutečných hodnot metody ad c) může být horší než ad a)