

## 10. Časové řady

### 10.1. Pojem a klasifikace časových řad

Důležitými statistickými daty jsou **časové řady** pomocí nichž můžeme zkoumat **dynamiku jevů v čase**. Časové řady jsou určeny především:

- ke **sledování a vyhodnocování změn**, k nimž dochází ve vývoji zkoumaných jevů v závislosti na čase,
- **pro analýzu příčin**, které na tyto jevy působily a ovlivňovaly jejich chování v minulosti,
- **pro předvídání jejich budoucího vývoje**.

**Časovou řadou** (dynamická řada, vývojová řada) **rozumíme** v čase **uspořádané číselné (kvantitativní) údaje**.

Hodnoty časové řady označujeme symbolem  $Y_t$ , kde  $t$  představuje čas. Odhadnutou hodnotu časové řady označujeme  $\hat{Y}_t$ .

Množinu hodnot časové řady až do časového bodu  $t$  značíme  $Y_1, Y_2, \dots, Y_{t-1}, Y_t$ .

Pracujeme-li s více časovými řadami najednou, používáme pro jejich označení další písmena z konce abecedy –  $Z, X$  atd.

V **matematickém vyjádření** je časová řada řadou pozorovaných hodnot číselného statistického znaku uspořádaná v přirozené souvislé časové posloupnosti  $y_1, y_2, \dots, y_t, \dots, y_n$ ,

pro  $y_t, t = 1, 2, \dots, n$ , kde  $n$  je **délka časové řady**.

Rozdíl  $n - t$  nazývá **věk pozorování**.

**Časové řady mohou být spojitě a nespojitě**. Mnoho řad, které mají nespojitý charakter často převádíme na řady spojitě sčítáním, průměrováním apod.

Často tak činíme i u ekonomických časových řad.

Například výroba v podniku (zajímá nás výroba za měsíc, čtvrtletí, nikoliv však výroba za den či po hodinách - ta však může být zajímavá pro samotného výrobce), průměrná denní teplota, tlak apod.

#### Problémy časových řad

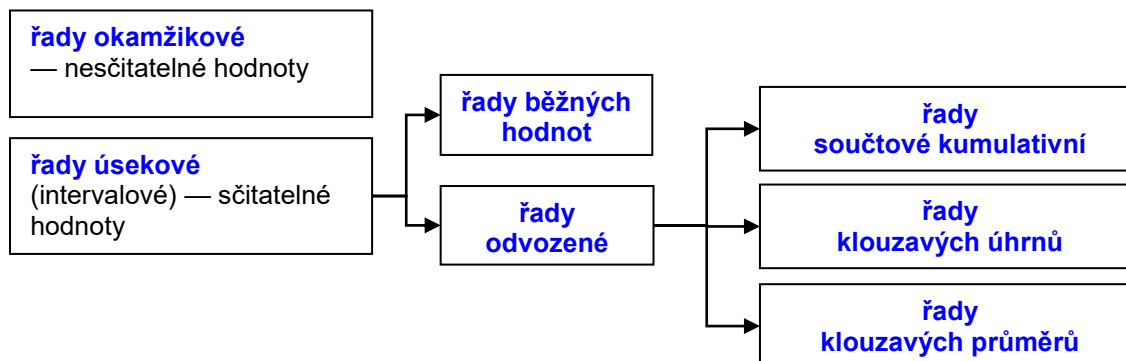
Při zpracování dat ve formě časové řady se potýkáme s množstvím, které jsou právě pro časové řady specifické. Jedná se především o

- problémy s volbou časových bodů pozorování,
- problémy s kalendářem,
  - různá délka měsíců,
  - různý počet víkendů v měsíci,
  - různý počet pracovních dnů v měsíci,
  - pohyblivé svátky,
- problémy s délkou časových řad,
- problémy nesrovnatelností dat,

## 10.2. Klasifikace časových řad

### 10.2.1. Časové řady absolutních veličin

Základní **dělení časových řad absolutních veličin** poskytuje následující schéma:



**Údaje okamžikových časových řad** se vztahují vždy k určitému časovému okamžiku např. počet pracovníků k prvnímu dni v jednotlivých měsících, stav zásob materiálu k 1.1. v jednotlivých letech. Jde o **nesčitatelné hodnoty**.

**Údaje úsekových časových řad** se vztahují vždy k určitému časovému úseku. Velikost údajů je tedy v přímé závislosti s délkou časových úseků, např. výroba v jednotlivých měsících roku, počet narozených dětí v jednotlivých letech. Typické je sčítání (**kumulování**) údajů. Jde o **sčitatelné hodnoty**.

**Úsekové řady** můžeme **podrobněji dělit na**:

- **řady běžných hodnot**,
- **řady odvozené**,
  - **řady součtové — kumulativní**, umožňují sledovat *postupné narůstání ukazatele od prvního časového úseku až po poslední*
  - **řady klouzavých úhrnů** - hodnoty ukazatele za období sestávající z určitého počtu dílčích úseků, přičemž každý další *úhrn v řadě přibírá údaj dalšího úseku a vypouští údaj nejstaršího úseku*
  - **řady klouzavých průměrů** - řady *klouzavých úhrnů dělené počtem úseků*, za které jsou klouzavé úhrny počítány

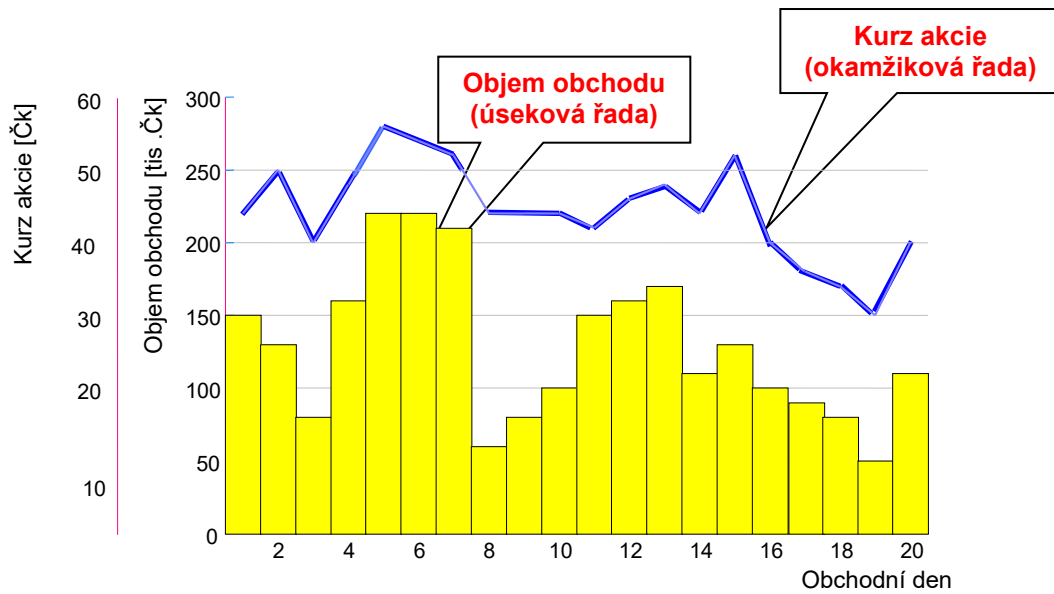
Při **grafickém znázorňování** úsekových časových řad se používají zejména sloupcové grafy, stupňovité čáry a spojnicové grafy (vynášení hodnot nad středy úseků). Z **kombinace řady běžných hodnot, kumulované řady a řady klouzavých úhrnů** se sestavuje tzv. **Z – diagram**

V ekonomické oblasti jsou typické úsekové a okamžikové časové řady **denních, týdenních, měsíčních, čtvrtletních, ročních** údajů.

### 10.2.2. Časové řady odvozených veličin

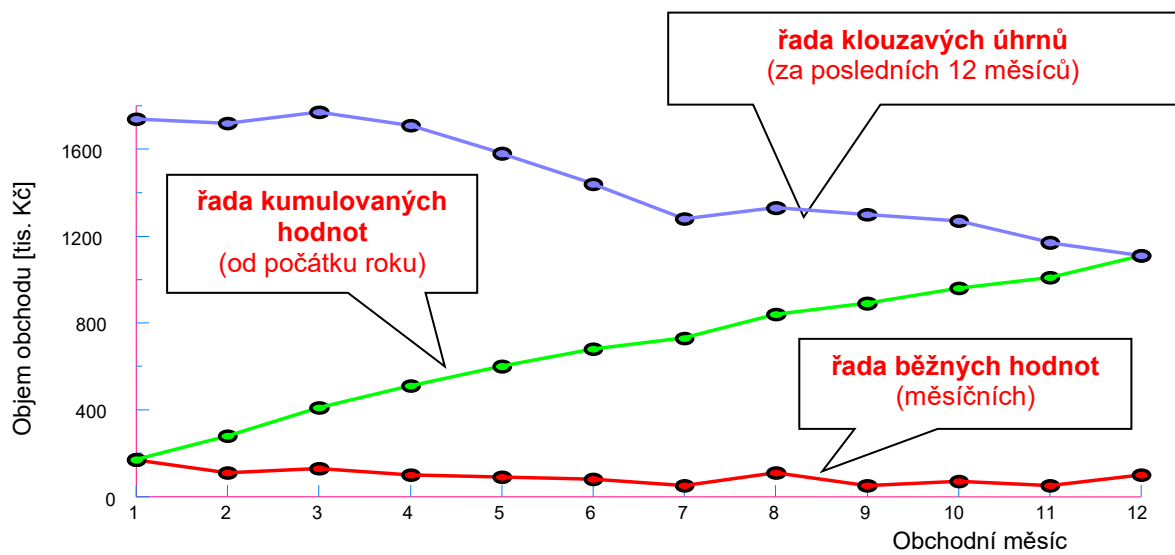
- **časové řady poměrných veličin** - např. plnění plánu v jednotlivých měsících, produktivita práce dosažená v jednotlivých letech,
- **časové řady průměrných veličin** - např. průměrná mzda pracovníků v jednotlivých letech, průměrná spotřeba masa na jednoho obyvatele v jednotlivých letech

**Příklad úsekové a okamžikové řady:**



Obr. 10.1. Kurz akcií a objem obchodu ve 20 obchodních dnech

**Příklad odvozených řad — Z–diagram — pro objem obchodování akcií:**



Obr. 10.1. Z diagram pro objem obchodování akcií

### 10.3. Měření úrovně časových řad

**Úsekové řady** — k měření úrovně se využívá **prostý aritmetický průměr** (vzhledem ke sčitatelnosti údajů lze např. z měsíčních údajů určit roční úhrn a jeho vydělením počtem měsíců stanovit průměrnou hodnotu připadající na jeden měsíc).

**Okamžikové řady** — vzhledem k nesčitatelnosti údajů se okamžiková řada o délce  $n$  převádí na úsekovou řadu o délce  $n-1$ , jejíž jednotlivé hodnoty jsou dány jako průměry sousedních hodnot původní řady  $\frac{y_{t-1} + y_t}{2}$ .

**Prostý nebo vážený aritmetický průměr** z těchto hodnot se nazývá **chronologický průměr**.

- **prostý chronologický průměr** při stálé vzdálenosti mezi okamžiky zjišťování

$$\bar{y}_{ch} = \frac{1}{n-1} \sum_{t=2}^n \frac{y_{t-1} + y_t}{2} = \frac{1}{n-1} \left( \frac{y_1}{2} + y_2 + \dots + y_{n-1} + \frac{y_n}{2} \right),$$

- **vážený chronologický průměr**, jsou-li vzdálenosti mezi okamžiky zjišťování pohyblivé a rovné  $w_t$  (pro vzdálenost mezi  $t$ -tým a  $(t-1)$  okamžikem)

$$\bar{y}_{ch} = \frac{1}{\sum_{t=2}^n w_t} \sum_{t=2}^n \frac{y_{t-1} + y_t}{2} w_t$$

### 10.4. Míry dynamiky časových řad

- **absolutní přírůstek**  $\Delta y_t = y_t - y_{t-1}$ ,  $t = 2, 3, \dots, n$

- **průměrný absolutní přírůstek**

$$\bar{\Delta} = \frac{\sum \Delta y_t}{n-1} = \frac{(y_2 - y_1) + (y_3 - y_2) + \dots + (y_n - y_{n-1})}{n-1} = \frac{y_n - y_1}{n-1}$$

- **relativní přírůstek**

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

- **průměrný koeficient růstu**

$$\bar{k} = \sqrt[n-1]{k_1 k_2 \dots k_n} = \sqrt[n-1]{\frac{y_2}{y_1} \frac{y_3}{y_2} \frac{y_4}{y_3} \dots \frac{y_n}{y_{n-1}}} = \sqrt[n-1]{\frac{y_n}{y_1}}$$

## 10.5. Analýza časových řad

Cílem analýzy je většinou konstrukce vhodného modelu. Pokud budeme schopni sestavit dobrý model, umožní nám to porozumět mechanismu, na jehož základě vznikají hodnoty časové řady, a porozumět podmínkám, které vznik těchto hodnot ovlivňují. To nám umožní tyto podmínky ovlivňovat a v některých případech ovlivnit i vývoj časové řady.

Dalším velmi častým cílem je konstrukce předpovědí.

Při klasické analýze časových řad se vychází z předpokladu, že **každá časová řada může obsahovat čtyři složky:**

- **trend,**
- **sezónní složku,**
- **cyklickou složku,**
- **náhodnou složku.**

**Trend** je obecná tendence vývoje zkoumaného jevu za dlouhé období. Je výsledkem dlouhodobých a stálých procesů.

Trend může být **rostoucí, klesající** nebo může existovat řada **bez trendu**.

**Sezónní složka** je pravidelně se opakující odchylka od trendové složky. Perioda této složky je menší než celková velikost sledovaného období.

**Cyklická složka** udává kolísání okolo trendu v důsledku dlouhodobého cyklického vývoje (požíváno spíše v makroekonomických úvahách).

**Náhodná (stochastická) složka** se nedá popsat žádnou funkcí času. "Zbývá" po vyloučení trendu, sezónní a cyklické složky.

Nejčastěji se při analýze časové řady předpokládá **aditivní model** popisu chování řady.

Předpokládá se, že jednotlivé složky vývoje se sčítají  $y_t$ , takže platí:

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

kde na pravé straně po řadě vystupují složky trendová  $T_t$ , sezónní  $S_t$ , cyklická  $C_t$  a náhodná  $\varepsilon_t$ .

Různé modifikace modelů vzniknou, když některou složku z úvah vypustíme.

Analýza složky kteréhokoliv typu se provádí v podstatě klasickou regresní analýzou.

Podstatný rozdíl je jen v tom, že nezávisle proměnná, je v tomto případě proměnná časová a můžeme ji vcelku libovolně vyjádřit v jakýchkoliv časových jednotkách s libovolným počátkem.

### 10.5.1. Analýza trendové složky

**Analýza trendové složky** je zřejmě nejdůležitější částí analýzy časových řad.

V průběhu let se potvrdilo, že při **výběru trendových funkcí** většinou vystačíme s úzkou nabídkou funkcí.

Nejčastěji používané trendové funkce jsou:

lineární trend	$y_t = a_0 + a_1 t$	Parametr $a_1$ představuje přírůstek hodnoty $y$ připadající na jednotkovou změnu časové proměnné.
polynomický trend	$y_t = a_0 + a_1 t + a_2 t^2 + \dots + a_k t^k$	Umožňuje najít trendovou funkci, která má extrém.
exponenciální trend	$y_t = a_0 a_1^t$	Parametr $a_1$ představuje průměrný přírůstek hodnot $y_t$ . (Ty se chovají jako členy geometrické posloupnosti.)
modifikovaný exponenciální trend	$y_t = k + a_0 a_1^t$	Funkce má vodorovnou asymptotu a dá se pomocí ní snáze modelovat vývoj jevů, které vycházejí z omezených zdrojů růstu a u kterých existuje určitá mez nasycení, daná např. zájmem nebo potřebou určitého výrobku.
logistický trend, logistika	$y_t = \frac{1}{k + a_0 a_1^t}$ , nebo $\frac{1}{y_t} = k + a_0 a_1^t$	Křivka má tři úseky, první je charakterizován pozvolným vzestupem, druhá v okolí inflexního bodu prudkým růstem a třetí určitou vrcholovou stagnací (nasycením). Uvedený tvar je jeden z mnoha různých funkčních předpisů popisujících křivku s charakteristickým průběhem ve tvaru písmena S.
Gompertzova křivka	$y_t = k a_0^{a_1^t}$	Křivka s podobným esovitým průběhem jako logistika, ale na rozdíl od ní je asymetrická. Těžiště hodnot je až za inflexním bodem.

První tři jmenované jsou v regresní analýze běžně užívané, při čemž u exponenciály se standardně přistupuje k linearizaci logaritmováním funkčního předpisu, což poněkud získanou exponenciálu degraduje.

V ostatních případech už linearizace není možná.

K odhadu koeficientů trendových funkcí se používá různých chytrých algoritmů, které většinou byly vymyšleny v předpočítačové éře, kdy představovaly jedinou šanci aspoň nějakého odhadu dosáhnout. Dnes se dají tyto metody využít pro určení kvalifikovaných výchozích hodnot pro nejrůznější numerické metody.

### 10.5.2. Analýza sezónní složky

**Analýza sezónní složky** se často provádí až po očištění dat od trendové složky. Jde o určení časového úseku, po jehož uplynutí mají data zase stejnou hodnotu, příp. ovlivněnou trendovou a náhodnou složkou.

Pro studium sezónní složky se používá několika typů modelů.

V ekonomických modelech bývá zpravidla zřejmá velikost periody (čtvrtletí, měsíc), v jiných případech je nutno i tuto délku odhadovat (v hydrogeologii např. u výšky hladiny spodních vod). Používá se tu i harmonické analýzy, která modeluje průběh dat pomocí několika členů Fourierovy řady. Parametry se určují použitím numerických metod.

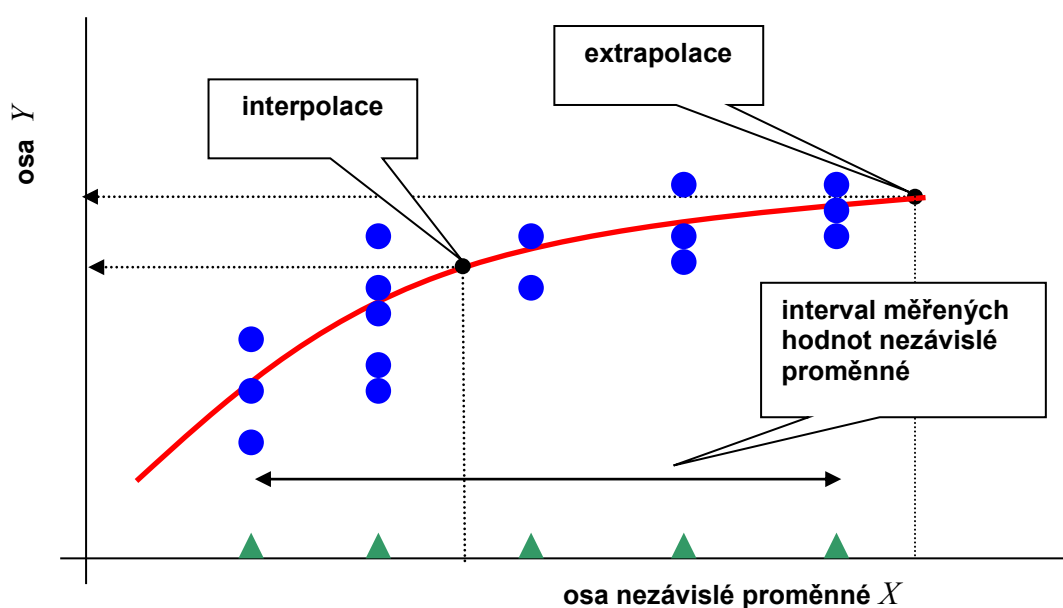
### 10.5.3. Interpolace a extrapolace

Výsledků analýzy časových řad a obecně i regresní analýzy vůbec se využívá k nalezení údajů, pro které není k dispozici výsledek měření nebo pozorování.

Pokud jde o chybějící údaj závislé veličiny  $y$  pro některou hodnotu  $x$  uvnitř intervalu známých hodnot  $x$ , jde o **interpolaci**. Ta zpravidla vede k dobrým výsledkům a nepřináší velká rizika chyb odhadované veličiny  $y$ .

Pokud však je nutno odhadnout výsledek  $y$  pro údaj  $x$  vně intervalu experimentálně udaných hodnot  $x$ , jde o **extrapolaci**.

V tomto případě je nutno být opatrný, neboť matematické prostředky použité pro určení charakteru regresní závislosti nemohou zpravidla zodpovědně odhadnout budoucí nebo minulý vývoj. Uvědomte si např., že třeba rostoucí oblouk křivky třetího stupně může velmi dobře popisovat nějakou závislost, za uvažovaným intervalem hodnot  $x$  však může dojít k nežádoucímu propadu této kubické křivky do lokálního minima (pozor na polynomu v Excel).



## 10.5.4. Schematické příklady k analýze časových řad

### Příklad na absolutní úroveň okamžikové časové řady

Počet pracovníků k 1.dni měsíce v podniku A v roce 1999

Datum	1.1.1999	1.2.1999	1.3.1999	1.4.1999	1.7.1999	1.1.2000
Počet pracovníků	148	153	142	138	133	154

**Výpočet průměrného počtu pracovníků - chronologický průměr:**

a) v prvním čtvrtletí: 
$$\bar{y}_{ch} = \frac{\frac{1}{2}y_1 + y_2 + y_3 + \dots + y_{n-1} + \frac{1}{2}y_n}{n-1} =$$
$$= \frac{\frac{1}{2}148 + 153 + 142 + \frac{1}{2}138}{4-1} = 146$$

b) v prvním pololetí: 
$$\bar{y}_{ch} = \frac{\frac{y_1+y_2}{2} \cdot v_I + \frac{y_2+y_3}{2} \cdot v_{II} + \dots}{\sum v_i} =$$
$$= \frac{\frac{148+153}{2} \cdot 1 + \frac{153+142}{2} \cdot 1 + \frac{142+138}{2} \cdot 1 + \frac{138+133}{2} \cdot 3}{1+1+1+3} = 145,75$$

c) ve druhém pololetí: 
$$\bar{y}_{ch} = \frac{133+154}{2} = 143,5$$

d) v celém roce: 
$$\bar{y}_{ch} = \frac{\frac{148+153}{2} \cdot 1 + \frac{153+142}{2} \cdot 1 + \frac{142+138}{2} \cdot 1 + \frac{138+133}{2} \cdot 3 + \frac{133+154}{2} \cdot 6}{1+1+1+3+6} = 142,125$$



### Příklad na absolutní úroveň úsekové časové řady

Výroba určitého produktu v podniku A v roce 1999:

Čas.úsek	leden	únor	březen	duben - červen	červenec - prosinec
Výroba	25	21	35	90	198

**Výpočet průměrné výroby připadající na 1 měsíc - aritmetický průměr:**

a) v prvním čtvrtletí:  $\bar{y} = \frac{\sum y_i}{n} = \frac{25 + 21 + 35}{3} = 27$

b) v prvním pololetí:  $\bar{y} = \frac{\sum y_i}{\sum n_i} = \frac{25 + 21 + 35 + 90}{1 + 1 + 1 + 3} = 28,5$

c) ve druhém pololetí:  $\bar{y} = \frac{\sum y_i}{\sum n_i} = \frac{198}{6} = 33$

d) v celém roce:  $\bar{y} = \frac{\sum y_i}{\sum n_i} = \frac{25 + 21 + 35 + 90 + 198}{1 + 1 + 1 + 3 + 6} = 30,75$

### Příklad na dynamiku časových řad

Výroba ve firmě A v letech 1993-1999

Rok	Výroba $y_i$	Absolutní přírůstek $\Delta_i$	Koeficient růstu $k_i$	$k_i'$	Koeficient růstu $k_i$ (%)	$k_i'$ (%)
1993	$y_0$ 40	-	-	-	-	-
1994	$y_1$ 35	-5	0,8750	-0,1250	87,50	-12,50
1995	$y_2$ 43	8	1,2286	0,2286	122,86	22,86
1996	$y_3$ 42	-1	0,9767	-0,0233	97,67	-2,33
1997	$y_4$ 50	8	1,1904	0,1904	119,04	19,04
1998	$y_5$ 52	2	1,0400	0,0400	104,00	4,00
1999	$y_6$ 48	-4	0,9231	-0,0769	92,31	-7,69

Průměrný absolutní přírůstek:  $\bar{\Delta} = \frac{\sum \Delta_i}{n} = \frac{8}{6} = 1,33$

Průměrný koeficient růstu:  $\bar{k} = \sqrt[n]{\frac{y_n}{y_0}} = \sqrt[6]{\frac{48}{40}} = 1,030853$

### Příklad na trend (celkový směr vývoje)

Výroba podniku A v letech 1993-1999

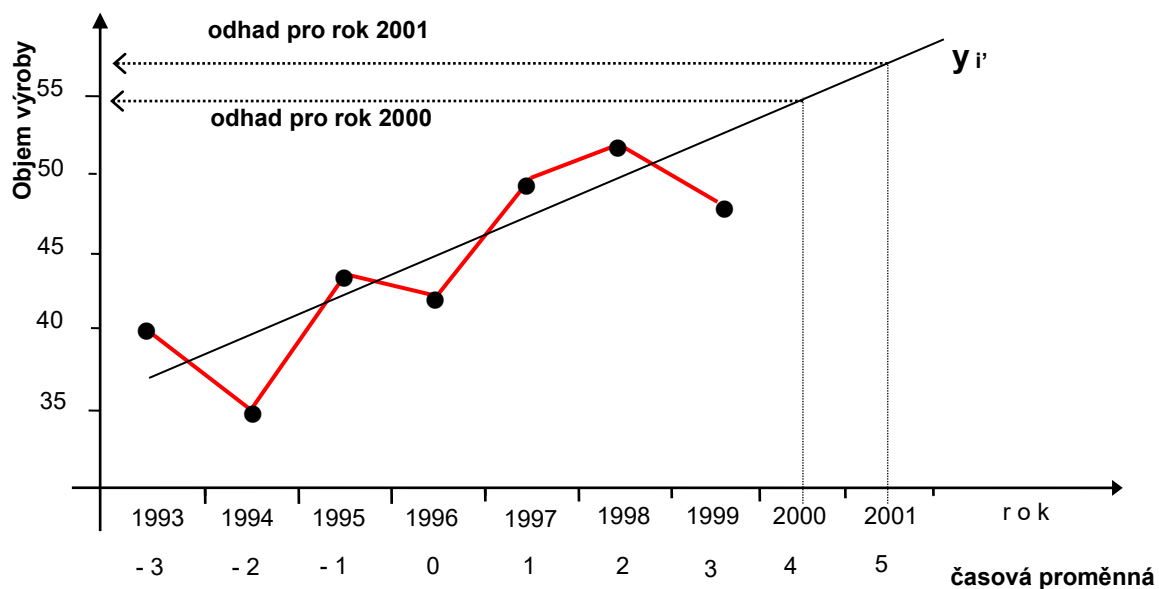
R o k	Objem výroby $y_i$	Časová proměnná $t_i$	Pomocné výpočty	
			$y_i t_i$	$t_i^2$
1993	40	- 3	- 120	9
1994	35	- 2	- 70	4
1995	43	- 1	- 43	1
1996	42	0	0	0
1997	50	1	50	1
1998	52	2	104	4
1999	48	3	144	9
$\Sigma$	<b>310</b>	<b>0</b>	<b>65</b>	<b>28</b>

Trendová funkce (přímka):  $y_i' = a + b \cdot t_i$

$$a = \frac{\sum y_i}{n} = \frac{310}{7} = 44,28$$

$$b = \frac{\sum y_i t_i}{\sum t_i^2} = \frac{65}{28} = 2,32$$

$$y_i' = 44,28 + 2,32 t_i$$



## Příklad na sezónnost (sezónní indexy)

Úrazovost v regionu A v letech 1998-2000

Rok	Čtvrtletí	Počet pracovních úrazů $y_i$	Časová proměnná $t_i$	Pomocné výpočty		Vyrovnané hodnoty
				$y_i t_i$	$t_i^2$	
1998	I	912	- 5,5	- 5016,0	30,25	1 166,5
	II	1 148	- 4,5	- 5166,0	20,25	1 157,7
	III	1 510	- 3,5	- 5285,0	12,25	1 148,9
	IV	1 115	- 2,5	- 2767,5	6,25	1 140,1
1999	I	1 010	- 1,5	- 1615,0	2,25	1 131,3
	II	1 224	- 0,5	- 612,0	0,25	1 122,5
	III	1 312	0,5	656,0	0,25	1 113,7
	IV	966	1,5	1 449,0	2,25	1 104,9
2000	I	895	2,5	2 237,5	5,25	1 096,1
	II	1 102	3,5	3 857,0	12,25	1 087,3
	III	1 203	4,5	5 413,5	20,25	1 078,5
	IV	2 020	5,5	5 610,0	30,25	1 069,7
$\Sigma$		13 417	0	- 1258,5	143,00	-

$$a = \frac{\sum y_i}{n} = \frac{13\,417}{12} = 1118,083 \quad b = \frac{\sum y_i t_i}{\sum t_i^2} = \frac{-1\,258,5}{143} = -8,8007$$

Trendová funkce (přímka):  $y_i' = 1118,1 + 8,8 \cdot t_i$

$$\text{Sezónní index } (I_s) = \frac{\text{skutečná hodnota}}{\text{vyrovnaná hodnota}} = \frac{y_i}{y_i'}$$

Čtvrtletí	Sezónní indexy				
	1998	1999	2000	1998 -2000	
				neopravené	opravené
I	78,18	89,65	81,65	83,04	83,04
II	99,16	109,04	101,35	103,18	103,18
III	131,43	117,81	111,54	120,26	120,25
IV	97,80	87,43	95,35	93,53	93,53
Celkem				400,01	400,00
Průměr					100,00

