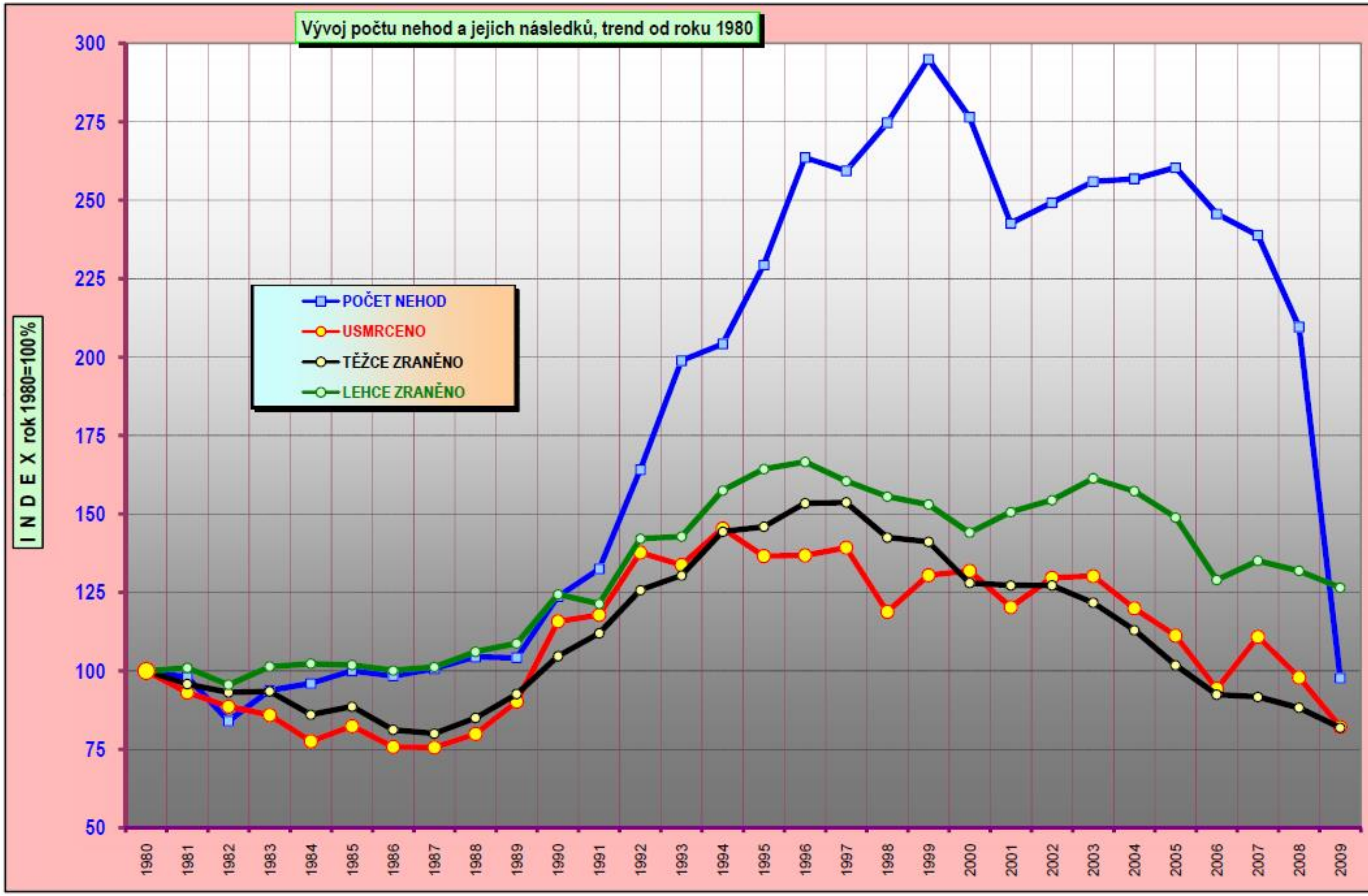
Část V. – Analýza časových řad

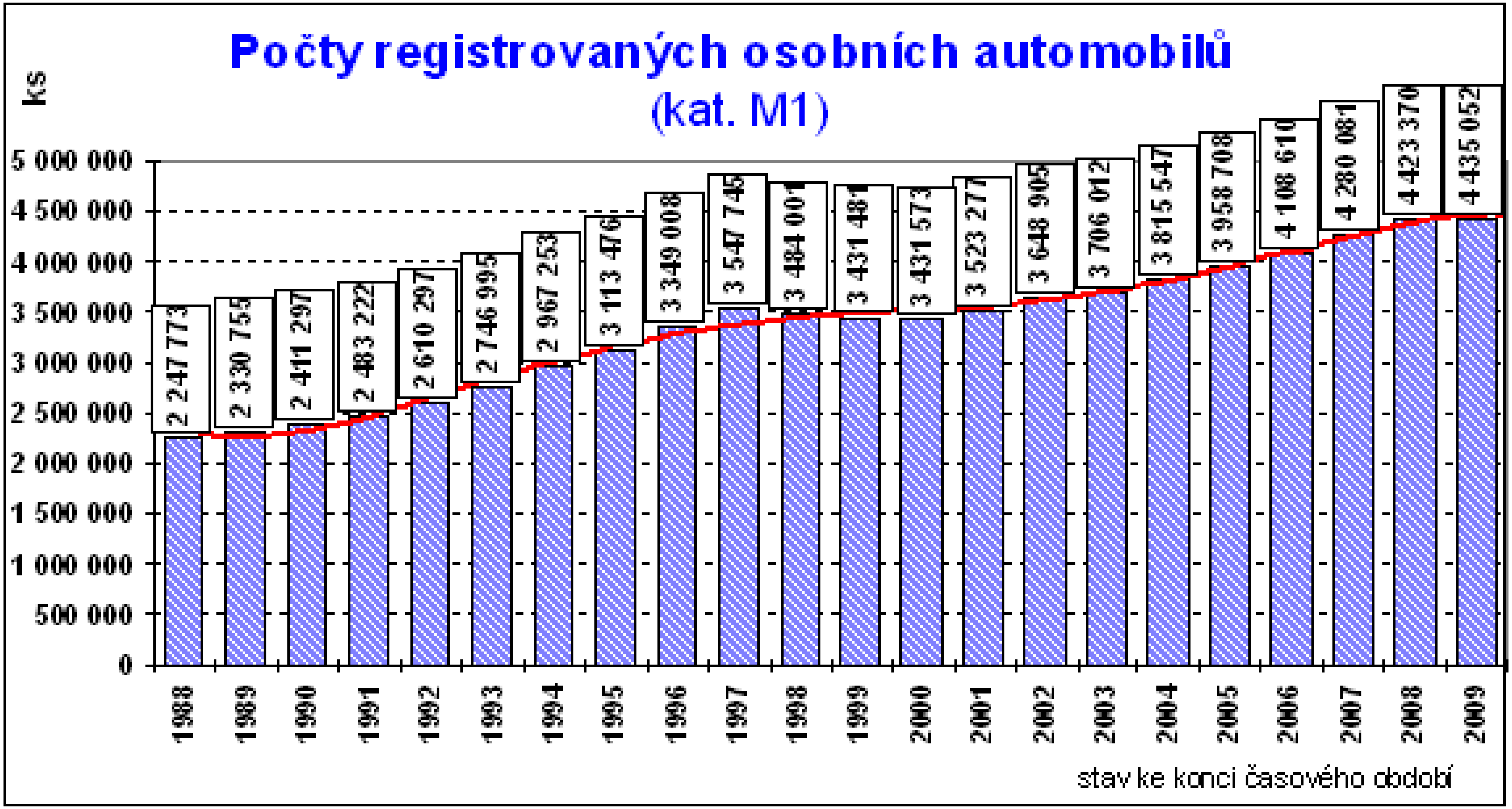
Analýza časových řad

- **Časovou řadou** rozumíme posloupnost věcně a prostorově srovnatelných pozorování (dat), která jsou jednoznačně uspořádána z hlediska času ve směru minulost – přítomnost.
- Časovou řadou v oblasti dopravy může být např. počet dopravních nehod v jednotlivých letech, počet registrovaných vozidel v jednotlivých letech apod.

Analýza časových řad



Analýza časových řad



Analýza časových řad

- Časové řady lze členit podle různých kritérií:
  1. Podle rozhodného časového hlediska.
  2. Podle periodicity, s jakou jsou údaje sledovány.
  3. Podle druhu sledovaných ukazatelů.
  4. Podle způsobu vyjádření údajů.

Analýza časových řad

1) Podle rozhodného časového hlediska rozlišujeme:
   - **Časové řady intervalové** (resp. časové řady intervalových ukazatelů).
   - **Časové řady okamžikové** (resp. časové řady okamžikových ukazatelů).

Analýza časových řad

- **Intervalovou časovou řadou** rozumíme řadu takového ukazatele, jehož velikost závisí na délce intervalu, za který je sledován.
  - Pro ukazatele tohoto typu má smysl tvořit součty, ukázkou intervalové časové řady může být řada zobrazující vývoj počtu dopravních nehod v jednotlivých letech.

Analýza časových řad

– Sledované údaje se mají vztahovat ke stejně dlouhým časovým intervalům, provádíme tzv. **očištění časových řad od vlivů kalendářních variací** (sledované údaje přepočítáváme na jednotkový časový interval)

Analýza časových řad

- Údaje očištěné na **kalendářní dny** získáme podle vztahu:

$$y_t^{(o)} = y_t \cdot \frac{\bar{k}_t}{k_t},$$

kde: y_t je hodnota očišťovaného ukazatele,

k_t je počet kalendářních dní v daném období,

$\bar{k}_t$ je průměrný počet kalendářních dní v daném období.

Analýza časových řad

$$\bar{k}_t = \frac{\sum_{t=1}^{n} k_t}{n} = \frac{366}{12} = 30{,}5$$

Měsíc	Počet dní měsíce k_t	Počet nehod y_t	Očištěný počet nehod $y_t^{(0)}$
Leden	31	18 939	18 634
Únor	29	16 137	16 972
Březen	31	17 849	17 561
Duben	30	15 724	15 986
Květen	31	17 694	17 409
Červen	30	17 914	18 213
Červenec	31	16 699	16 430
Srpen	31	17 386	17 106
Září	30	16 829	17 109
Říjen	31	19 105	18 797
Listopad	30	18 644	18 955
Prosinec	31	18 596	18 296

např. $y_1^{(0)} = y_1 \cdot \frac{\bar{k}_t}{k_1} = 18939 \cdot \frac{30{,}5}{31} \doteq 18634$

Analýza časových řad

- Údaje očištěné na **pracovní dny** získáme podle vztahu:

$$y_t^{(o)} = y_t \cdot \frac{\overline{p}_t}{p_t},$$

kde: y_t je hodnota očišťovaného ukazatele,

p_t je počet pracovních dní v daném období,

$\overline{p}_t$ je průměrný počet pracovních dní v daném období.

Analýza časových řad

- **Okamžikové časové řady** jsou tvořeny z údajů, které se vztahují k určitému okamžiku.
  - Příkladem může být počet evidovaných vozidel v ČR k 31. 12. každého roku.
  - U těchto řad nemá smysl stanovovat součty.
  - Řady tohoto typu se shrnují pomocí **chronologického průměru**.

Analýza časových řad

- V případě, že je délka mezi jednotlivými časovými okamžiky stejná, počítáme prostý chronologický průměr:

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

kde: $y_1, y_2, \ldots, y_{n-1}, y_n$ jsou jednotlivé hodnoty okamžikového ukazatele.

Analýza časových řad

- V případě, že délka mezi jednotlivými časovými okamžiky není konstantní, počítáme vážený chronologický průměr:

$$\bar{y} = \frac{\frac{y_1 + y_2}{2} \cdot d_1 + \frac{y_2 + y_3}{2} \cdot d_2 + \ldots + \frac{y_{n-1} + y_n}{2} \cdot d_{n-1}}{d_1 + d_2 + \ldots + d_{n-1}},$$

kde: $y_1, y_2, \ldots, y_{n-1}, y_n$ jsou jednotlivé hodnoty okamžikového ukazatele,

$d_1, d_2, \ldots, d_{n-1}$ jsou délky jednotlivých časových intervalů.

Analýza časových řad

Datum	Počet zaměstnanců y_t	Délka časové mezery d_t
1.1.2009	152	31
1.2.2009	164	28
1.3.2009	158	31
1.4.2009	174	30
1.5.2009	176	31
1.6.2009	171	

$$\bar{y} = \frac{\frac{152+164}{2}\cdot 31 + \frac{164+158}{2}\cdot 28 + \frac{158+174}{2}\cdot 31 + \frac{174+176}{2}\cdot 30 + \frac{176+171}{2}\cdot 31}{31+28+31+30+31} \doteq 167$$

Analýza časových řad

2) Podle periodicity, s jakou jsou údaje sledovány, rozlišujeme:
   - **Krátkodobé časové řady** (periodicita je kratší než 1 rok) – zpravidla 1 měsíc.
   - **Roční (dlouhodobé) časové řady** (periodicita je roční nebo ještě delší).

Analýza časových řad

3) Podle druhu sledovaných ukazatelů rozlišujeme:
   - **Časovou řadu absolutních hodnot** (zpravidla časová řada očištěná od kalendářních variací).
   - **Časovou řadu odvozených charakteristik** – vznikají na základě absolutních údajů, např. časové řady součtové (např. časová řada klouzavých ročních úhrnů)

Analýza časových řad

- **Klouzavým ročním úhrnem** rozumíme hodnotu intervalového ukazatele za celé roční období, které končí sledovaným měsícem.

Analýza časových řad

Měsíc	Počet nehod		Rozdíl roku	Klouzavé roční úhrny
	2005	2006	2006 - 2005	
Leden	16 961	17 219	258	199 262 + 258 = 199 520
Únor	16 375	16 789	414	199 520 + 414 = 199 934
Březen	15 527	17 748	2 221	199 934 + 2 221 = 202 155
Duben	14 168	15 598	1 430	202 155 + 1 430 = 203 585
Květen	16 827	17 031	204	203 585 + 204 = 203 789
Červen	16 707	17 996	1 289	203 789 + 1 289 = 205 078
Červenec	15 937	11 746	-4 191	205 078 - 4 191 = 200 887
Srpen	17 065	13 595	-3 470	200 887 - 3 470 = 197 417
Září	16 536	13 854	-2 682	197 417 - 2 682 = 194 735
Říjen	16 721	15 841	-880	194 735 - 880 = 193 855
Listopad	17 693	15 632	-2 061	193 855 - 2 061 = 191 794
Prosinec	18 745	14 916	-3 829	191 794 - 3 829 = 187 965
∑	199 262	187 965		

Analýza časových řad

4) Podle způsobu vyjádření údajů rozlišujeme časové řady:
   - **Naturálních ukazatelů** (hodnoty příslušného ukazatele jsou vyjádřeny naturálním kritériem).
   - **Peněžních ukazatelů** (hodnoty ukazatele jsou vyjádřeny v peněžní formě).

Analýza časových řad

- Mezi základní charakteristiky časových řad zařazujeme:
  1. Diference jednotlivých řádů (zejména 1. a 2. řádu).
  2. Tempa růstu (řetězové indexy).
  3. Průměrné tempo růstu.
  4. Průměry hodnot časové řady.

Analýza časových řad

Mějme hodnoty sledovaného ukazatele

y_t pro $t = 1,2,\ldots,n$.

1) Diferenci 1. řádu stanovíme dle vztahu:

$$D1_t = y_t - y_{t-1} \text{ pro } t = 2,\ldots,n.$$

Diferenci 2. řádu určíme podle vztahu:

$$D2_t = D1_t - D1_{t-1} \text{ pro } t = 3,\ldots,n.$$

Analýza časových řad

2) Tempa růstu určíme dle vztahu:

$$k_t = \frac{y_t}{y_{t-1}} \text{ pro } t = 2,\ldots,n \,.$$

Pokud potřebujeme tempo růstu vyjádřit v procentech, potom:

$$k_t^{\%} = 100 \cdot \left(\frac{y_t}{y_{t-1}} - 1 \right) = 100 \cdot (k_t - 1) \text{ pro } t = 2,\ldots,n \quad [\%] \,.$$

Analýza časových řad

3) Průměrné tempo růstu se stanoví jako geometrický průměr jednotlivých temp růstu:

$$\bar{k} = \left(k_2 \cdot k_3 \cdot \ldots \cdot k_n\right)^{\frac{1}{n-1}} .$$

Průměrné tempo růstu vyjádřené v % získáme podle vztahu:

$$\bar{k}^{\%} = 100 \cdot \left(\bar{k} - 1\right) [\%]$$

Analýza časových řad

4) V případě intervalové řady očištěné od vlivu kalendářních variací stanovíme průměr všech hodnot ukazatele jako aritmetický průměr:

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

V případě okamžikové řady použijeme vztahy pro výpočet chronologického průměru (viz dříve).

Analýza časových řad

Měsíc	***t***	**Počet nehod 2006**	$D1_t$	$D2_t$	k_t [-]	$k_t^{\%}$ [%]
Leden	1	17 219	-	-	-	-
Únor	2	16 789	-430	-	0,9750	-2,50
Březen	3	17 748	959	1 389	1,0571	5,71
Duben	4	15 598	-2 150	-3 109	0,8789	-12,11
Květen	5	17 031	1 433	3 583	1,0919	9,19
Červen	6	17 996	965	-468	1,0567	5,67
Červenec	7	11 746	-6 250	-7 215	0,6527	-34,73
Srpen	8	13 595	1 849	8 099	1,1574	15,74
Září	9	13 854	259	-1 590	1,0191	1,91
Říjen	10	15 841	1 987	1 728	1,1434	14,34
Listopad	11	15 632	-209	-2 196	0,9868	-1,32
Prosinec	12	14 916	-716	-507	0,9542	-4,58
∑	78	187 965				

např.

$$D1_2 = y_2 - y_1 = $$
$$= 16789 - 17219 = $$
$$= -430$$

$$D2_3 = D1_3 - D1_2 = $$
$$= 959 - 959 = $$
$$= 1389$$

Analýza časových řad

např.

$$k_2 = \frac{y_2}{y_1} = \frac{16789}{17219} \doteq 0{,}98$$

$$k_2^{\%} = 100 \cdot \left(\frac{y_2}{y_1} - 1 \right) = 100 \cdot \left(\frac{16789}{17219} - 1 \right) \doteq -2{,}50\%$$

$$\bar{k} = (k_2 \cdot k_3 \cdot \ldots \cdot k_{12})^{\frac{1}{12-1}} = (0{,}9750 \cdot 1{,}0571 \cdot \ldots \cdot 0{,}9542)^{\frac{1}{12-1}} \doteq 0{,}9870$$

$$\bar{k}^{\%} = 100 \cdot (0{,}9870 - 1) \doteq -1{,}30\%$$

Analýza časových řad

- Za základní princip modelu časové řady se používá jednorozměrný model:

  $y_t = f(t, \varepsilon_t)$,

  kde y_t je hodnota ukazatele v čase t, kde $t = 1, 2, \ldots, n$ a ε_t je hodnota náhodné složky v čase t.

Analýza časových řad

- K tomuto modelu lze přistupovat více způsoby, zpravidla se užívá **klasický (formální) model**, který dekomponuje časovou řadu na složku:
  - Trendovou (T_t).
  - Sezónní (S_t).
  - Cyklickou (C_t).
  - Náhodnou (ε_t).

Analýza časových řad

- Vlastní rozklad časové řady v aditivním tvaru potom vypadá:

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

  kde Y_t se nazývá teoretická (deterministická) složka.

- **Trendem** rozumíme hlavní tendenci dlouhodobého vývoje sledovaného ukazatele v čase – rostoucí trend, klesající trend, řada bez trendu.

Analýza časových řad

- **Sezónní složka** je pravidelně se opakující odchylka od trendové složky vyskytující se u časových řad s periodicitou menší než 1 rok.

- **Cyklickou složkou** rozumíme kolísání okolo trendu v důsledku dlouhodobého cyklického vývoje s délkou vlny delší než 1 rok.

Analýza časových řad

- **Náhodná složka** je složka, kterou nelze popsat žádnou funkcí času, jejím zdrojem jsou drobné a nepopsatelné příčiny.

- Nyní nás bude zajímat popis trendové složky pomocí trendových funkcí.

Analýza časových řad

- Nejčastěji se využívají tyto trendové funkce:
  1. Lineární trend.
  2. Parabolický trend.
  3. Exponenciální trend.
  4. Modifikovaný (posunutý) exponenciální trend.
  5. Logistický trend.
  6. Gompertzova křivka.

Analýza časových řad

1) Lineární trend

- Mějme hodnoty sledovaného ukazatele y_t pro $t = 1,2,\ldots,n$.
- Skutečný průběh $Y_t = \beta_0 + \beta_1 \cdot t$ neznáme, provádíme pouze odhad tohoto trendu ve tvaru:

$$\widehat{y}_t = b_0 + b_1 \cdot t.$$

Analýza časových řad

- Pro odhad parametrů lze použít metodu nejmenších čtverců, tedy:

$$\varphi = \sum_{t=1}^{n} \left(y_t - \widehat{y}_t\right)^2 = \sum_{t=1}^{n} \left(y_t - b_0 - b_1 \cdot t\right)^2 \rightarrow \min.$$

- Položíme-li parciální derivací rovny nule a upravíme, dostaneme:

$$1) \sum_{t=1}^{n} y_t - n b_0 - b_1 \sum_{t=1}^{n} t = 0,$$

$$2) \sum_{t=1}^{n} t \cdot y_t - b_0 \sum_{t=1}^{n} t - b_1 \sum_{t=1}^{n} t^2 = 0.$$

Analýza časových řad

- Z první rovnice vyjádříme:

$$b_0 = \frac{\sum_{t=1}^{n} y_t}{n} - b_1 \cdot \frac{\sum_{t=1}^{n} t}{n} = \bar{y}_t - b_1 \cdot \bar{t}\,.$$

- Dosazením do druhé rovnice a algebraickými úpravami dostaneme:

$$b_1 = \frac{\sum_{t=1}^{n} t \cdot y_t - \bar{y}_t \sum_{t=1}^{n} t}{\sum_{t=1}^{n} t^2 - \bar{t} \sum_{t=1}^{n} t}.$$

Analýza časových řad

- Jelikož platí:

$$\bar{y}_t = \frac{\sum_{t=1}^{n} y_t}{n} \text{ a } \sum_{t=1}^{n} t = n \cdot \bar{t},$$

můžeme psát:

$$b_1 = \frac{\sum_{t=1}^{n} t \cdot y_t - \bar{y}_t \sum_{t=1}^{n} t}{\sum_{t=1}^{n} t^2 - \bar{t} \sum_{t=1}^{n} t} = \frac{\sum_{t=1}^{n} t \cdot y_t - \bar{t} \sum_{t=1}^{n} y_t}{\sum_{t=1}^{n} t^2 - n \cdot \bar{t}^2}.$$

Analýza časových řad

- Uvedené vztahy pro výpočet odhadů parametrů modelu lze zjednodušit následující úpravou.
- Počátek časové proměnné, tedy $t = 1$ umisťujeme tam, kde máme z chronologického hlediska první pozorování.
- Zaveďme si proměnnou t' tak, aby platilo:

$$\sum_{t'} t' = 0.$$

Analýza časových řad

- Pro sudý počet pozorování např.

	leden	únor	březen	duben	květen	červen
t	1	2	3	4	5	6
t'	-3	-2	-1	1	2	3

- Pro lichý počet pozorování např.

	leden	únor	březen	duben	květen
t	1	2	3	4	5
t'	-2	-1	0	1	2

Analýza časových řad

- Jelikož platí:

$$\sum_{t'} (t')^k = 0 \text{ pro } k = 1,3,5,\ldots \text{ a } \bar{t} = 0,$$

dostaneme zjednodušené vztahy pro odhad parametrů modelu ve tvaru:

$$b_0 = \bar{y}_{t'} \quad \text{a} \quad b_1 = \frac{\sum_{t'} t' \cdot y_{t'}}{\sum_{t'} (t')^2}.$$

Analýza časových řad

- **Př.:** Byly sledovány počty prodaných automobilů v jednom roce. Stanovte rovnici lineárního trendu pro tuto časovou řadu. Dále proveďte předpověď počtu prodaných automobilů pro další dva nadcházející měsíce.

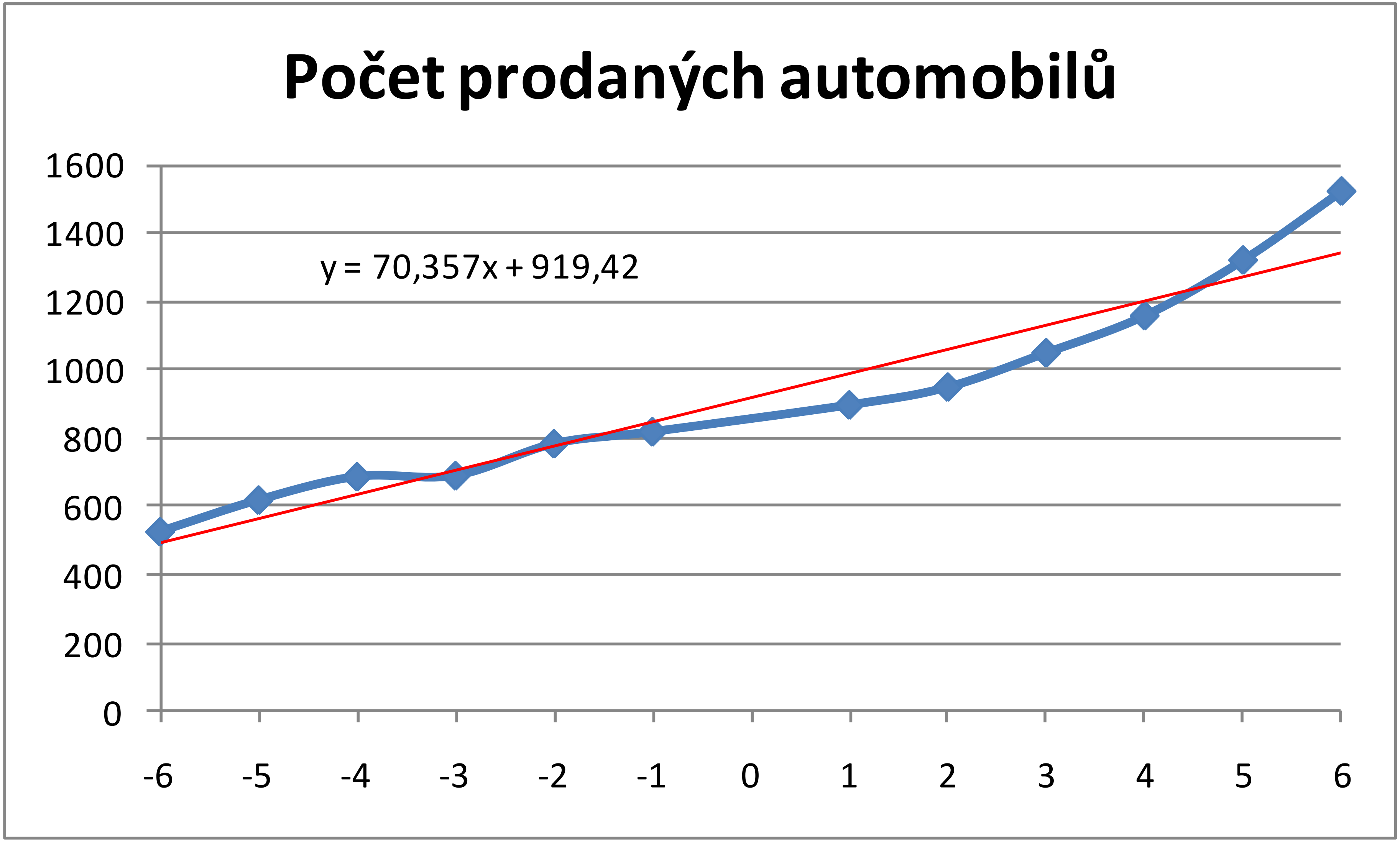
Analýza časových řad

Měsíc	t'	$y_{t'}$	$t' \cdot y_{t'}$	$(t')^2$
Leden	-6	529	-3174	36
Únor	-5	621	-3105	25
Březen	-4	689	-2756	16
Duben	-3	692	-2076	9
Květen	-2	785	-1570	4
Červen	-1	820	-820	1
Červenec	1	898	898	1
Srpen	2	950	1900	4
Září	3	1050	3150	9
Říjen	4	1158	4632	16
Listopad	5	1320	6600	25
Prosinec	6	1521	9126	36
∑	0	11033	12805	182

b_0	**919,42**
b_1	**70,36**

$$\hat{y}_{t'} = 919{,}42 + 70{,}36t'$$

Analýza časových řad



Analýza časových řad

- Předpověď pro první následující měsíc získáme dosazením $t' = 7$:

$$\hat{y}_{t'} = 919{,}42 + 70{,}36 \cdot 7 = 1411 \text{ automobilů}$$

- Předpověď pro druhý následující měsíc získáme dosazením $t' = 8$:

$$\hat{y}_{t'} = 919{,}42 + 70{,}36 \cdot 8 = 1482 \text{ automobilů}$$

Analýza časových řad

- **Př.:** Na základě předchozích sčítání intenzit je známa hodnota RPDI pro předcházející období. Odhadněte rovnici lineárního trendu pro RPDI a extrapolací odhadněte předpokládanou hodnotu RPDI v příštím období.

Analýza časových řad

Rok	t	y_t
1980	1	11523
1985	2	12201
1990	3	12948
1995	4	13578
2000	5	14987
2005	6	16012
2010	7	17065

$$b_0 = \bar{y}_t - b_1 \cdot \bar{t}$$

$$b_1 = \frac{\sum_{t=1}^{n} t \cdot y_t - \bar{y}_t \cdot \sum_{t=1}^{n} t}{\sum_{t=1}^{n} t^2 - \bar{t} \cdot \sum_{t=1}^{n} t}.$$

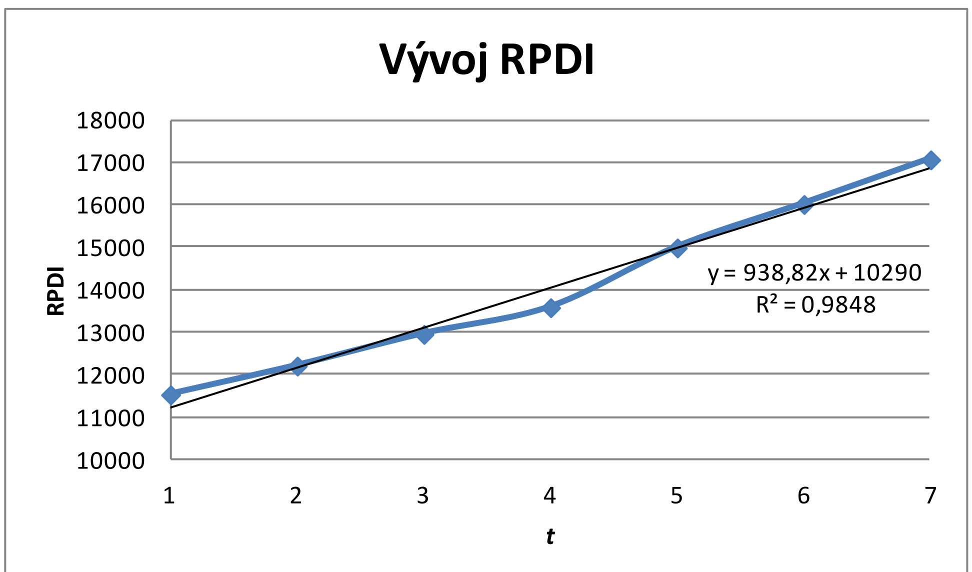
$$R^2 = \frac{\sum_{t=1}^{n} \left(\hat{y}_t - \bar{y}_t\right)^2}{\sum_{t=1}^{n} \left(y_t - \bar{y}_t\right)^2}$$

Analýza časových řad

Rok	t	y_t	$t \cdot y_t$	t^2	$\hat{y}_t$	$(\hat{y}_t - y_p)^2$	$(y_t - y_p)^2$
1980	1	11523	11523	1	11228	7932471,07	6359763,45
1985	2	12201	24402	4	12167	3525542,70	3399809,16
1990	3	12948	38844	9	13106	881385,67	1203095,59
1995	4	13578	54312	16	14045	0,00	217955,59
2000	5	14987	74935	25	14984	881385,67	887633,16
2005	6	16012	96072	36	15923	3525542,70	3869651,02
2010	7	17065	119455	49	16861	7932471,07	9121262,88
∑	28		419543	140		24678798,89	25059170,86
Průměr	4	14044,86					

b_0	10289,57
b_1	938,82
R^2	0,98

$$\hat{y}_t = 10289{,}57 + 938{,}82 \cdot t$$



Analýza časových řad

- Předpověď pro první následující období (rok 2015) získáme dosazením t=8 do rovnice trendu:

$$\hat{y}_8 = 10289{,}57 + 938{,}82 \cdot 8 \doteq 17800 \text{ vozidel} \cdot \text{den}^{-1}.$$

Analýza časových řad

2) Parabolický trend

- Pro odhad průběhu trendu lze psát:

$$\widehat{y}_t = b_0 + b_1 \cdot t + b_2 \cdot t^2 \,,$$

resp. po provedení transformace časové proměnné:

$$\widehat{y}_{t'} = b_0 + b_1 \cdot t' + b_2 \cdot (t')^2 \,.$$

Analýza časových řad

- Aplikací metody nejmenších čtverců dostaneme:

$$\varphi = \sum_{t'} \left(y_{t'} - \hat{y}_{t'}\right)^2 = \sum_{t'} \left[y_{t'} - b_0 - b_1 \cdot t' - b_2 \cdot (t')^2\right]^2 \rightarrow \min.$$

- Položíme-li parciální derivací rovny nule a upravíme, dostaneme:

$$\sum_{t'} y_{t'} - nb_0 - b_1 \sum_{t'} t' - b_2 \sum_{t'} (t')^2 = 0,$$

$$\sum_{t'} t' \cdot y_{t'} - b_0 \sum_{t'} t' - b_1 \sum_{t'} (t')^2 - b_2 \sum_{t'} (t')^3 = 0,$$

$$\sum_{t'} (t')^2 \cdot y_{t'} - b_0 \sum_{t'} (t')^2 - b_1 \sum_{t'} (t')^3 - b_2 \sum_{t'} (t')^4 = 0.$$

Analýza časových řad

- Úpravami získáme:

$$b_0 = \frac{\sum_{t'} y_{t'} \cdot \sum_{t'} (t')^4 - \sum_{t'} (t')^2 \cdot \sum_{t'} y_{t'} \cdot (t')^2}{n \sum_{t'} (t')^4 - \left[\sum_{t'} (t')^2 \right]^2},$$

$$b_1 = \frac{\sum_{t'} t' \cdot y_{t'}}{\sum_{t'} (t')^2},$$

$$b_2 = \frac{n \sum_{t'} (t')^2 \cdot y_{t'} - \sum_{t'} y_{t'} \cdot \sum_{t'} (t')^2}{n \sum_{t'} (t')^4 - \left[\sum_{t'} (t')^2 \right]^2}.$$

Analýza časových řad

- **Př.:** Byly sledovány počty prodaných automobilů v jednom roce. Stanovte rovnici parabolického trendu pro tuto časovou řadu.

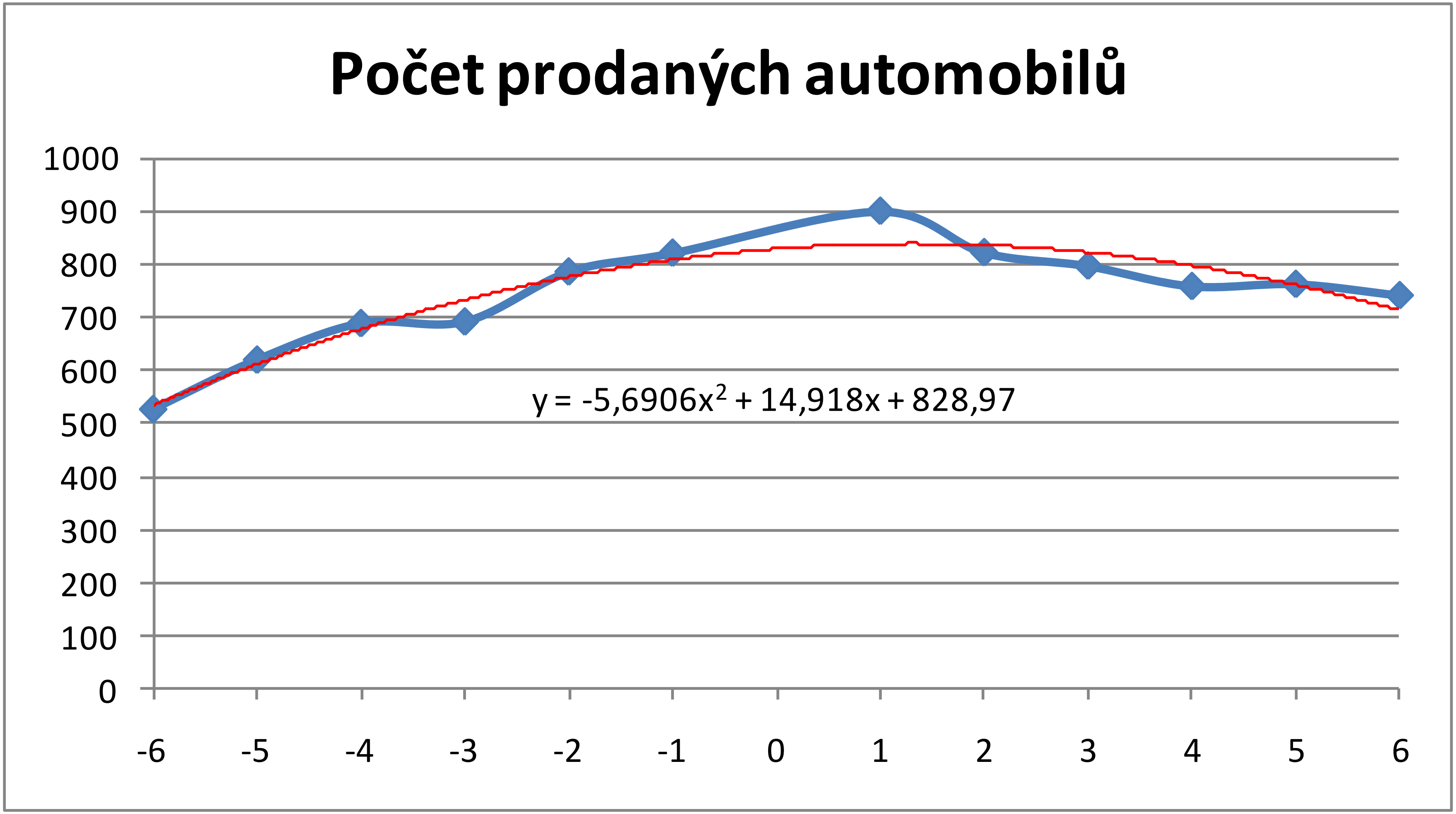
Analýza časových řad

Měsíc	t'	$y_{t'}$	$t' \cdot y_{t'}$	$(t')^2$	$(t')^4$	$(t')^2 \cdot y_{t'}$
Leden	-6	529	-3174	36	1296	19044
Únor	-5	621	-3105	25	625	15525
Březen	-4	689	-2756	16	256	11024
Duben	-3	692	-2076	9	81	6228
Květen	-2	785	-1570	4	16	3140
Červen	-1	820	-820	1	1	820
Červenec	1	898	898	1	1	898
Srpen	2	821	1642	4	16	3284
Září	3	796	2388	9	81	7164
Říjen	4	758	3032	16	256	12128
Listopad	5	762	3810	25	625	19050
Prosinec	6	741	4446	36	1296	26676
∑	0	8912	2715	182	4550	124981

b_0	828,97
b_1	14,92
b_2	-5,69

$$\hat{y}_{t'} = 828{,}97 + 14{,}92t' - 5{,}69(t')^2$$

Analýza časových řad



Analýza časových řad

- Pro posouzení kvality modelu se opět používá index determinace, který je definován stejně jako u lineárního trendu.

Analýza časových řad

3) Exponenciální trend

- Pro odhad průběhu trendu lze psát:

$\widehat{y}_t = b_0 \cdot b_1^t$ pro $b_1 > 0$.

- Tento model není lineární v parametrech, nelze přímo použít metodu nejmenších čtverců.

Analýza časových řad

- Odhad parametrů modelu lze získat:
  a) Metodou linearizující transformace a aplikací metody nejmenších čtverců.
  b) Metodou linearizující transformace a aplikací vážené metody nejmenších čtverců.

- K odhadu parametrů lze použít i jiných metod než metoda nejmenších čtverců – Metoda vybraných bodů.

Analýza časových řad

a) Nejdříve provedeme linearizující transformaci (budeme pracovat s transformovanou proměnnou *t'*):

$$\log \hat{y}_{t'} = \log\left(b_0 \cdot b_1^{t'}\right),$$
$$\log \hat{y}_{t'} = \log b_0 + t' \cdot \log b_1.$$

Označme $A = \log b_0, B = \log b_1$,

potom můžeme psát:

$$\log \hat{y}_{t'} = A + t' \cdot B.$$

Analýza časových řad

- Nyní lze aplikovat metodu nejmenších čtverců v logaritmickém tvaru, tedy:

$$\varphi = \sum_{t'} \left(\log y_{t'} - \log \hat{y}_{t'}\right)^2 = \sum_{t'} \left(\log y_{t'} - A - B \cdot t'\right)^2 \to \min.$$

- Známým postupem dostaneme:

$$\sum_{t'} \log y_{t'} - n \log b_0 - \log b_1 \cdot \sum_{t'} t' = 0 \Rightarrow \log b_0 = \frac{\sum_{t'} \log y_{t'}}{n},$$

$$\sum_{t'} t' \cdot \log y_{t'} - \log b_0 \cdot \sum_{t'} t' - \log b_1 \cdot \sum_{t'} \left(t'\right)^2 = 0 \Rightarrow \log b_1 = \frac{\sum_{t'} t' \cdot \log y_{t'}}{\sum_{t'} \left(t'\right)^2}.$$

Analýza časových řad

- Jelikož jsme použili metodu nejmenších čtverců v logaritmické formě, je nutno přistoupit ke stanovení indexu determinace rovněž v logaritmické formě:

$$R^2 = \frac{\sum_{t=1}^{n} \left(\log \hat{y}_t - \overline{\log y_t} \right)^2}{\sum_{t=1}^{n} \left(\log y_t - \overline{\log y_t} \right)^2}.$$

Analýza časových řad

b) Odhad parametrů touto metodou nemá příliš dobré statistické vlastnosti, vhodnější je použít váženou metodu nejmenších čtverců:

$$\varphi = \sum_{t'} w_{t'} \cdot \left(\log y_{t'} - \log \hat{y}_{t'}\right)^2 = \sum_{t'} w_{t'} \cdot \left(\log y_{t'} - A - B \cdot t'\right)^2 \to \min,$$

kde

$$w_{t'} = y_{t'}^2 .$$

Analýza časových řad

- Známým postupem dostaneme:

$$\sum_{t'} y_{t'}^2 \cdot \log y_{t'} - \log b_0 \cdot \sum_{t'} y_{t'}^2 - \log b_1 \cdot \sum_{t'} t' \cdot y_{t'}^2 = 0,$$

$$\sum_{t'} y_{t'}^2 \cdot t' \cdot \log y_{t'} - \log b_0 \cdot \sum_{t'} t' \cdot y_{t'}^2 - \log b_1 \cdot \sum_{t'} \left(t'\right)^2 \cdot y_{t'}^2 = 0.$$

Analýza časových řad

- Řešením bychom získali vztahy pro odhad parametrů ve tvaru:

$$\log b_0 = \frac{\sum_{t'} y_{t'}^2 \cdot \log y_{t'} \cdot \sum_{t'} (t')^2 \cdot y_{t'}^2 - \sum_{t'} t' \cdot y_{t'}^2 \cdot \sum_{t'} t' \cdot y_{t'}^2 \cdot \log y_{t'}}{\sum_{t'} y_{t'}^2 \cdot \sum_{t'} (t')^2 \cdot y_{t'}^2 - \left(\sum_{t'} t' \cdot y_{t'}^2 \right)^2},$$

$$\log b_1 = \frac{\sum_{t'} y_{t'}^2 \cdot \sum_{t'} t' \cdot y_{t'}^2 \cdot \log y_{t'} - \sum_{t'} y_{t'}^2 \cdot \log y_{t'} \cdot \sum_{t'} t' \cdot y_{t'}^2}{\sum_{t'} y_{t'}^2 \cdot \sum_{t'} (t')^2 \cdot y_{t'}^2 - \left(\sum_{t'} t' \cdot y_{t'}^2 \right)^2}.$$

Analýza časových řad

- **Př.:** Byly sledovány počty prodaných automobilů v jednom roce. Stanovte rovnici exponenciálního trendu metodou nejmenších čtverců a váženou metodou nejmenších čtverců pro tuto časovou řadu a dosažený výsledky porovnejte.

Analýza časových řad

Měsíc	t'	$y_{t'}$	$\log y_{t'}$	$t' \cdot \log y_{t'}$	$(t')^2$	$(y_{t'} - \hat{y}_{t'})^2$
Leden	-6	52	1.72	-10.30	36	17290.54
Únor	-5	258	2.41	-12.06	25	84.28
Březen	-4	529	2.72	-10.89	16	19590.38
Duben	-3	985	2.99	-8.98	9	175173.17
Květen	-2	1524	3.18	-6.37	4	488863.18
Červen	-1	2152	3.33	-3.33	1	904422.84
Červenec	1	2654	3.42	3.42	1	11603.65
Srpen	2	3215	3.51	7.01	4	242628.46
Září	3	4502	3.65	10.96	9	803718.29
Říjen	4	6215	3.79	15.17	16	2708077.41
Listopad	5	9821	3.99	19.96	25	2639502.72
Prosinec	6	15214	4.18	25.09	36	2107511.80
∑	0	47121	38.91	29.70	182	8513042.34

$\log b_0$	3,24
$\log b_1$	0,16
b_0	1748,73
b_1	1,46
$\ln b_1$	0,38

$$\log_a x = y \Rightarrow$$

$$\Rightarrow x = a^y$$

Analýza časových řad

Měsíc	t'	$(t')^2$	$y_{t'}$	$(y_{t'})^2$	$\log y_{t'}$	$t'\cdot(y_{t'})^2$	$(t')^2\cdot(y_{t'})^2$	$t'\cdot(y_{t'})^2\cdot\log y_{t'}$	$(y_{t'})^2\cdot\log y_{t'}$	$(y_{t'}-\hat{y}_{t'})^2$
Leden	-6	36	52	2704	1.72	-16224	97344	-27840.44	4640.07	32153.93
Únor	-5	25	258	66564	2.41	-332820	1664100	-802635.27	160527.05	4708.77
Březen	-4	16	529	279841	2.72	-1119364	4477456	-3048538.23	762134.56	4597.79
Duben	-3	9	985	970225	2.99	-2910675	8732025	-8712920.00	2904306.67	111414.83
Květen	-2	4	1524	2322576	3.18	-4645152	9290304	-14785448.99	7392724.49	365395.78
Červen	-1	1	2152	4631104	3.33	-4631104	4631104	-15434739.15	15434739.15	728673.34
Červenec	1	1	2654	7043716	3.42	7043716	7043716	24116985.68	24116985.68	4266.06
Srpen	2	4	3215	10336225	3.51	20672450	41344900	72502023.39	36251011.70	193830.52
Září	3	9	4502	20268004	3.65	60804012	182412036	222141711.30	74047237.10	434656.05
Říjen	4	16	6215	38626225	3.79	154504900	618019600	586105242.91	146526310.73	1150923.38
Listopad	5	25	9821	96452041	3.99	482260205	2411301025	1925257831.60	385051566.32	220428.55
Prosinec	6	36	15214	231465796	4.18	1388794776	8332768656	5808277802.43	968046300.41	467394.42
∑	0	182	47121	412465021	38.91	2100424720	11621782266	8595589475.23	1660698483.93	3718443.43

$\log b_0$	3,26
$\log b_1$	0,15
b_0	1833,33
b_1	1,41
$\ln b_1$	0,35

Analýza časových řad

- Metodou nejmenších čtverců jsme získali:

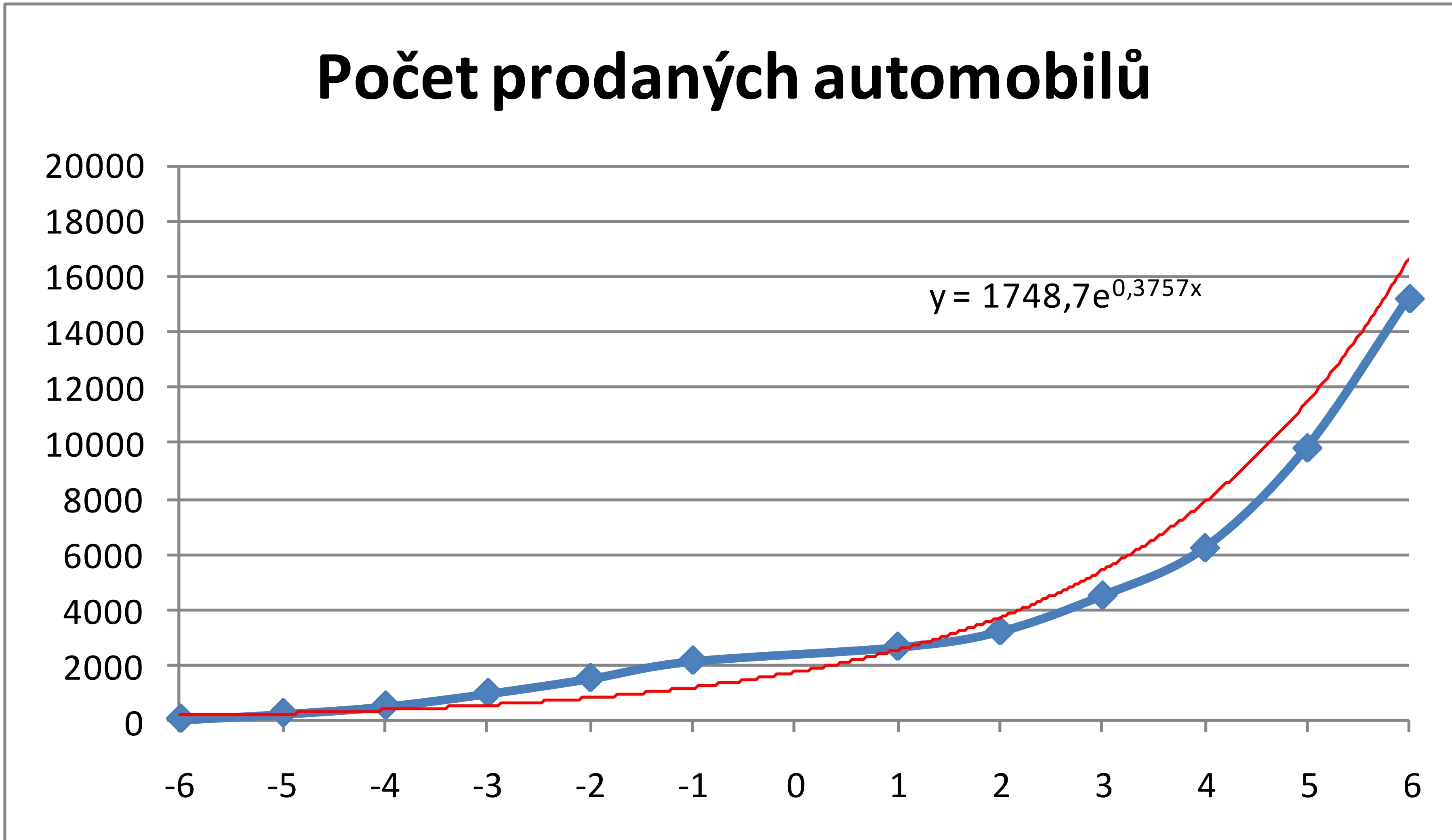
$$\varphi = \sum_{t'} \left(y_{t'} - \hat{y}_{t'} \right)^2 \doteq 8513042{,}34.$$

- Váženou metodou nejmenších čtverců jsme získali:

$$\varphi = \sum_{t'} \left(y_{t'} - \hat{y}_{t'} \right)^2 \doteq 3718443{,}43.$$

- Vidíme, že reziduální součet čtverců je v druhém případě podstatně nižší, proto bychom za odhad parametrů zvolili tyto výsledky.

Analýza časových řad



Analýza časových řad

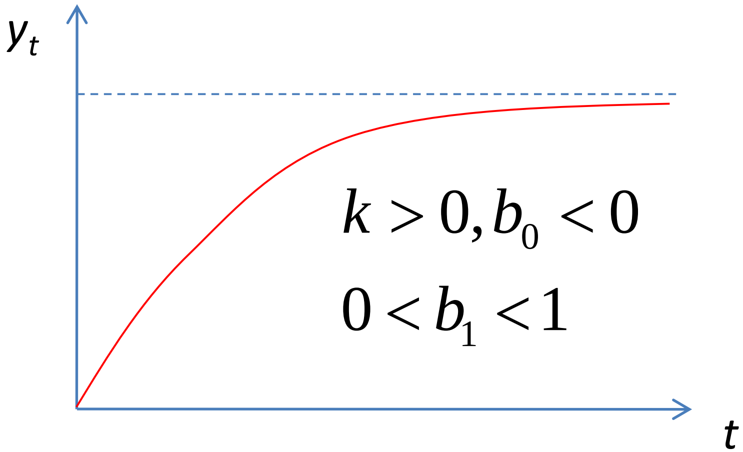
4) Modifikovaný exponenciální trend

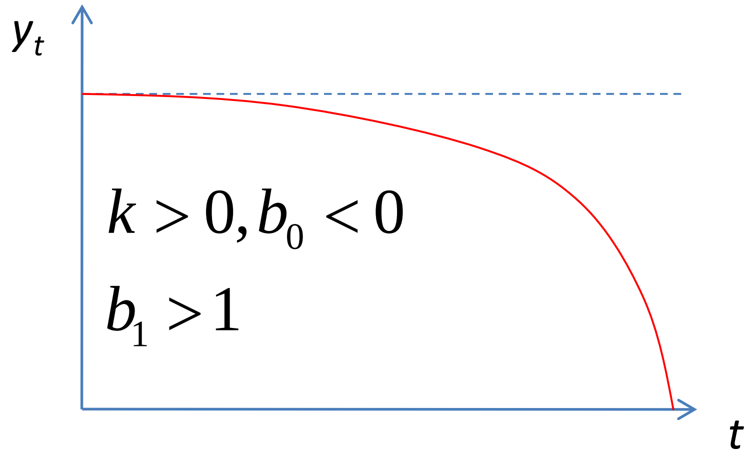
- Pro odhad průběhu trendu lze psát:

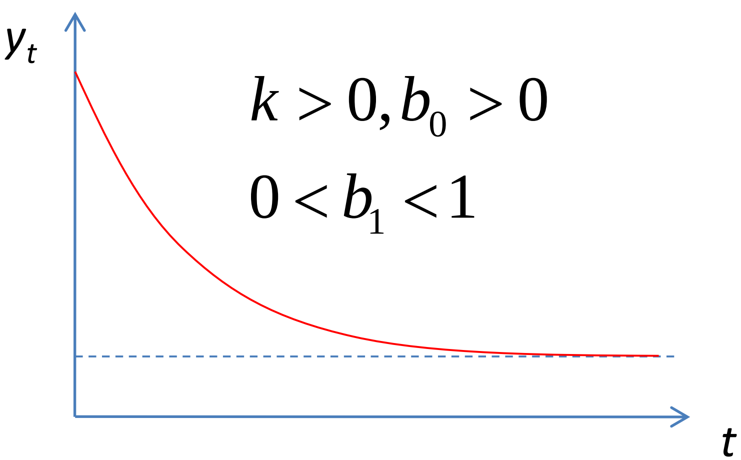
$\widehat{y}_t = k + b_0 \cdot b_1^t$ pro $b_1 > 0$.

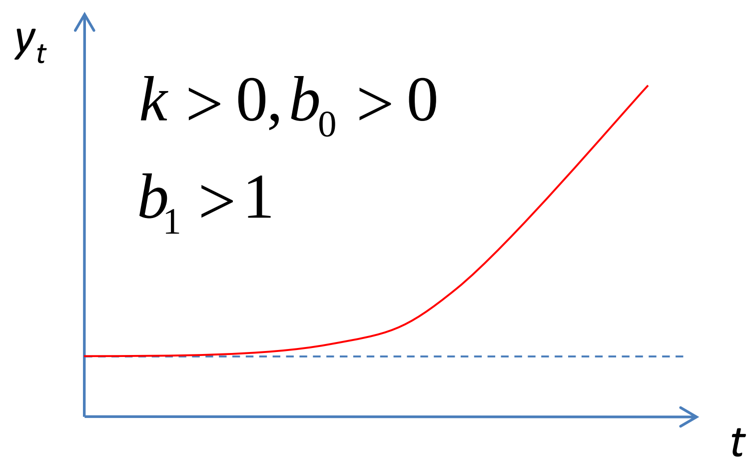
- Odhad parametrů je již složitější, protože trendovou funkci nemůžeme linearizovat pro použití metody nejmenších čtverců.

Analýza časových řad









Analýza časových řad

- Metody, které se používají pro odhad parametrů modifikované exponenciální trendové funkce, jsou:
  a) **Metoda částečných součtů**.
  b) **Metoda dílčích průměrů**.
  c) **Metoda vybraných bodů**.

Analýza časových řad

a) Metoda částečných součtů je založena na vytvoření tří na sebe navazujících a současně disjunktních součtů S_1, S_2 a S_3 o délce *m*, přičemž platí, že $n=3m$, kde *n* je počet pozorování. Platí tedy:

$$S_1 = \sum_{t=n-3m+1}^{n-2m} y_t \,, S_2 = \sum_{t=n-2m+1}^{n-m} y_t \,, S_3 = \sum_{t=n-m+1}^{n} y_t \,.$$

Analýza časových řad

- Není-li počet pozorování dělitelný 3, potom se vynechává potřebný počet pozorování na začátku časové řady – pro potřeby odhadu parametrů vynecháme prvních n-3m pozorování, zbývajícím pozorováním potom přiřadíme pořadí 1 až 3m.
- Nyní dosadíme do částečných součtů předpisy pro odhad modifikované exponenciální trendové funkce.

Analýza časových řad

$$S_1 = \sum_{t=n-3m+1}^{n-2m} y_t = \sum_{t=n-3m+1}^{n-2m} \left(k + b_0 \cdot b_1^t\right) = m \cdot k + b_0 \cdot \sum_{t=n-3m+1}^{n-2m} b_1^t{}^{*},$$

$$S_2 = \sum_{t=n-2m+1}^{n-m} y_t = \sum_{t=n-2m+1}^{n-m} \left(k + b_0 \cdot b_1^t\right) = m \cdot k + b_0 \cdot \sum_{t=n-2m+1}^{n-m} b_1^t{}^{*},$$

$$S_3 = \sum_{t=n-m+1}^{n} y_t = \sum_{t=n-m+1}^{n} \left(k + b_0 \cdot b_1^t\right) = m \cdot k + b_0 \cdot \sum_{t=n-m+1}^{n} b_1^t{}^{*}.$$

---

* Je třeba si uvědomit, že tyto výrazy reprezentují součet *m* členů geometrické posloupnosti. Pro tento součet obecně platí:

$$S_m = a_1 \cdot \frac{q^m - 1}{q - 1}.$$

Analýza časových řad

- Využijeme-li znalostí o součtu prvních *m* členů geometrické posloupnosti, dostaneme:

$$S_1 = m \cdot k + b_0 \cdot b_1^{n-3m+1} \cdot \frac{b_1^m - 1}{b_1 - 1} = m \cdot k + b_0 \cdot b_1 \cdot \frac{b_1^m - 1}{b_1 - 1},$$

$$S_2 = m \cdot k + b_0 \cdot b_1^{n-2m+1} \cdot \frac{b_1^m - 1}{b_1 - 1} = m \cdot k + b_0 \cdot b_1^{m+1} \cdot \frac{b_1^m - 1}{b_1 - 1},$$

$$S_3 = m \cdot k + b_0 \cdot b_1^{n-m+1} \cdot \frac{b_1^m - 1}{b_1 - 1} = m \cdot k + b_0 \cdot b_1^{2m+1} \cdot \frac{b_1^m - 1}{b_1 - 1}.$$

Analýza časových řad

- Máme soustavu 3 rovnic o 3 neznámých, kterými jsou odhady parametrů modifikovaného exponenciálního trendu.
- Vynásobme první rovnici (-1) a přičtěme ji k druhé rovnici. Po menších úpravách dostaneme:

$$S_2 - S_1 = b_0 \cdot \frac{b_1^m - 1}{b_1 - 1} \cdot \left(b_1^{m+1} - b_1\right) = b_0 \cdot \frac{b_1^m - 1}{b_1 - 1} \cdot b_1 \cdot \left(b_1^m - 1\right) =$$

$$= b_0 \cdot \frac{\left(b_1^m - 1\right)^2}{b_1 - 1} \cdot b_1 .$$

Analýza časových řad

- Z tohoto výrazu již můžeme vyjádřit:

$$b_0 = \frac{b_1 - 1}{b_1 \cdot \left(b_1^m - 1\right)^2} \cdot \left(S_2 - S_1\right).$$

- Nyní vynásobme druhou rovnici (-1) a přičtěme ji ke třetí rovnici. Po menších úpravách dostaneme:

$$S_3 - S_2 = b_0 \cdot \frac{b_1^m - 1}{b_1 - 1} \cdot \left(b_1^{2m+1} - b_1^{m+1}\right) = b_0 \cdot \frac{b_1^m - 1}{b_1 - 1} \cdot b_1^{m+1} \cdot \left(b_1^m - 1\right) =$$

$$= b_0 \cdot \frac{\left(b_1^m - 1\right)^2}{b_1 - 1} \cdot b_1^{m+1}.$$

Analýza časových řad

- Dosadíme-li do získaného výrazu vztah pro výpočet parametru b_0, dostaneme:

$$S_3 - S_2 = \frac{b_1 - 1}{b_1 \cdot \left(b_1^m - 1\right)^2} \cdot \left(S_2 - S_1\right) \cdot \frac{\left(b_1^m - 1\right)^2}{b_1 - 1} \cdot b_1^{m+1}.$$

- Po úpravách můžeme vyjádřit b_1 ve tvaru

$$b_1 = \left[\frac{\left(S_3 - S_2\right)}{\left(S_2 - S_1\right)}\right]^{\frac{1}{m}}.$$

Analýza časových řad

- Poslední parametr *k* můžeme potom vyjádřit např. z první rovnice. Dostaneme:

$$k = \frac{S_1 - b_0 \cdot b_1 \cdot \dfrac{b_1^m - 1}{b_1 - 1}}{m}.$$

Analýza časových řad

- **Př.:** Je zadána časová řada čítající 9 pozorování. Metodou částečných součtů proveďte odhad parametrů posunutého exponenciálního trendu.

t	y_t
1	3
2	10
3	15
4	21
5	35
6	42
7	58
8	81
9	110

Analýza časových řad

- Odvodili jsme si, že odhady parametrů získáme podle vztahů:

$$b_0 = \frac{b_1 - 1}{b_1 \cdot \left(b_1^m - 1\right)^2} \cdot \left(S_2 - S_1\right),$$

$$b_1 = \left[\frac{\left(S_3 - S_2\right)}{\left(S_2 - S_1\right)}\right]^{\frac{1}{m}},$$

a $k = \dfrac{S_1 - b_0 \cdot b_1 \cdot \dfrac{b_1^m - 1}{b_1 - 1}}{m}.$

Analýza časových řad

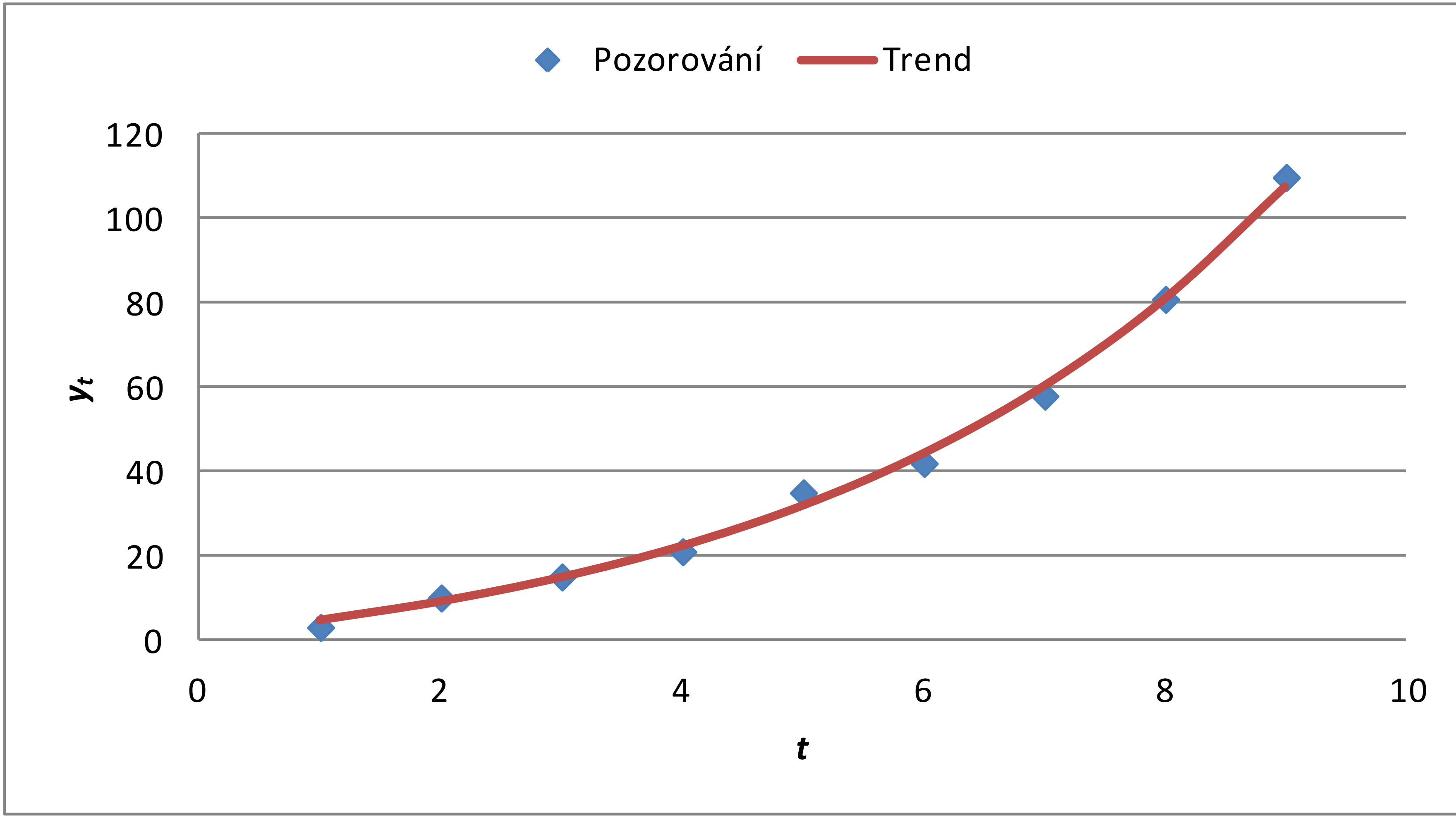
- Dosazením do těchto vztahů dostaneme následující výsledky.

m	3
S_1	28
S_2	98
S_3	249
$\boldsymbol{b_0}$	**11,82**
$\boldsymbol{b_1}$	**1,29**
$\boldsymbol{k}$	**-10,83**

t	y_t	$\hat{y}_t$
1	3	4,44
2	10	8,90
3	15	14,66
4	21	22,11
5	35	31,73
6	42	44,16
7	58	60,22
8	81	80,98
9	110	107,80

$$\hat{y}_t = -10{,}83 + 11{,}82 \cdot 1{,}29^t$$

Analýza časových řad



Analýza časových řad

b) Metoda dílčích průměrů je modifikací metody předchozí. Tato metoda zavádí dolní dílčí součet S_d, horní dílčí součet S_h a prostřední součet S_p. Pro první dva součty platí:

$$S_d = \frac{1}{5} \cdot \sum_{t=1}^{5} y_t \, ,$$

$$S_h = \frac{1}{5} \cdot \sum_{t=n-4}^{n} y_t \, .$$

Analýza časových řad

- Máme-li lichý počet pozorování *n*, potom pro prostřední součet platí:

$$S_p = \frac{1}{5} \cdot \sum_{t=\frac{(n+1)}{2}-2}^{\frac{(n+1)}{2}+2} y_t \,,$$

je-li *n* sudé, potom platí:

$$S_p = \frac{1}{6} \cdot \sum_{t=\frac{n}{2}-2}^{\frac{n}{2}+3} y_t \,.$$

Analýza časových řad

- Parametr trendu b_1 potom stanovíme dle vztahu:

$$b_1 = \left(\frac{S_h - S_p}{S_p - S_d} \right)^{\frac{2}{(n-5)}}$$

- Známe-li parametr trendu b_1, potom lze trendovou funkci považovat za lineární v parametrech a můžeme použít metodu nejmenších čtverců.

Analýza časových řad

- Můžeme tedy psát:

$$\varphi = \sum_{t=1}^{n} \left(y_t - k - b_0 \cdot b_1^t \right)^2 \rightarrow \min .$$

- Známým postupem dostaneme:

$$\sum_{t=1}^{n} y_t - n \cdot k - b_0 \cdot \sum_{t=1}^{n} b_1^t = 0 ,$$

$$\sum_{t=1}^{n} y_t \cdot b_1^t - k \cdot \sum_{t=1}^{n} b_1^t - b_0 \cdot \sum_{t=1}^{n} \left(b_1^t \right)^2 = 0 .$$

Analýza časových řad

- Z první rovnice můžeme vyjádřit b_0 ve tvaru:

$$b_0 = \frac{\sum_{t=1}^{n} y_t - n \cdot k}{\sum_{t=1}^{n} b_1^t}.$$

- Dosadíme-li tento výraz do druhé rovnice, dostaneme po úpravách:

$$k = \frac{\sum_{t=1}^{n} y_t \cdot b_1^t \cdot \sum_{t=1}^{n} b_1^t - \sum_{t=1}^{n} y_t \cdot \sum_{t=1}^{n} \left(b_1^t\right)^2}{\left(\sum_{t=1}^{n} b_1^t\right)^2 - n \cdot \sum_{t=1}^{n} \left(b_1^t\right)^2}.$$

Analýza časových řad

c) Metoda vybraných bodů je založena na výběru počátečního bodu *t*, který pro jednoduchost zpravidla volíme *t*=0. Další dva body volíme *t*+*m* a *t*+2*m*, kde opět platí, že:

$$m = \frac{n}{3}.$$

Analýza časových řad

- Parametry trendové funkce potom stanovíme dle vztahů:

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

$$b_0 = \frac{y_{t+m} - y_t}{b_1^{t+m} - 1},$$

$$k = y_t - b_0 \cdot b_1.$$

Analýza časových řad

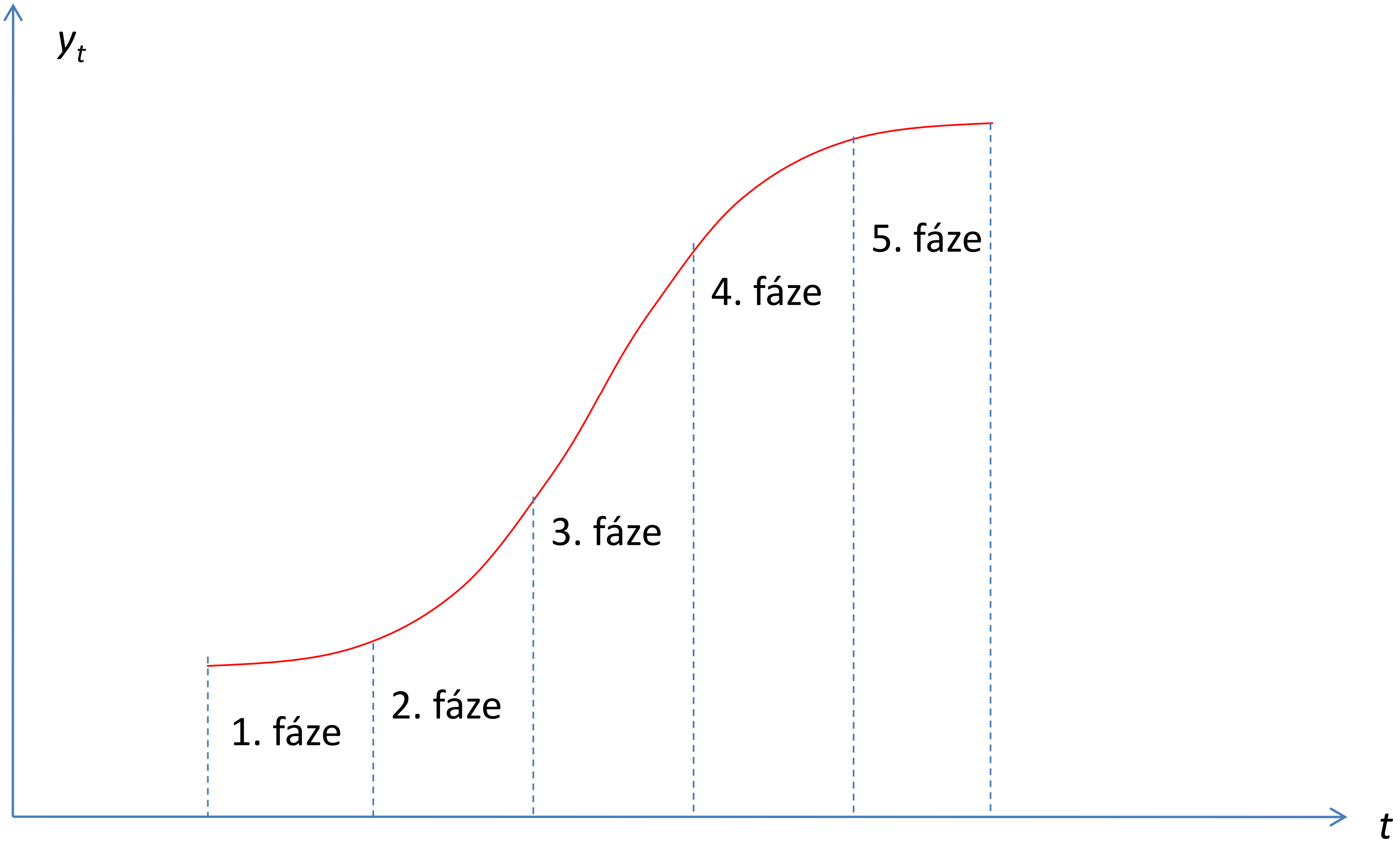
5) Logistický trend

- Pro odhad průběhu trendu lze psát:

$$\widehat{y}_t = \frac{k}{1 + b_0 \cdot b_1^{\,t}}.$$

- Tento funkční předpis je jeden z možných předpisů pro logistický trend.
- Logistická křivka se někdy také nazývá S-křivka.

Analýza časových řad



Analýza časových řad

- Průběh logistického trendu lze rozdělit do 5 fází:
  - 1. fáze: vznik nových výrobků a inovací, které se začínají pozvolna prosazovat (rozvoj je zpomalován existencí starých výrobků, které si zatím zachovávají svůj vliv).
  - 2. fáze: dochází k výraznému prosazení nových výrobků a inovací.

Analýza časových řad

- 3. fáze: nové výrobky a inovace plně ovládly další vývoj, nicméně dochází k náznaku změny trendu – objevují se novější výrobky a další inovace (v této fázi se nachází inflexní bod).
- 4. fáze: dochází k útlumu a postupnému nahrazování novějšími výrobky a inovacemi.
- 5. fáze: dochází k úplnému útlumu a nahrazení novějšími výrobky a inovacemi.

Analýza časových řad

- Odhad parametrů logistického trendu lze provést více způsoby a to:
  a) Metodou částečných součtů.
  b) Metodou vybraných bodů.

Analýza časových řad

a) Při odhadu parametrů metodou částečných součtů zavedeme pro t=1,2,...,n substituci:

$$\hat{x}_t = \frac{1}{\hat{y}_t}.$$

- Potom můžeme psát:

$$\hat{x}_t = \frac{1}{\dfrac{k}{1+b_0 \cdot b_1^t}} = \frac{1}{k} + \frac{b_0}{k} \cdot b_1^t,$$

což je zápis modifikované exponenciální trendové funkce.

Analýza časových řad

- Dále tedy postupujeme jako u odhadu parametrů pro modifikovaný exponenciální trend.
- Zaveďme následující značení:

$$K = \frac{1}{k},$$

$$B_0 = \frac{b_0}{k}.$$

Analýza časových řad

- Pro funkci logistického trendu můžeme tedy psát:

$$\hat{x}_t = K + B_0 \cdot b_1^t .$$

- Nyní vytvoříme 3 částečné součty, každý o délce *m* pozorování:

$$S_1 = \sum_{t=n-3m+1}^{n-2m} \frac{1}{y_t}, S_2 = \sum_{t=n-2m+1}^{n-m} \frac{1}{y_t}, S_3 = \sum_{t=n-m+1}^{n} \frac{1}{y_t}.$$

Analýza časových řad

- Známými vztahy spočítáme odhady koeficientů modelu:

$$B_0 = \frac{b_1 - 1}{b_1 \cdot \left(b_1^m - 1\right)^2} \cdot \left(S_2 - S_1\right),$$

$$b_1 = \left[\frac{\left(S_3 - S_2\right)}{\left(S_2 - S_1\right)}\right]^{\frac{1}{m}},$$

$$\text{a } K = \frac{S_1 - B_0 \cdot b_1 \cdot \dfrac{b_1^m - 1}{b_1 - 1}}{m}.$$

Analýza časových řad

- Odhady původních parametrů modelu potom získáme zpětnou transformací:

$$k = \frac{1}{K},$$

$$b_0 = k \cdot B_0 .$$

- Tento postup lze použít pouze tehdy, pokud je mají rozdíly S_2-S_1 a S_3-S_2 stejná znaménka.

Analýza časových řad

- **Př.:** V tabulce je dána časová řada vývoje stupně automobilizace v ČR v letech 1990 – 2008. Nalezněte odhad logistické trendové funkce popisující tento vývoj a odhadněte stupeň automobilizace v roce 2015.

Rok	t	Počet osobních vozidel na 1000 obyvatel
1990	1	233
1991	2	241
1992	3	250
1993	4	266
1994	5	283
1995	6	295
1996	7	309
1997	8	329
1998	9	339
1999	10	335
2000	11	335
2001	12	345
2002	13	358
2003	14	363
2004	15	374
2005	16	387
2006	17	399
2007	18	412
2008	19	424

Analýza časových řad

- Jelikož nemáme počet pozorování dělitelný 3, musíme vynechat 1. pozorování, potom budeme tvořit částečné součty o délce m=6 pozorování, pro potřeby těchto výpočtů zavedeme potřebnou transformaci a přečíslujeme si časovou proměnnou t.

Analýza časových řad

t	$1/y_t$
1	0,00415
2	0,00400
3	0,00376
4	0,00353
5	0,00339
6	0,00324
7	0,00304
8	0,00295
9	0,00299
10	0,00299
11	0,00290
12	0,00279
13	0,00275
14	0,00267
15	0,00258
16	0,00251
17	0,00243
18	0,00236

m	6
S_1	0,02207
S_2	0,01765
S_3	0,01530

Analýza časových řad

- Odvozenými vztahy spočítáme odhady parametrů pro substituovaný trend:

B_0	0,00224
b_1	0,89996
K	0,00211

$$B_0 = \frac{b_1 - 1}{b_1 \cdot \left(b_1^m - 1\right)^2} \cdot \left(S_2 - S_1\right),$$

$$b_1 = \left[\frac{\left(S_3 - S_2\right)}{\left(S_2 - S_1\right)}\right]^{\frac{1}{m}},$$

$$\text{a } K = \frac{S_1 - B_0 \cdot b_1 \cdot \frac{b_1^m - 1}{b_1 - 1}}{m}.$$

Analýza časových řad

- A nakonec zpětnou transformací získáme odhady parametrů logistického trendu:

b_0	1,06
b_1	0,90
k	474,53

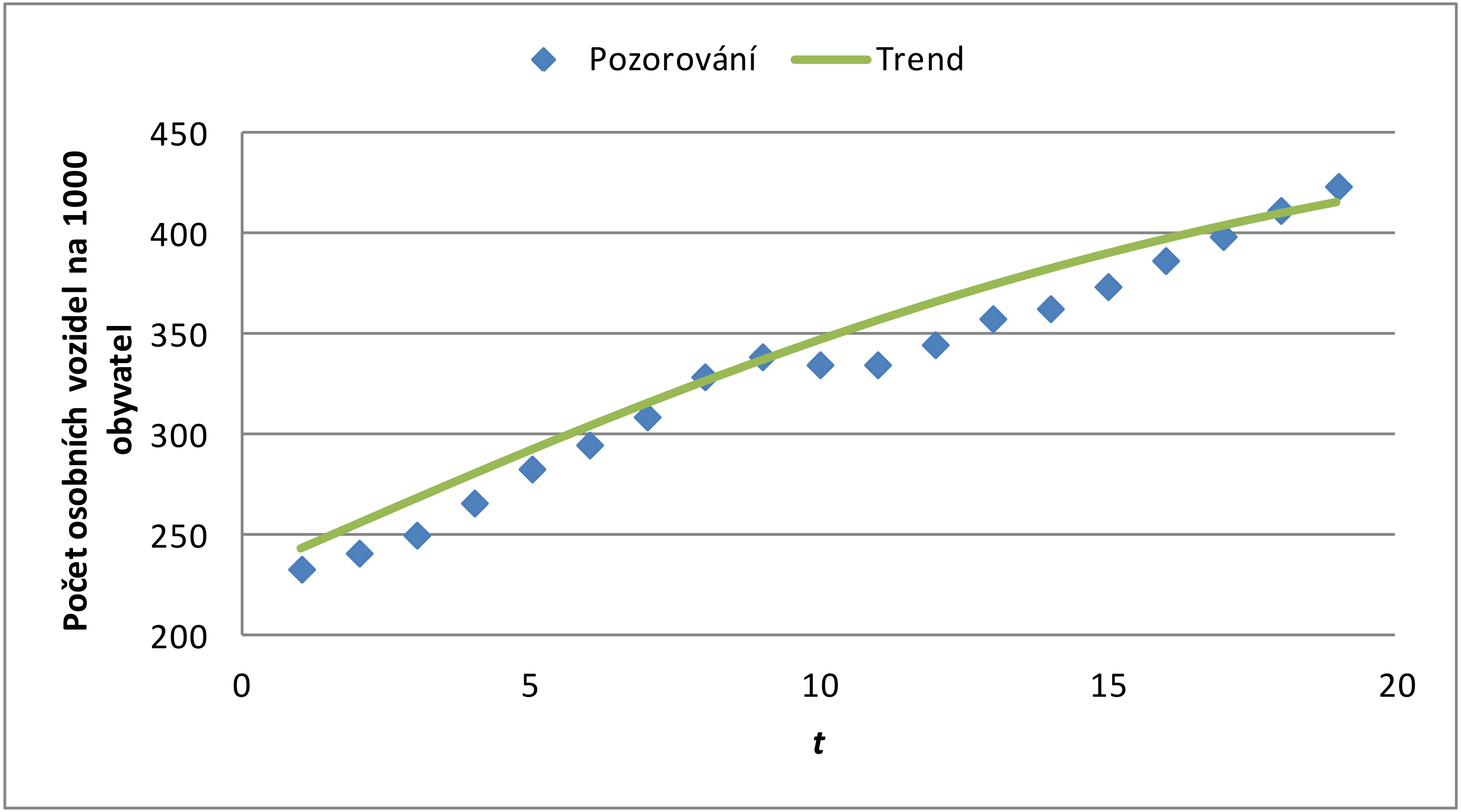
$$k = \frac{1}{K},$$

$$b_0 = k \cdot B_0 .$$

- Dostáváme rovnici trendu ve tvaru:

$$\hat{y}_t = \frac{474{,}53}{1 + 1{,}06 \cdot 0{,}9^t}.$$

Analýza časových řad



Analýza časových řad

- Odhad stupně automobilizace pro rok 2015 dostaneme dosazením za t=26:

$$\hat{y}_{26} = \frac{474{,}53}{1+1{,}06\cdot 0{,}9^{26}} = 444\,\text{automobilů na}\,1000\,\text{obyvatel.}$$

Analýza časových řad

b) Při metodě vybraných bodů opět vybereme počáteční bod *t*, který pro jednoduchost volíme *t*=0. Další dva body volíme *t*+*m* a *t*+2*m*.

- Dosazením do funkčního předpisu pro logistický trend dostaneme:

$$y_0 = \frac{k}{1+b_0}, y_m = \frac{k}{1+b_0 \cdot b_1^m}, y_{2m} = \frac{k}{1+b_0 \cdot b_1^{2m}}.$$

Analýza časových řad

- Z prvního vztahu vyjádříme b_0:

$$b_0 = \frac{k - y_0}{y_0}.$$

- Dosazením do druhého vztahu a následnými úpravami můžeme vyjádřit b_1 ve tvaru:

$$b_1 = \left[\frac{y_0 \cdot (k - y_m)}{y_m \cdot (k - y_0)}\right]^{\frac{1}{m}}.$$

Analýza časových řad

- Dosadíme-li oba parametry do poslední rovnice, můžeme vyjádřit parametr *k*:

$$k = \frac{2y_0 \cdot y_m \cdot y_{2m} - y_m^2 \cdot (y_0 + y_{2m})}{y_0 \cdot y_{2m} - y_m^2}.$$

Analýza časových řad

- Odhad parametrů logistického trendu metodou vybraných bodů lze realizovat i dalším způsobem. Uvažujme opět body t, $t+m$ a $t+2m$ (pro jednoduchost opět zvolíme $t=0$).

Analýza časových řad

- Zaveďme pomocné veličiny:

$$S_1 = \frac{1}{y_{t+2m}} - \frac{1}{y_{t+m}} = \frac{1 + b_0 \cdot b_1^{t+2m}}{k} - \frac{1 + b_0 \cdot b_1^{t+m}}{k},$$

$$S_2 = \frac{1}{y_{t+m}} - \frac{1}{y_t} = \frac{1 + b_0 \cdot b_1^{t+m}}{k} - \frac{1 + b_0 \cdot b_1^{t}}{k},$$

$$S_3 = \frac{y_{t+m}}{y_t} = \frac{1 + b_0 \cdot b_1^{t}}{1 + b_0 \cdot b_1^{t+m}}.$$

Analýza časových řad

- Dosadíme-li za t=0, můžeme psát:

$$S_1 = \frac{1+b_0 \cdot b_1^{2m}}{k} - \frac{1+b_0 \cdot b_1^{m}}{k} = b_1^m \cdot \left(\frac{b_0 \cdot b_1^m}{k} - \frac{b_0}{k} \right),$$

$$S_2 = \frac{1+b_0 \cdot b_1^{m}}{k} - \frac{1+b_0}{k} = \frac{b_0 \cdot b_1^m}{k} - \frac{b_0}{k},$$

$$S_3 = \frac{1+b_0}{1+b_0 \cdot b_1^m}.$$

Analýza časových řad

- Dosadíme-li do 1. rovnice 2. rovnici, dostaneme:

$$S_1 = b_1^m \cdot S_2 \,,$$

z čehož už snadno vyjádříme:

$$b_1 = \left(\frac{S_1}{S_2} \right)^{\frac{1}{m}} .$$

Analýza časových řad

- Dosadíme-li b_1 do 3. rovnice, můžeme psát:

$$S_3 = \frac{1+b_0}{1+b_0 \cdot \left[\left(\frac{S_1}{S_2}\right)^{\frac{1}{m}}\right]^m} = \frac{1+b_0}{1+b_0 \cdot \frac{S_1}{S_2}},$$

z čehož vyjádříme b_0:

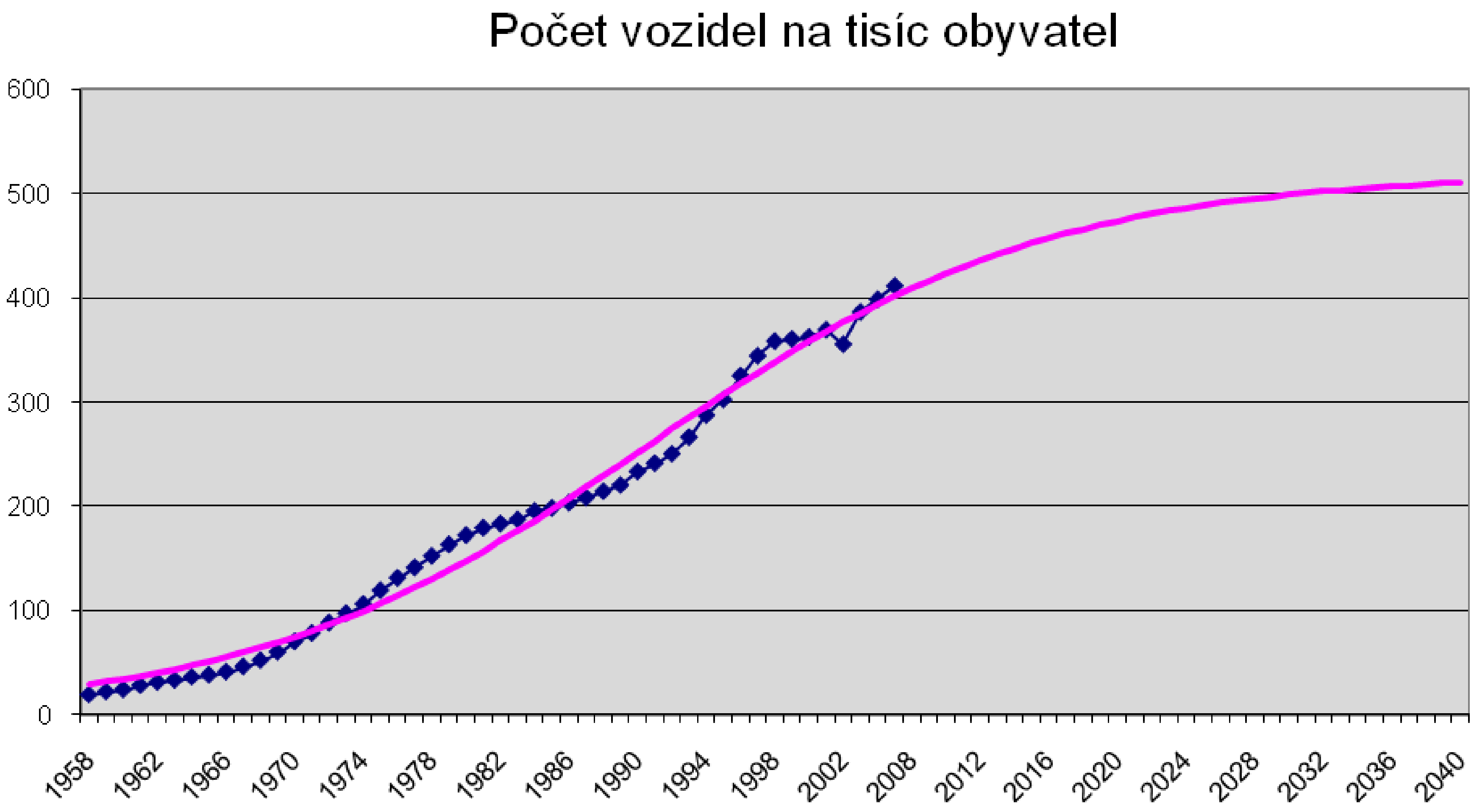
$$b_0 = \frac{S_3 - 1}{1 - \frac{S_1 \cdot S_3}{S_2}}.$$

Analýza časových řad

- Známe-li parametry b_0 a b_1, můžeme parametr k vyjádřit z funkčního předpisu logistické trendové funkce, kde dosadíme t=0:

$$y_0 = \frac{k}{1 + b_0 \cdot b_1^0} \Rightarrow k = y_0 \cdot \left(1 + b_0\right).$$

Analýza časových řad



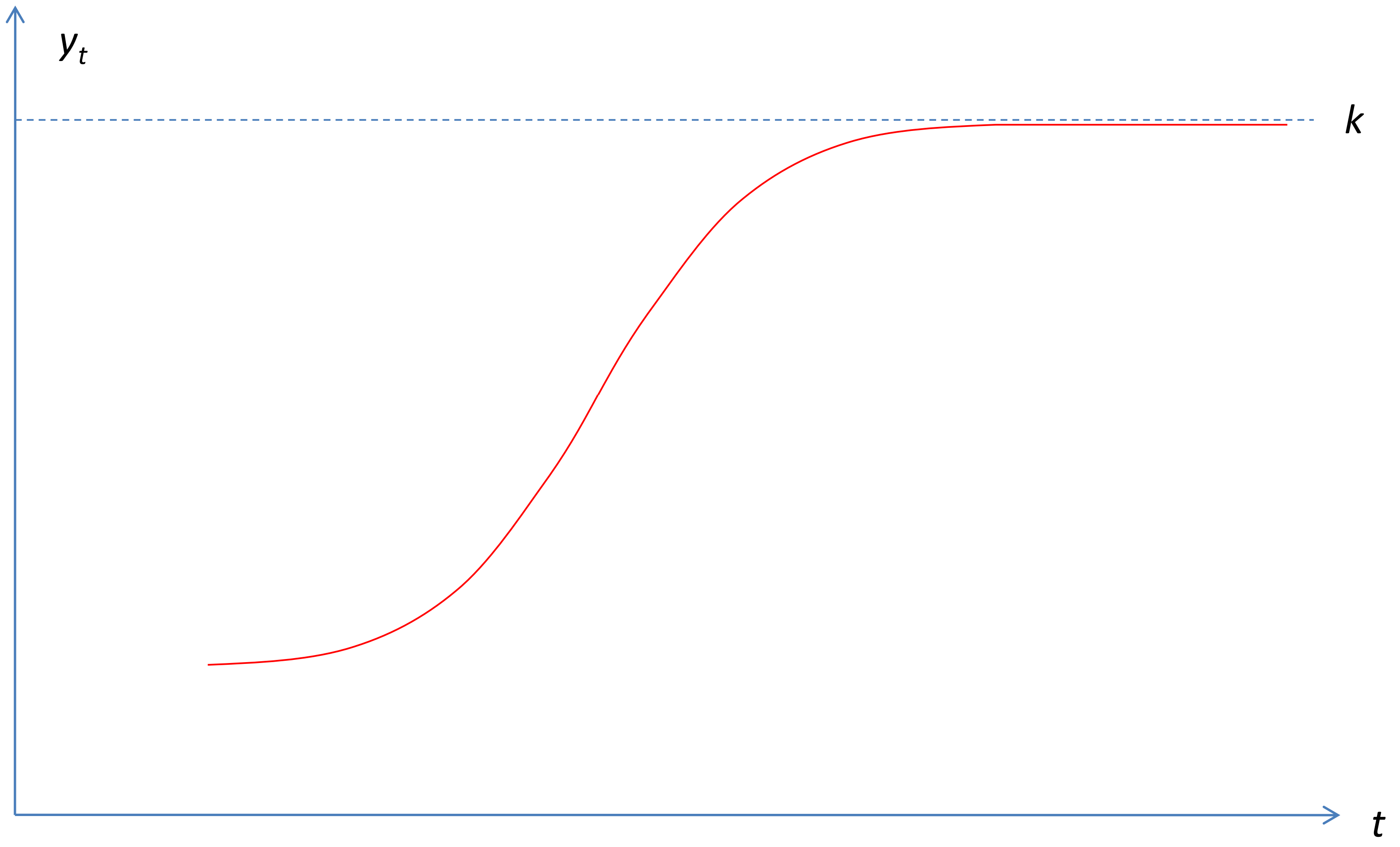
Analýza časových řad

6) Gompertzova křivka

- Má podobný průběh jako logistická křivka, ale není symetrická.
- Pro odhad průběhu trendu lze psát:

$$\widehat{y}_t = k \cdot b_0^{b_1^t}.$$

Analýza časových řad



Analýza časových řad

- Chceme-li provést odhad parametrů Gompertzovy funkce, zlogaritmováním převedeme funkční předpis do podoby:

$$\ln \hat{y}_t = \ln k \cdot b_0^{b_1^t} \,,$$

$$\ln \hat{y}_t = \ln k + b_1^t \cdot \ln b_0 \,.$$

- Touto úpravou jsme v podstatě získali funkci modifikovaného exponenciálního trendu.

Analýza časových řad

- Zavedeme-li substituce $K=\ln k$ a $B_0=\ln b_0$, můžeme parametry trendové funkce odhadnout stejnými metodami jako u modifikovaného exponenciálního trendu.

Analýza časových řad

- Nyní se zaměřme na to, na základě jakých kritérií zvolit vhodný typ trendu. Vhodný typ trendu lze volit:
  1. Na základě analýzy grafu studované časové řady (zda jde o rostoucí či klesající trend, zda přichází v úvahu inflexní bod, zda jde o funkci rostoucí do nekonečna nebo rostoucí k nějaké konečné limitě apod.)

Analýza časových řad

2) Dále lze vhodný typ trendové funkce vybrat na základě hodnoty reziduálního součtu čtverců, kdy z možných trendových funkcí vybereme tu s minimálním reziduálním součtem čtverců. Dalším kritériem může být index determinace známý z regresní analýzy, jako vhodný typ trendové funkce vybereme takový, u kterého je index determinace nejvyšší. Snahou je ale použít co nejjednodušší model trendové funkce.

Analýza časových řad

3) Rozhodujeme-li se mezi lineárním, parabolickým nebo exponenciálním trendem, lze použít **analýzu diferencí** časové řady. Diference příslušných řádů definujeme:

$$D1_t = y_t - y_{t-1} \text{ pro } t = 2,\ldots,n,$$
$$D2_t = D1_t - D1_{t-1} \text{ pro } t = 3,\ldots,n,$$
$$D3_t = D2_t - D2_{t-1} \text{ pro } t = 4,\ldots,n,$$

atd.

Analýza časových řad

– Pro lineární trend je typické, že diference prvního řádu jsou přibližně stejné a diference druhého řádu jsou přibližně nulové.

t	$y_t = b_0 + b_1 \cdot t$	$D1_t$	$D2_t$
1	$b_0 + b_1 \cdot 1$	-	-
2	$b_0 + b_1 \cdot 2$	b_1	-
3	$b_0 + b_1 \cdot 3$	b_1	0
4	$b_0 + b_1 \cdot 4$	b_1	0

Analýza časových řad

– Pro parabolický trend je typické, že diference prvního řádu vykazují lineární trend, diference druhého řádu jsou přibližně stejné a diference třetího řádu přibližně nulové.

t	$y_t = b_0 + b_1 \cdot t + b_2 \cdot t^2$	$D1_t$	$D2_t$	$D3_t$
1	$b_0 + b_1 \cdot 1 + b_2 \cdot 1^2$	-	-	-
2	$b_0 + b_1 \cdot 2 + b_2 \cdot 2^2$	$b_1 + b_2 \cdot 3$	-	-
3	$b_0 + b_1 \cdot 3 + b_2 \cdot 3^2$	$b_1 + b_2 \cdot 5$	$b_2 \cdot 2$	-
4	$b_0 + b_1 \cdot 4 + b_2 \cdot 4^2$	$b_1 + b_2 \cdot 7$	$b_2 \cdot 2$	0
5	$b_0 + b_1 \cdot 5 + b_2 \cdot 5^2$	$b_1 + b_2 \cdot 9$	$b_2 \cdot 2$	0

Analýza časových řad

– Na exponenciální trend budeme usuzovat na základě relativních diferencí prvního řádu definovaných podílem:

$$\frac{D1_t}{D1_{t-1}} \text{ pro t} = 3,\ldots,\text{n}$$

a nebo na základě temp růstu definovaných:

$$k_t = \frac{y_t}{y_{t-1}} \text{ pro } t = 2,\ldots,n.$$

Analýza časových řad

– Budou-li tyto charakteristiky kolísat kolem konstanty, lze pro popis trendové složky časové řady použít exponenciální trend.

Analýza časových řad

4) Vhodný typ trendové funkce lze provést na základě analýzy růstových charakteristik. Předpokladem je očištění časové řady od náhodných výkyvů a výpočet průměrných růstových charakteristik.

   Očištění časové řady od nahodilého kolísání se nejčastěji provádí pomocí lineárních filtrů, nejčastěji pomocí **klouzavých průměrů**.

Analýza časových řad

- Výpočet klouzavých průměrů:
  - Zvolme liché $m < n$, kde *n* je počet pozorování.
  - Postupně spočítáme průměr pro prvních *m* pozorování – tedy pro $y_1, y_2, \ldots, y_m$, pak pro dalších *m* pozorování – tedy pro $y_2, y_3, \ldots, y_{m+1}$ atd. Obecně můžeme psát:

$$\bar{y}_t = \frac{y_{t-p} + y_{t-p+1} + \ldots + y_{t+p}}{m},$$

pro $t = p+1, p+2, \ldots, n-p$, kde $m = 2p+1 \Rightarrow p = \frac{m-1}{2}$.

Analýza časových řad

- Hodnotu *p* volíme zpravidla 2, 3 nebo 4.
- Zaveďme dále průměrnou růstovou charakteristiku počítanou klouzavým způsobem z *m* pozorování:

$$\overline{\Delta}_t = \frac{-\frac{m-1}{2}\cdot y_{t-\frac{m-1}{2}} - \ldots - y_{t-1} + y_{t+1} + \ldots + \frac{m-1}{2}\cdot y_{t+\frac{m-1}{2}}}{\frac{m}{3}\cdot\frac{m-1}{2}\cdot\frac{m+1}{2}}$$

pro $t = p+1, p+2, \ldots, n-p$.

Analýza časových řad

- Pro $m = 5$ dostaneme:

$$\overline{\Delta}_t = \frac{-2y_{t-2} - y_{t-1} + y_{t+1} + 2y_{t+2}}{10}.$$

- Pro $m = 7$ dostaneme:

$$\overline{\Delta}_t = \frac{-3y_{t-3} - 2y_{t-2} - y_{t-1} + y_{t+1} + 2y_{t+2} + 3y_{t+3}}{28}.$$

- Pro $m = 9$ dostaneme:

$$\overline{\Delta}_t = \frac{-4y_{t-4} - 3y_{t-3} - 2y_{t-2} - y_{t-1} + y_{t+1} + 2y_{t+2} + 3y_{t+3} + 4y_{t+4}}{60}.$$

Analýza časových řad

- Vhodný typ trendu pak stanovíme na základě chování průměrných růstových a z nich odvozených charakteristik.

Analýza časových řad

Růstová charakteristika	Charakter změny růstové charakteristiky	Vhodný typ trendové funkce
$\overline{\Delta}_t$	Přibližně stejná	Lineární trend
$\overline{\Delta}_t$	Lineárně roste	Parabolický trend
$\dfrac{\overline{\Delta}_t}{\overline{y}_t}$	Přibližně stejná	Exponenciální trend
$\log \overline{\Delta}_t$	Lineárně klesá	Modifikovaný exponenciální trend
$\log\left(\dfrac{\overline{\Delta}_t}{\overline{y}_t}\right)$	Lineárně klesá	Gompertzova křivka
$\log\left(\dfrac{\overline{\Delta}_t}{\overline{y}_t^2}\right)$	Lineárně klesá	Logistický trend

Analýza časových řad

- **Př.:** V tabulce jsou uvedeny počty prodaných osobních automobilů značky X v tisících kusů za rok. Na základě rozboru růstových charakteristik vyberte vhodný typ trendové funkce.

Analýza časových řad

Rok	t	y_t	$\bar{y}_t$
1993	1	23	-
1994	2	38	-
1995	3	33	39,0
1996	4	50	45,8
1997	5	51	55,2
1998	6	57	66,2
1999	7	85	77,4
2000	8	88	95,2
2001	9	106	114,2
2002	10	140	137,2
2003	11	152	164,6
2004	12	200	-
2005	13	225	-

Studovanou časovou řadu nejprve očistíme od nahodilého kolísání pomocí pětičlenných klouzavých průměrů:

$$m = 2p + 1 = 5 \Rightarrow p = \frac{5-1}{2} = 2.$$

Pro výpočet klouzavých průměrů platí:

$$\bar{y}_t = \frac{y_{t-p} + y_{t-p+1} + \ldots + y_{t+p}}{m}.$$

Analýza časových řad

- Jednotlivé pětičlenné klouzavé průměry stanovíme:

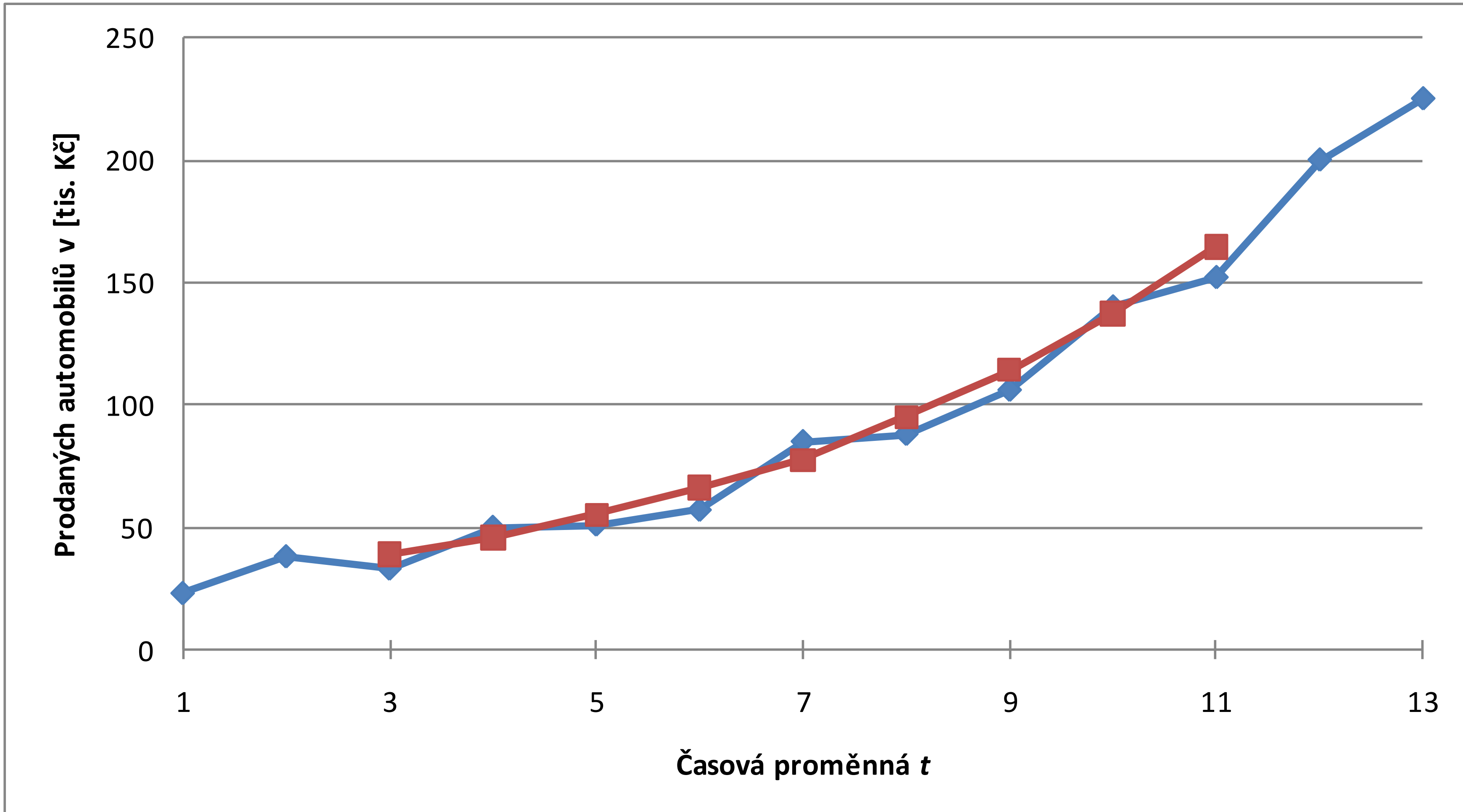
$$\bar{y}_3 = \frac{y_1 + y_2 + y_3 + y_4 + y_5}{5} = \frac{23 + 38 + 33 + 50 + 51}{5} = 39,$$

$$\bar{y}_4 = \frac{y_2 + y_3 + y_4 + y_5 + y_6}{5} = \frac{38 + 33 + 50 + 51 + 57}{5} = 45{,}8,$$

$$\vdots$$

$$\bar{y}_{11} = \frac{y_9 + y_{10} + y_{11} + y_{12} + y_{13}}{5} = \frac{106 + 140 + 152 + 200 + 225}{5} = 164{,}6.$$

Analýza časových řad



Analýza časových řad

- Nyní provedeme výpočet průměrných růstových charakteristik počítaných z 5 pozorování dle vztahu:

$$\overline{\Delta}_t = \frac{-2y_{t-2} - y_{t-1} + y_{t+1} + 2y_{t+2}}{10}.$$

Analýza časových řad

- Postupně dostaneme:

$$\overline{\Delta}_3 = \frac{-2y_1 - y_2 + y_4 + 2y_5}{10} = \frac{(-2)\cdot 23 - 38 + 50 + 2\cdot 51}{10} = 6{,}8,$$

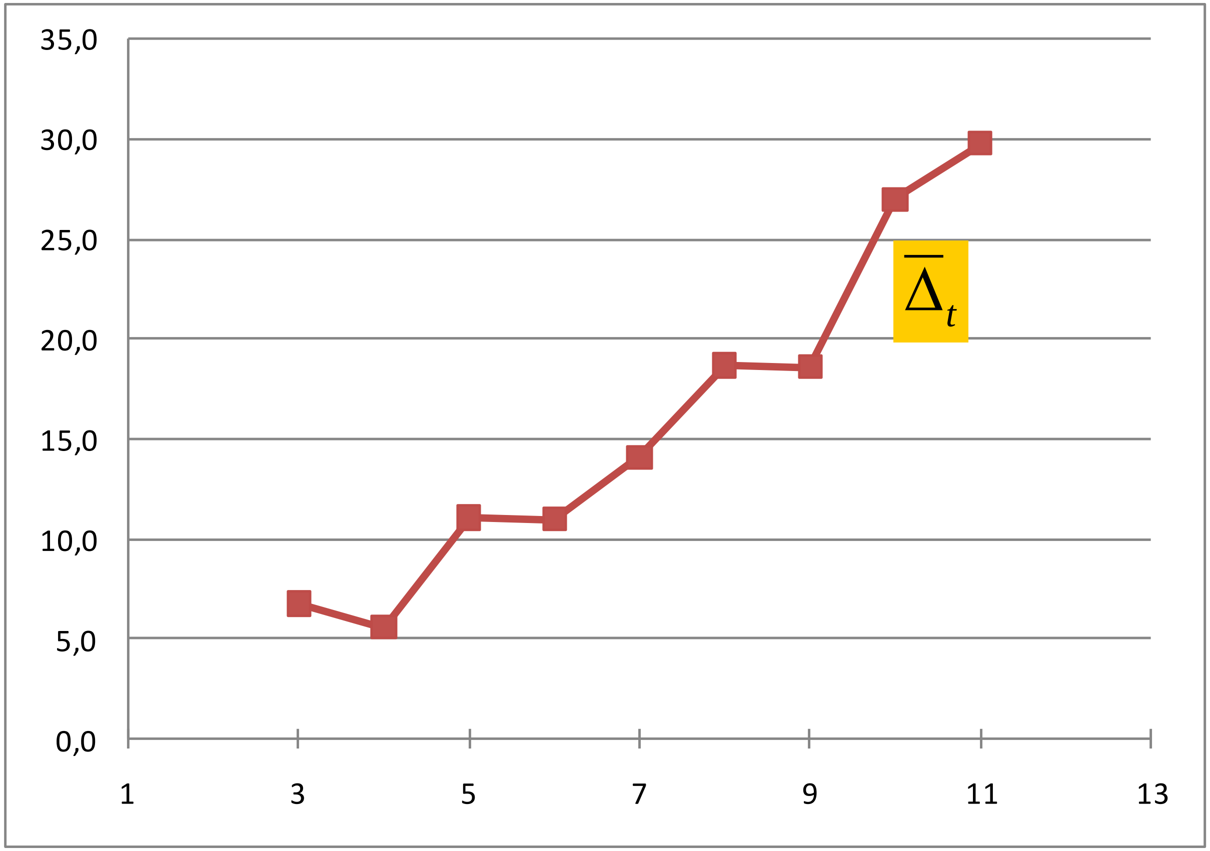
$$\overline{\Delta}_4 = \frac{-2y_2 - y_3 + y_5 + 2y_6}{10} = \frac{(-2)\cdot 38 - 33 + 51 + 2\cdot 57}{10} = 5{,}6,$$

$$\vdots$$

$$\overline{\Delta}_{11} = \frac{-2y_9 - y_{10} + y_{12} + 2y_{13}}{10} = \frac{(-2)\cdot 106 - 140 + 200 + 2\cdot 225}{10} = 29{,}8.$$

Analýza časových řad

t	y_t	$\bar{y}_t$	$\bar{\Delta}_t$
1	23	-	-
2	38	-	-
3	33	39,0	6,8
4	50	45,8	5,6
5	51	55,2	11,1
6	57	66,2	11,0
7	85	77,4	14,1
8	88	95,2	18,7
9	106	114,2	18,6
10	140	137,2	27,0
11	152	164,6	29,8
12	200	-	-
13	225	-	-



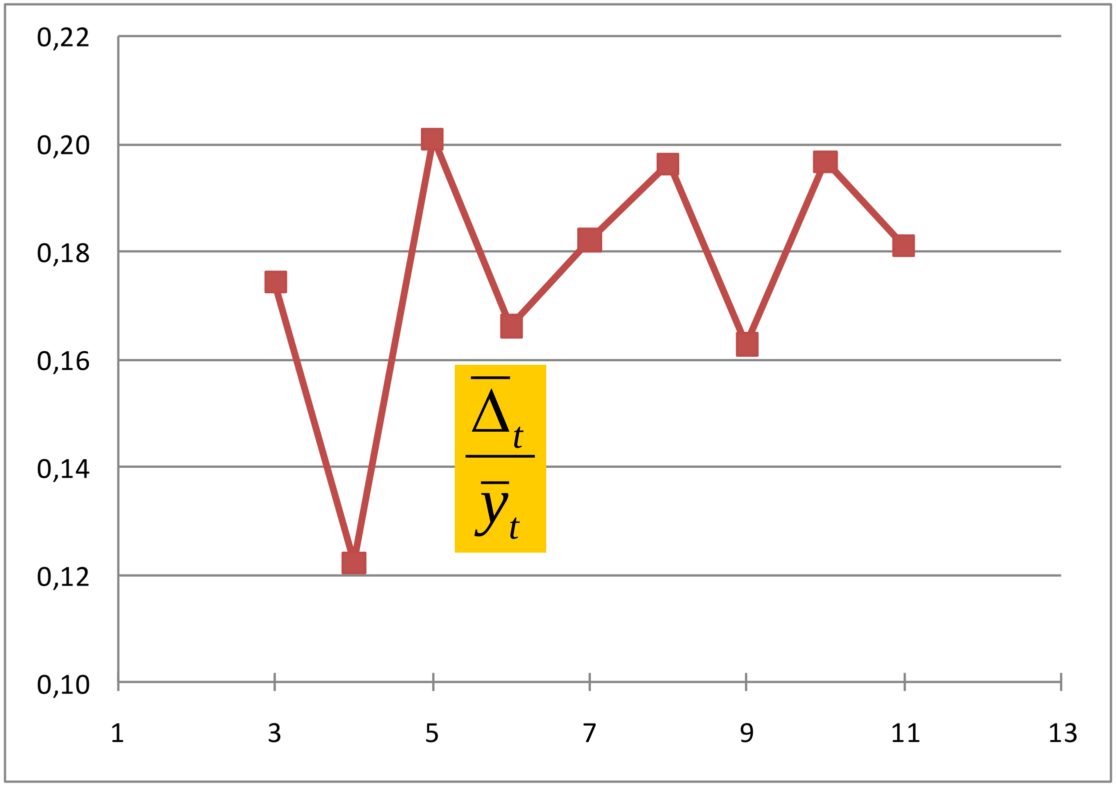
Analýza časových řad

- Z průběhu růstové charakteristiky vidíme, že její průběh zhruba lineárně roste, použití lineární trendové funkce se tedy nehodí, parabolická trendová funkce v úvahu přichází.

Analýza časových řad

t	y_t	$\bar{y}_t$	$\bar{\Delta}_t$	$\frac{\bar{\Delta}_t}{\bar{y}_t}$
1	23	-	-	-
2	38	-	-	-
3	33	39,0	6,8	0,17436
4	50	45,8	5,6	0,12227
5	51	55,2	11,1	0,20109
6	57	66,2	11,0	0,16616
7	85	77,4	14,1	0,18217
8	88	95,2	18,7	0,19643
9	106	114,2	18,6	0,16287
10	140	137,2	27,0	0,19679
11	152	164,6	29,8	0,18104
12	200	-	-	-
13	225	-	-	-

$$\frac{\bar{\Delta}_t}{\bar{y}_t}$$

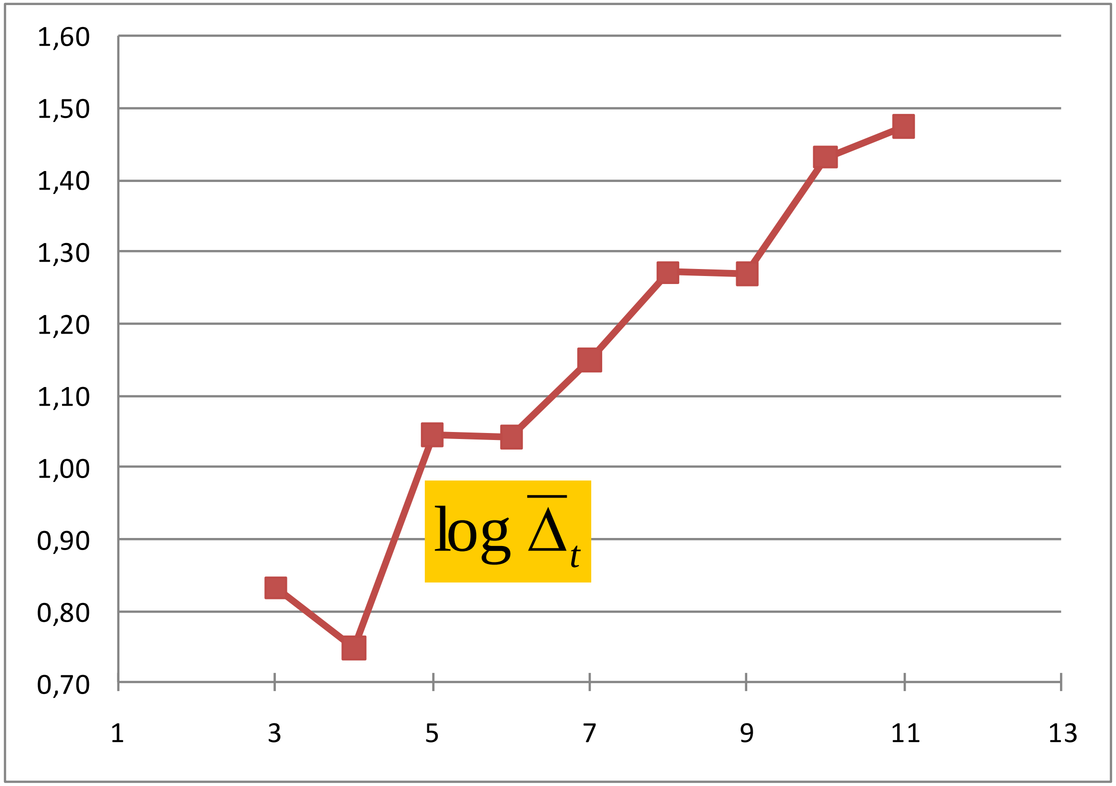


Analýza časových řad

- Z průběhu růstové charakteristiky vidíme, že její průběh zhruba kolísá kolem jedné hodnoty, použití exponenciálního trendu tedy přichází rovněž v úvahu.

Analýza časových řad

t	y_t	$\bar{y}_t$	$\bar{\Delta}_t$	$\log \bar{\Delta}_t$
1	23	-	-	-
2	38	-	-	-
3	33	39,0	6,8	0,83251
4	50	45,8	5,6	0,74819
5	51	55,2	11,1	1,04532
6	57	66,2	11,0	1,04139
7	85	77,4	14,1	1,14922
8	88	95,2	18,7	1,27184
9	106	114,2	18,6	1,26951
10	140	137,2	27,0	1,43136
11	152	164,6	29,8	1,47422
12	200	-	-	-
13	225	-	-	-

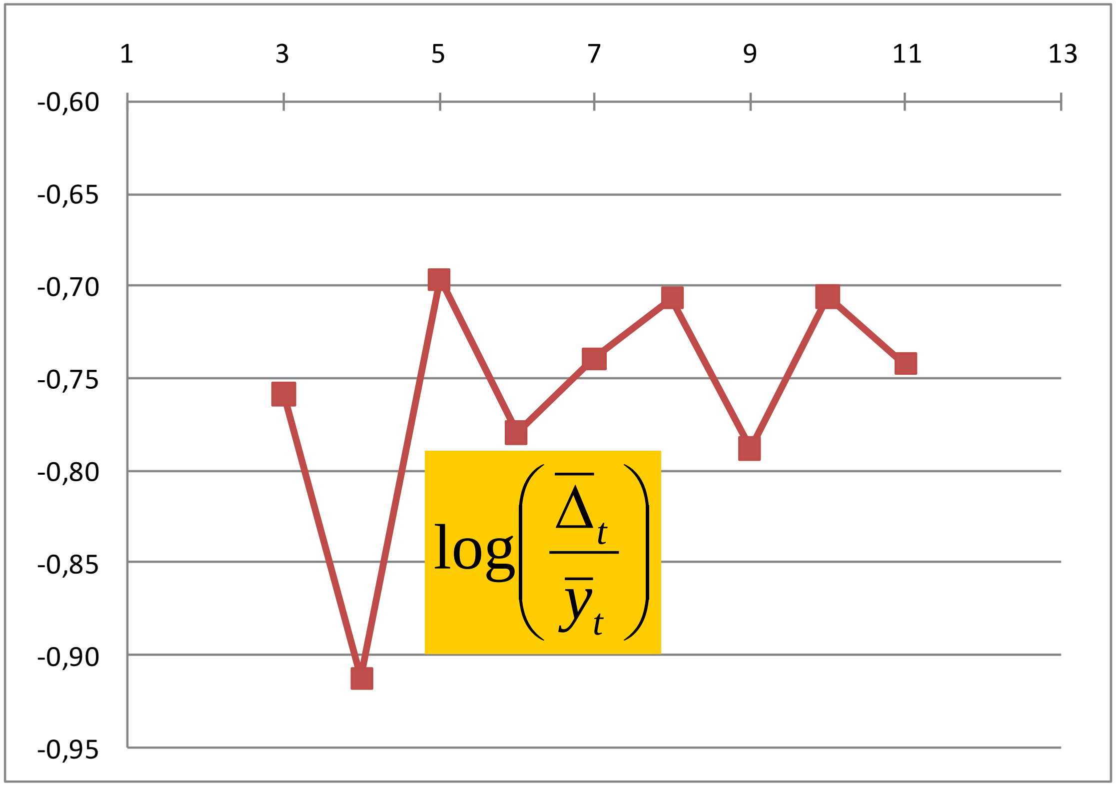


Analýza časových řad

- Jelikož růstová charakteristika v tomto případě neklesá lineárně, nýbrž naopak roste, lze modifikovaný exponenciální trend vyloučit.

Analýza časových řad

t	y_t	$\bar{y}_t$	$\bar{\Delta}_t$	$\log\left(\frac{\bar{\Delta}_t}{\bar{y}_t}\right)$
1	23	-	-	-
2	38	-	-	-
3	33	39,0	6,8	-0,75856
4	50	45,8	5,6	-0,91268
5	51	55,2	11,1	-0,69662
6	57	66,2	11,0	-0,77947
7	85	77,4	14,1	-0,73952
8	88	95,2	18,7	-0,70680
9	106	114,2	18,6	-0,78815
10	140	137,2	27,0	-0,70599
11	152	164,6	29,8	-0,74221
12	200	-	-	-
13	225	-	-	-



Analýza časových řad

- Jelikož růstová charakteristika v tomto případě lineárně neklesá, lze vyloučit použití Gompertzovy křivky.

Analýza časových řad

$$\log\left(\frac{\overline{\Delta}_t}{\bar{y}_t^2}\right)$$

t	y_t	$\bar{y}_t$	$\overline{\Delta}_t$	$\log\left(\frac{\overline{\Delta}_t}{\bar{y}_t^2}\right)$
1	23	-	-	-
2	38	-	-	-
3	33	39,0	6,8	-2,34962
4	50	45,8	5,6	-2,57354
5	51	55,2	11,1	-2,43856
6	57	66,2	11,0	-2,60032
7	85	77,4	14,1	-2,62826
8	88	95,2	18,7	-2,68543
9	106	114,2	18,6	-2,84582
10	140	137,2	27,0	-2,84334
11	152	164,6	29,8	-2,95864
12	200	-	-	-
13	225	-	-	-

Analýza časových řad

- Jelikož růstová charakteristika v tomto případě zhruba lineárně klesá, přichází rovněž v úvahu logistická trendová funkce.

- Analýzou růstových charakteristik jsme jako možné trendové funkce vybrali parabolický, exponenciální a logistický trend.

Analýza časových řad

- Vhodnou křivku bychom vybrali na základě toho, zda modelovaný jev může růst do nekonečna, zda existuje nějaká hranice nasycení apod.
- Dále bychom jako další kritérium mohli vzít součet čtverců reziduí, který chceme minimalizovat.