Regresní analýza trendu v časových řadách

Pojem časové řady: Časovou řadou rozumíme řadu hodnot $y_{t_1}, \ldots, y_{t_n}$ určitého ukazatele uspořádanou podle přirozené časové posloupnosti $t_1 < \ldots < t_n$. Jsou-li časové intervaly $(t_1, t_2), \ldots, (t_{n-1}, t_n)$ stejně dlouhé (ekvidistantní), zjednodušeně zapisujeme časovou řadu jako $y_1, \ldots, y_n$. Přitom ukazatel je veličina, která charakterizuje nějaký jev v určitém prostoru a určitém čase (okamžiku či intervalu).

Druhy časových řad

a) Časová řada okamžiková: příslušný ukazatel udává, kolik jevů existuje v daném časovém okamžiku (např. počet obyvatelstva k určitému dnu).

b) Časová řada intervalová: příslušný ukazatel udává, kolik jevů vzniklo či zaniklo v určitém časovém intervalu (např. počet sňatků během roku). Nejsou-li jednotlivé časové intervaly ekvidistantní, musíme provést očištění časové řady od důsledků kalendářních variací.

Příklad: Máme k dispozici údaje o tržbě obchodní organizace (v tis. Kč) v jednotlivých měsících roku 2013: 2400, 2134, 2407, 2445, 2894, 3354, 3515, 3515, 3225, 3063, 2694, 2600. Vypočtěte očištěné údaje.

Řešení: Průměrná délka měsíce je 365/12 dne. Očištěná hodnota

pro leden $y_1^{(o)} = 2400 \cdot \dfrac{365}{12 \cdot 31} = 2354{,}84$,

pro únor $y_2^{(o)} = 2134 \cdot \dfrac{365}{12 \cdot 28} = 2318{,}18$.

Pro ostatní měsíce analogicky dostaneme

2361,71; 2478,96; 2839,54; 3400,58, 3448,86; 3448,86; 3269,79; 3005,36; 2731,42; 2551,08.

Výpočet pomocí systému STATISTICA:

Vytvoříme nový datový soubor o třech proměnných: trzba, dm (délky jednotlivých měsíců) a ot (očištěná tržba) a 12 případech. Do proměnné trzba zapíšeme zjištěné hodnoty. Do proměnné dm vložíme délky jednotlivých měsíců, tj. 31, 28, 30, …, 31. Do Dlouhého jména proměnné ot napíšeme =trzba*365/(12*dm).

	1 trzba	2 dm	3 ot
1	2400	31	2354,839
2	2134	28	2318,185
3	2407	31	2361,707
4	2445	30	2478,958
5	2894	31	2839,543
6	3354	30	3400,583
7	3515	31	3448,858
8	3515	31	3448,858
9	3225	30	3269,792
10	3063	31	3005,363
11	2694	30	2731,417
12	2600	31	2551,075

Grafické znázornění okamžikové časové řady

Použijeme spojnicový diagram. Na vodorovnou osu vynášíme časové okamžiky $t_1, \dots, t_n$, na svislou osu odpovídající hodnoty $y_1, \dots, y_n$. Dvojice bodů (t_i, y_i), $i = 1, \dots, n$ spojíme úsečkami.

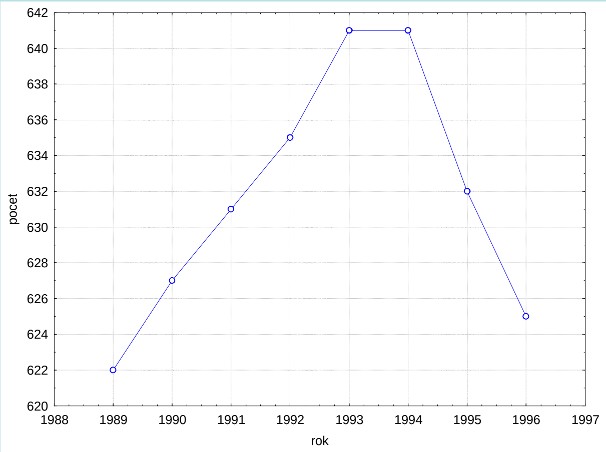
Příklad: Časová řada obsahuje údaje o počtu zaměstnanců určité akciové společnosti v letech 1989 – 1996 vždy k 31.12.

1989	1990	1991	1992	1993	1994	1995	1996
622	627	631	635	641	641	632	625

Znázorněte tuto časovou řadu graficky.

Řešení pomocí systému STATISTICA:

Vytvoříme datový soubor o dvou proměnných nazvaných rok a pocet a 8 případech.
Grafy – Bodové grafy – odškrtneme Lineární proložení – Proměnné X – rok, Y – počet – OK – OK. 2x klikneme na pozadí grafu – vybereme Graf: obecné – zaškrtneme Spojnice – OK.



Grafické znázornění intervalové časové řady

Použijeme sloupkový diagram. Je to soustava obdélníků, kde šířka obdélníku je rovna délce intervalu a výška odpovídá hodnotě ukazatele v daném intervalu. Ke znázornění intervalové časové řady lze použít i spojnicový diagram, přičemž na vodorovnou osu vynášíme středy příslušných intervalů.

Příklad: Máme k dispozici údaje o produkci určitého podniku (v tisících výrobků) v letech 1991-1996.

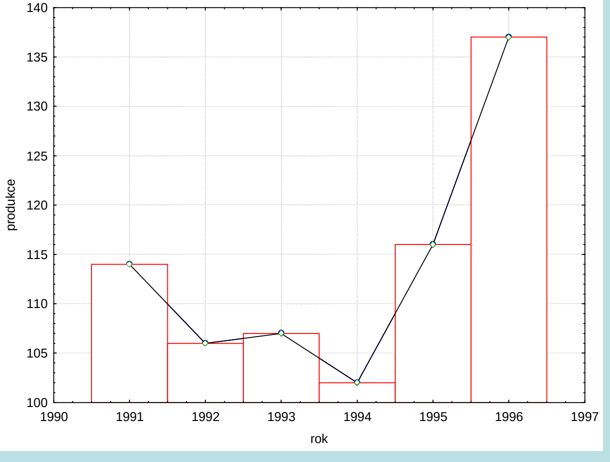
1991	1992	1993	1994	1995	1996
114	106	107	102	116	137

Znázorněte tuto časovou řadu graficky.

Řešení pomocí systému STATISTICA:

Vytvoříme datový soubor o dvou proměnných nazvaných rok a produkce a 6 případech.

Grafy – Bodové grafy – odškrtneme Lineární proložení – Proměnné X – rok, Y – produkce – OK – OK. 2x klikneme na pozadí grafu – vybereme Graf: obecné – zaškrtneme Spojnice – Přidat nový graf – typ Sloupcový graf – OK. Do sloupců označených jako Nový1, Nový2 okopírujeme hodnoty proměnných rok a produkce. Ve Všech možnostech: Sloupce upravíme šířku sloupce na 1.



Aditivní model časové řady

Předpokládejme, že pro časovou řadu $y_1, ..., y_n$ platí model

$y_t = f(t) + \varepsilon_t$, $t = 1, ..., n$, kde

$f(t)$ je neznámá trendová funkce (trend), kterou považujeme za systematickou (deterministickou) složku časové řady (popisuje hlavní tendenci dlouhodobého vývoje časové řady),

ε_t je náhodná složka časové řady zahrnující odchylky od trendu. Náhodná složka splňuje předpoklady

$E(\varepsilon_t) = 0$,

$D(\varepsilon_t) = \sigma^2$,

$C(\varepsilon_t, \varepsilon_{t+h}) = 0$,

$\varepsilon_t \sim N(0, \sigma^2)$ (říkáme, že ε_t je bílý šum).

Odhad trendu časové řady pomocí klouzavých průměrů

Podstata klouzavých průměrů

Předpokládáme, že časová řada se řídí aditivním modelem

$y_t = f(t) + \varepsilon_t$, $t = 1, ..., n$.

Odhad trendu v bodě t získáme určitým zprůměrováním původních pozorování z jistého okolí uvažovaného časového okamžiku t. Můžeme si představit, že podél dané časové řady klouže okénko, v jehož rámci se průměruje.

Nejprve musíme zvolit délku okna h > 1, v němž se bude počítat průměrná hodnota h pozorovaných hodnot y_i, $i = z, z+1, \dots, z+h-1$ v po sobě jdoucích časech $t = t_z$, $t = t_{z+1}$, …, $t = t_{z+h-1}$. Hodnota z představuje začátek okna.

a) Nechť šířka h vyhlazovacího okénka je liché číslo: h = 2d + 1. Pak $\hat{y}_{z+d} = \frac{1}{2d+1}\sum_{i=1}^{2d+1} y_{z+i-1}$ pro z = 1, …, n – 2d. Vypočtený průměr se přiřadí časovému okamžiku, který odpovídá středu okna.

b) Nechť šířka h vyhlazovacího okénka je sudé číslo: h = 2d. Pak $\hat{y}_{z+d} = \frac{1}{4d}\left(\sum_{i=1}^{2d} y_{z+i-1} + \sum_{i=1}^{2d} y_{z+i}\right)$ pro z = 1, …, n – 2d. Odhad trendu v bodě t se vypočte jako průměr ze dvou klouzavých průměrů, které příslušejí časovým okamžikům nejblíže vlevo a vpravo od středu okna. Tento průměr se přiřadí nejblíže většímu okamžiku od středu okna.

Hodnoty centrovaného klouzavého průměru a trendové funkce $\widehat{f}(t)$ jsou definovány v časech $t = t_{d+1}$, $t = t_{d+2}$, …, $t = t_{n-d}$.

Šířka vyhlazovacího okénka

Velmi důležitou otázkou je stanovení šířky vyhlazovacího okénka. Je-li okénko příliš široké, bude se odhad trendu blížit přímce (říkáme, že je přehlazen) a zároveň se ztratí velký počet členů na začátku a na konci časové řady. Je-li naopak okénko úzké, bude se odhad trendu blížit původním hodnotám (říkáme, že odhad je podhlazen). Nejčastěji se volí šířka okénka h = 3, 5, 7, pro čtvrtletní hodnoty pak 4.

Příklad: Časová řada -4,483, -39,534, -99,57, -152,99, -150,45, -80,239, -64,413, -120,825, -117,415, -70,811, -69,793, -26,438, 38,624, 39,761, 87,915, 67,246, 149,587, 121,241, 191,128, 305,71, 351,22 udává bilanci zahraničního obchodu ČR (v miliardách Kč) v letech 1993 až 2013.

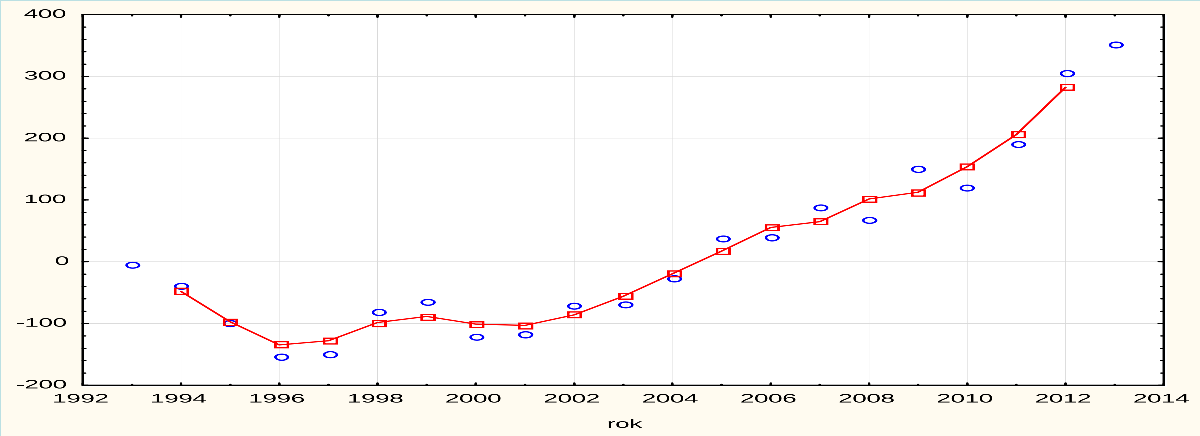
a) Odhadněte trend této časové řady pomocí klouzavých průměrů s vyhlazovacím okénkem šířky 3 a poté 5.
b) Graficky znázorněte průběh časové řady s odhadnutým trendem.

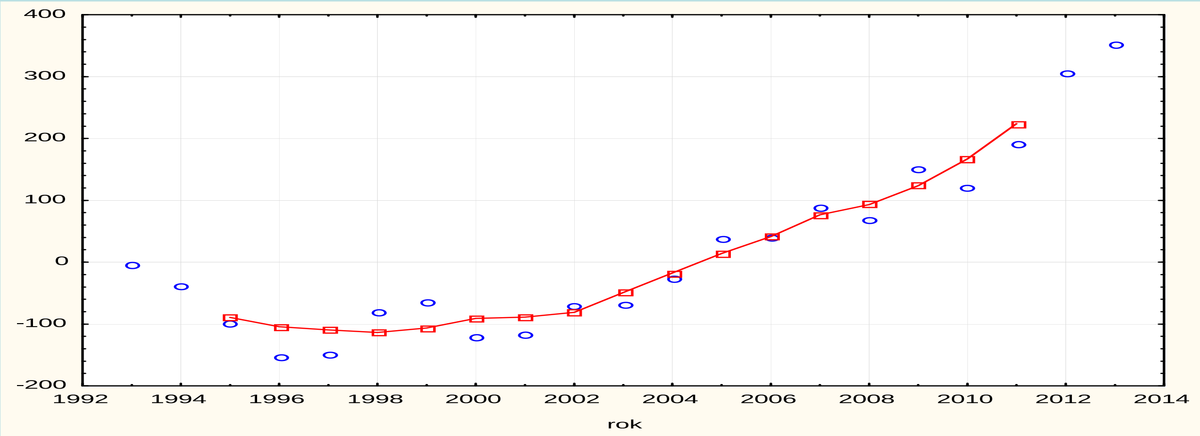
Řešení pomocí systému STATISTICA:

Načteme datový soubor bilance.sta o dvou proměnných rok a bilance a 21 případech.
Statistiky – Pokročilé lineární/nelineární modely – Časové řady/predikce – Proměnné bilance – OK– OK (transformace, autokorelace, kříž. korelace, grafy) – Vyhlazování – zaškrtneme N-bod. klouzavý průměr, N = 3 – OK (Transformovat vybrané řady) – vykreslí se graf, vrátíme se do Transformace proměnných – Uložit proměnné. Otevře se nový spreadsheet, kde v proměnné bilance_1 jsou uloženy klouzavé průměry pro N = 3. Totéž uděláme pro případ N = 5. Ve spreadsheetu se proměnná bilance_1 přepíše na bilance_2 a nová proměnná se uloží jako bilance_1. Nově vzniklé proměnné nazveme KP3 a KP5. K datovému souboru přidáme proměnnou rok, do jejíhož Dlouhého jména napíšeme =1992+v0.

	bilance.sta			
	1 rok	2 BILANCE	3 KP3	4 KP5
1	1993	-4,483		
2	1994	-39,534	-47,862	
3	1995	-99,570	-97,365	-89,405
4	1996	-152,990	-134,337	-104,557
5	1997	-150,450	-127,893	-109,532
6	1998	-80,239	-98,367	-113,783
7	1999	-64,413	-88,492	-106,668
8	2000	-120,825	-100,884	-90,741
9	2001	-117,415	-103,017	-88,651
10	2002	-70,811	-86,006	-81,056
11	2003	-69,793	-55,681	-49,167
12	2004	-26,438	-19,202	-17,731
13	2005	38,624	17,316	14,014
14	2006	39,761	55,433	41,422
15	2007	87,915	64,974	76,627
16	2008	67,246	101,583	93,150
17	2009	149,587	112,691	123,423
18	2010	121,241	153,985	166,982
19	2011	191,128	206,026	223,777
20	2012	305,710	282,686	
21	2013	351,220		

Grafické znázornění časové řady s odhadnutým trendem provedeme pomocí vícenásobných bodových grafů.





Cíl regresní analýzy trendu

Regresní analýza trendu má objasnit vztah mezi závisle proměnnou veličinou Y a časem t.

Předpokládáme, že trend f(t) závisí (lineárně či nelineárně) na neznámých parametrech $\beta_0, \beta_1, ..., \beta_k$ a známých funkcích $\varphi_0(t), \varphi_1(t), ..., \varphi_k(t)$, které již neobsahují žádné neznámé parametry, tj.

$f(t) = g(\beta_0, \beta_1, ..., \beta_k; \varphi_0(t), \varphi_1(t), ..., \varphi_k(t))$.

Odhady $b_0, b_1, ..., b_k$ neznámých parametrů $\beta_0, \beta_1, ..., \beta_k$ lze získat např. metodou nejmenších čtverců a pak vyjádřit odhad $\hat{f}(t)$ neznámého trendu v bodě t pomocí odhadů $b_0, b_1, ..., b_k$ a funkcí $\varphi_0(t), \varphi_1(t), ..., \varphi_k(t)$, tj.

$\hat{f}(t) = g(b_0, b_1, ..., b_k; \varphi_0(t), \varphi_1(t), ..., \varphi_k(t))$.

Nejdůležitější typy trendových funkcí

Volba typu trendové funkce se provádí

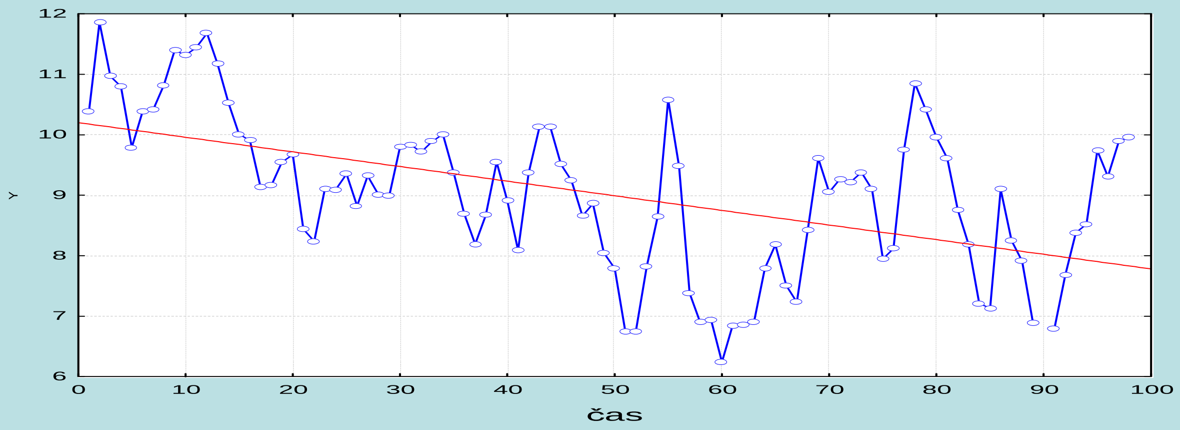
- na základě teoretických znalostí a zkušeností se zkoumanou veličinou Y_t
- pomocí grafu časové řady
- pomocí informativních testů založených na jednoduchých charakteristikách časové řady

a) Lineární trend

Analytické vyjádření: $f(t) = \beta_0 + \beta_1 t$

Informativní test: 1. diference ($\Delta y_t = y_t - y_{t-1}, t = 2, \ldots, n$) jsou přibližně konstantní.

Příklad lineárního trendu:

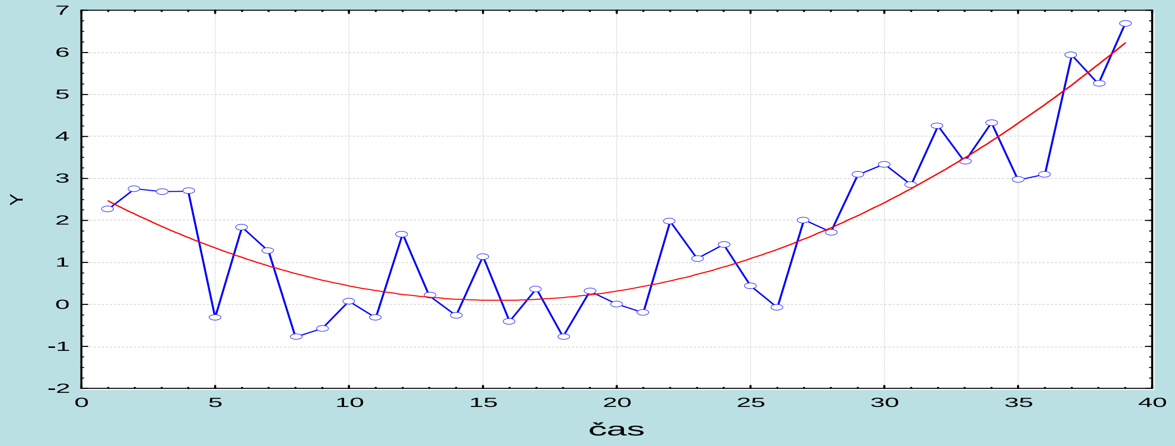


b) Kvadratický trend

Analytické vyjádření: $f(t) = \beta_0 + \beta_1 t + \beta_2 t^2$

Informativní test: 1. diference mají přibližně lineární trend, 2. diference ($\Delta^{(2)} y_t = \Delta y_t - \Delta y_{t-1} = y_t - 2y_{t-1} + y_{t-2}, t = 3, \ldots, n$) jsou přibližně konstantní.

Příklad kvadratického trendu:

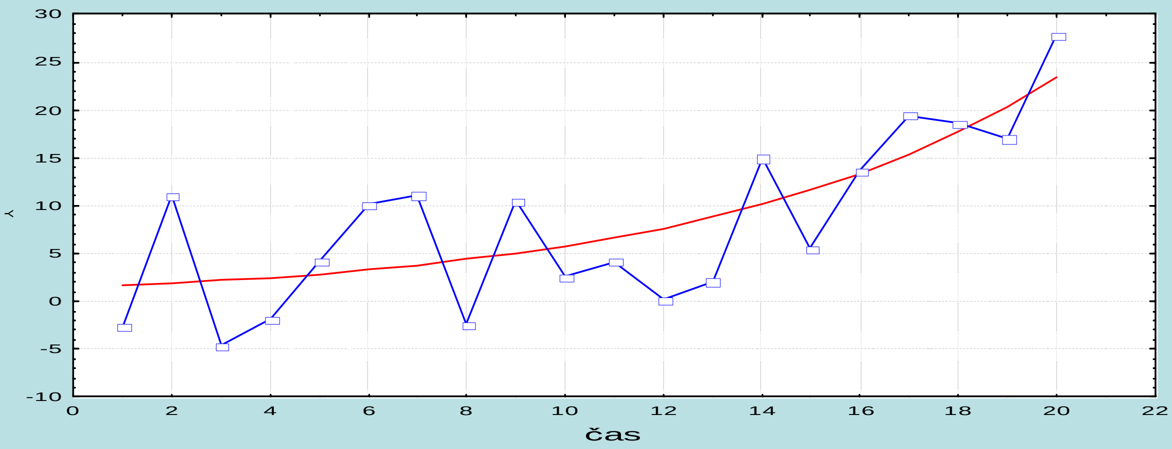


c) Exponenciální trend

Analytické vyjádření: $f(t) = \beta_0 \beta_1^{t}$.

Informativní test: koeficienty růstu $\left(k_t = \dfrac{y_t}{y_{t-1}}, t = 2, \ldots, n\right)$ jsou přibližně konstantní.

Příklad exponenciálního trendu:

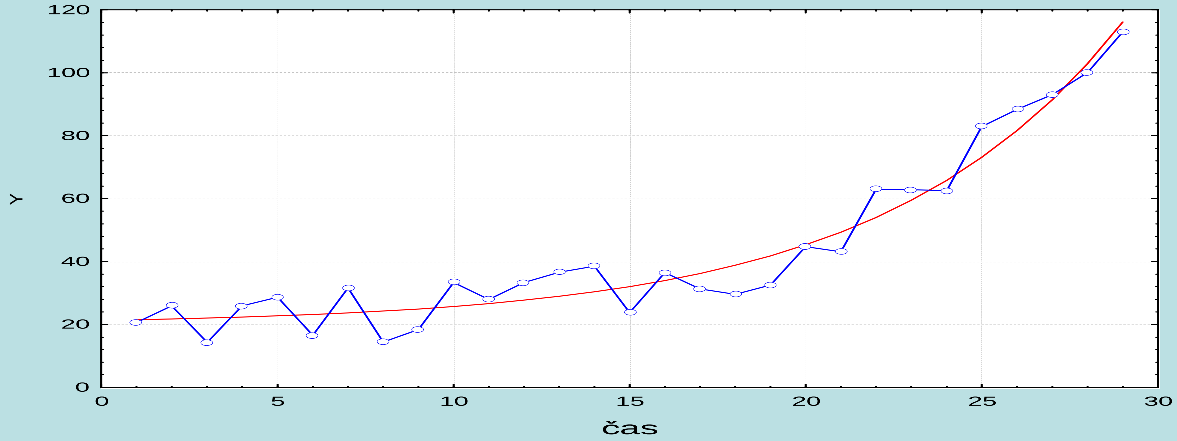


d) Modifikovaný exponenciální trend

Analytické vyjádření: $f(t) = \alpha + \beta_0 \beta_1^{\,t}$.

Informativní test: řada podílů sousedních 1. diferencí je přibližně konstantní.

Příklad modifikovaného exponenciálního trendu

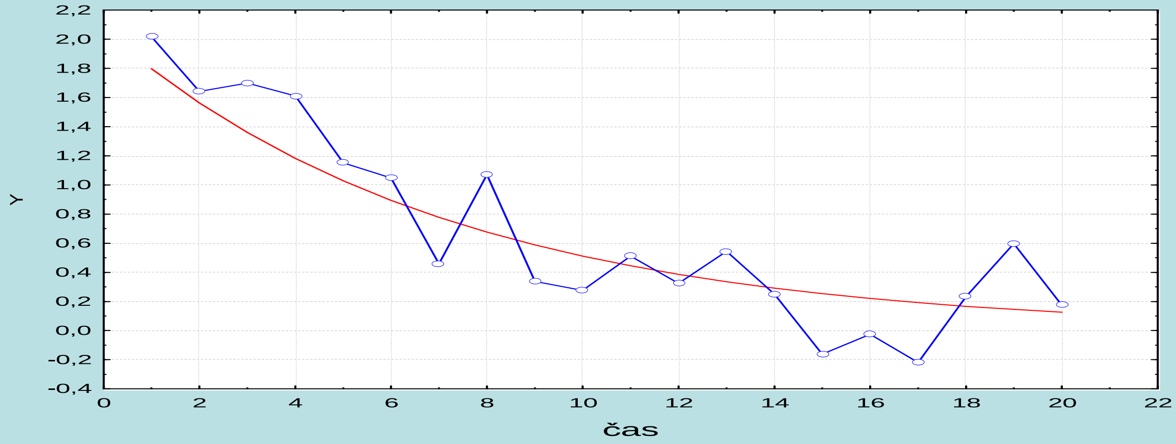


e) Logistický trend

Analytické vyjádření: $f(t) = \dfrac{\alpha}{1 + \beta_0 \beta_1^{\,t}}$

Informativní test: průběh 1. diferencí je podobný Gaussově křivce a podíly $\dfrac{1/y_{t+2} - 1/y_{t+1}}{1/y_{t+1} - 1/y_t}$ jsou přibližně konstantní.

Příklad logistického trendu:

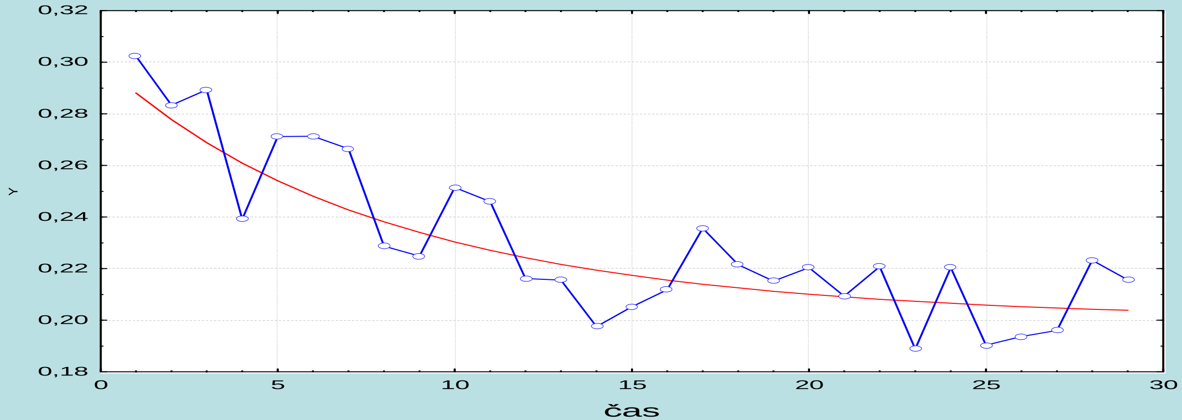


f) Gompertzova křivka

Analytické vyjádření: $f(t) = \alpha \beta_0^{\beta_1^t}$

Informativní test: podíly $\dfrac{\ln y_{t+2} - \ln y_{t+1}}{\ln y_{t+1} - \ln y_t}$ jsou přibližně konstantní.

Příklad Gompertzovy křivky



Příklad:
Je dána časová řada potratů (v tisících) v ČR v letech 1993 až 2013 – viz datový soubor potraty_CR.sta.

Předpokládejte, že tato časová řada má kubický trend. Odhadněte parametry trendové funkce.

Vypočtěte index determinace ID^2.

Proveďte celkový F-test.

Proveďte dílčí t-testy.

Proveďte analýzu reziduí.

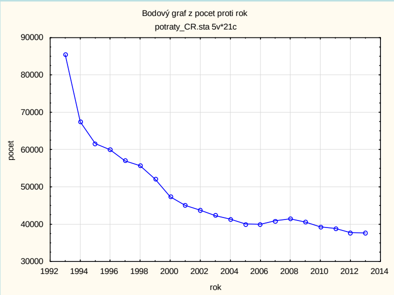
Sestrojte 95% intervaly spolehlivosti pro parametry trendové funkce.

Stanovte střední absolutní procentuální chybu predikce (MAPE).

Graficky znázorněte průběh časové řady s odhadnutým trendem, 95% pásem spolehlivosti a 95% predikčním pásem.

Řešení v systému STATISTICA:
Načteme datový soubor potraty_CR.sta o dvou proměnných rok, počet a 21 případech. Přidáme tři nové proměnné t, t2 a t3. Do proměnné t uložíme pořadová čísla 1 až 21, do proměnné t2 druhé mocniny pořadových čísel, do proměnné t3 třetí mocniny pořadových čísel.

Grafické znázornění časové řady:
Grafy – Bodové grafy – Proměnné X rok, Y pocet – OK – vypneme Lineární proložení – OK.
Formát – Všechny možnosti – Graf: Obecné – zaškrtneme Spojnice – OK. Vznikne spojnicový diagram.



Trendová funkce $\hat{f}(t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3$

Získání odhadů parametrů:
Statistiky – Vícenásobná regrese – Proměnné Závislé počet, Nezávislé t, t2, t3 - OK

	Výsledky regrese se závislou proměnnou : pocet (potraty_CR.sta) R= ,98289265 R2= ,96607796 Upravené R2= ,96009172 F(3,17)=161,38 p<,00000 Směrod. chyba odhadu : 2452,3					
N=21	b*	Sm.chyba z b*	b	Sm.chyba z b	t(17)	p-hodn.
Abs.člen			87825,07	2587,863	33,93729	0,000000
t	-4,43177	0,502979	-8767,76	995,088	-8,81104	0,000000
t2	6,36704	1,188830	556,08	103,828	5,35572	0,000052
t3	-2,84112	0,730560	-12,08	3,107	-3,88896	0,001180

Odhadnutá trendová funkce má tedy tvar:

$\hat{f}(t) = 87825{,}07 - 8767{,}76t + 556{,}08t^2 - 12{,}08t^3$, kde t = 1, …, 21.

Index determinace je 0,966, tedy kubická trendová funkce vysvětluje variabilitu dané časové řady z 96,6 %.

N=21	Výsledky regrese se závislou proměnnou : pocet (potraty_CR.sta) R= ,98289265 R2= ,96607796 Upravené R2= ,96009172 F(3,17)=161,38 p<,00000 Směrod. chyba odhadu : 2452,3					
	b*	Sm.chyba z b*	b	Sm.chyba z b	t(17)	p-hodn.
Abs.člen			87825,07	2587,863	33,93729	0,000000
t	-4,43177	0,502979	-8767,76	995,088	-8,81104	0,000000
t2	6,36704	1,188830	556,08	103,828	5,35572	0,000052
t3	-2,84112	0,730560	-12,08	3,107	-3,88896	0,001180

Testová statistika celkového F-testu je 161,38, p-hodnota je blízká 0, tedy na hladině významnosti 0,05 zamítáme hypotézu o nevýznamnosti modelu jako celku.

Všechny čtyři dílčí t-testy mají p-hodnoty menší než 0,05, tedy na hladině významnosti 0,05 zamítáme hypotézy o nulovosti parametrů β_0, β_1, β_2, β_3.

Posouzení nezávislosti reziduí pomocí Durbinovy – Watsonovy statistiky:

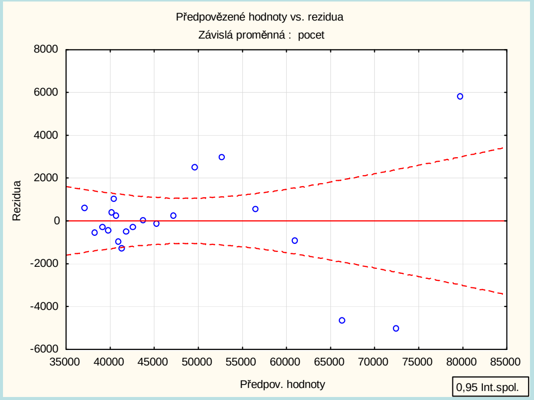
Statistiky – Vícenásobná regrese – proměnná Závislá: y, nezávislá t, t2, t3 – OK – na záložce Residua/předpoklady/předpovědi vybereme Reziduální analýza - Detaily – Durbin-Watsonova statistika:

	Durbin-Watson.d	Sériové korelace
Odhad	1,469592	0,096446

Hodnota D-W statistiky je poněkud nízká, ale podle testu autokorelace je vše v pořádku.

Posouzení homoskedasticity reziduí

Reziduální analýza – Bodové grafy – Předpovědi vs. rezidua



Testování nulovosti střední hodnoty reziduí:

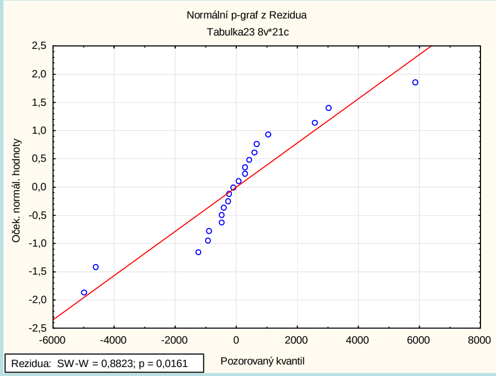
Pro proměnnou Rezidua z tabulky uložené pomocí Reziduální analýzy provedeme jednovýběrový t-test: Statistiky - Základní statistiky/tabulky – t-test, samost. vzorek – OK – proměnné Rezidua – OK.

Proměnná	Test průměrů vůči referenční konstantě (hodnotě) (Tabulka23) Průměr	Sm.odch.	N	Sm.chyba	Referenční konstanta	t	SV	p
Rezidua	0,000186	2260,908	21	493,3705	0,00	0,000000	20	1,000000

Na hladině významnosti 0,05 nezamítáme hypotézu, že střední hodnota reziduí je 0.

Ověření normality reziduí

Sestrojíme N-P plot reziduí a současně provedeme S-W test:



S-W test poskytuje p-hodnotu 0,0161, tedy na hladině významnosti 0,05 zamítáme hypotézu o normalitě reziduí.

Sestrojení 95% intervalů spolehlivosti pro parametry trendu:

Ve výstupní tabulce výsledků regrese přidáme za proměnnou p-hodn. dvě nové proměnné dm (pro dolní meze 95% intervalů spolehlivosti) a hm (pro horní meze 95% intervalů spolehlivosti). Do Dlouhého jména proměnné dm resp. hm napíšeme: =v3-v4*VStudent(0,975;17) resp. =v3+v4*VStudent(0,975;17)

	Výsledky regrese se závislou proměnnou : pocet (potraty_CR.sta) R= ,98289265 R2= ,96607796 Upravené R2= ,96009172 F(3,17)=161,38 p<,00000 Směrod. chyba odhadu : 2452,3							
N=21	b*	Sm.chyba z b*	b	Sm.chyba z b	t(17)	p-hodn.	dm =v3-v4*VSt	hm =v3+v4*VSt
Abs.člen			87825,07	2587,863	33,93729	0,000000	82365,1582	93284,9871
t	-4,43177	0,502979	-8767,76	995,088	-8,81104	0,000000	-10867,21	-6668,3052
t2	6,36704	1,188830	556,08	103,828	5,35572	0,000052	337,017114	775,134431
t3	-2,84112	0,730560	-12,08	3,107	-3,88896	0,001180	-18,637869	-5,5277001

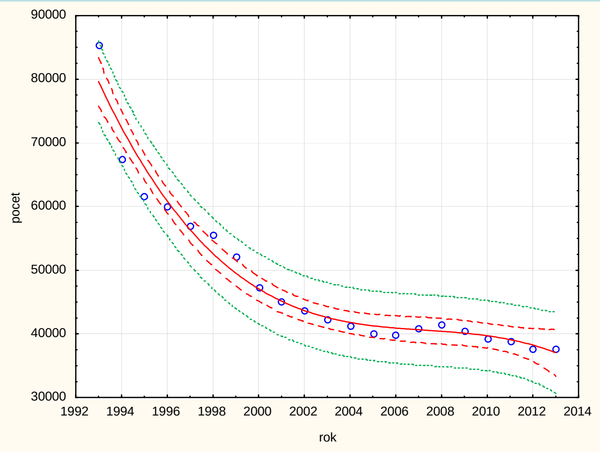
Vidíme, že $82365 < \beta_0 < 93285$ s pravděpodobností aspoň 0,95, $-10867 < \beta_1 < -6668$, $337 < \beta_2 < 775$, $-18{,}6 < \beta_3 < -5{,}2$ s pravděpodobností aspoň 0,95.

Výpočet MAPE:

Ve tabulce s uloženými rezidui odstraníme proměnné 8 – 13, přidáme proměnnou chyby a do jejího Dlouhého jména napíšeme =100*abs(v7/v2). Pak spočteme průměr této proměnné a zjistíme, že MAPE = 2,47 %.

Graf časové řady s proloženým kvadratickým trendem a pásem spolehlivosti a predikčním pásem získáme takto:

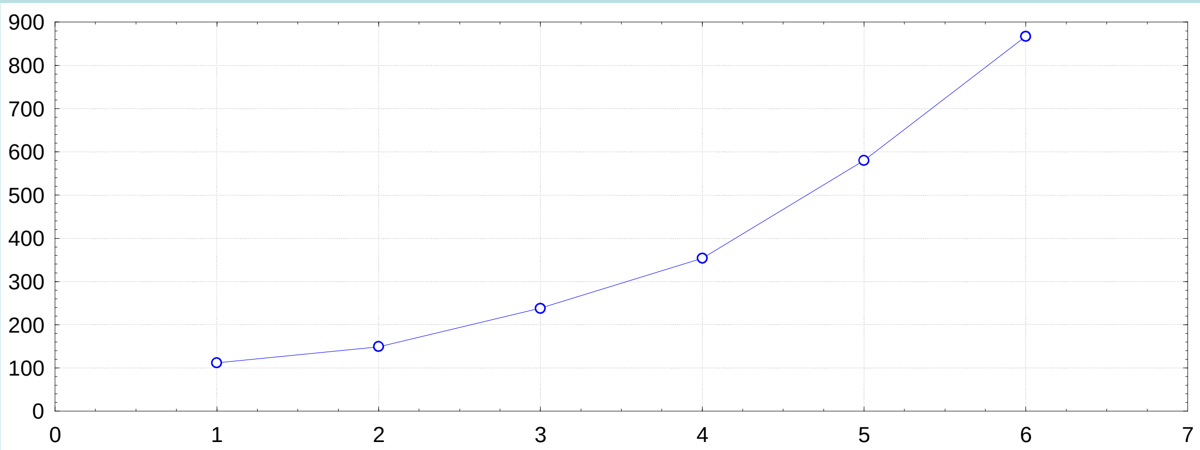
Grafy – Bodové grafy – Proměnné X rok, Y pocet – OK – Detaily Proložení Polynomiální. Ve vytvořeném grafu 2x klikneme na pozadí, vybereme Graf: Regresní pásy – Přidat nový pár pásů – Typ Spolehlivostní – OK. Totéž provedeme ještě jednou a nyní zaškrtneme Typ Predikční.



Příklad: Časová řada 112, 149, 238, 354, 580, 867 udává zisk (v tisících dolarů) jisté společnosti v prvních šesti letech její existence.
a) Graficky znázorněte průběh této časové řady.
b) Vypočtěte koeficienty růstu a graficky je znázorněte.
c) Z grafu časové řady a chování koeficientů růstu lze usoudit, že časová řada má exponenciální trend $f(t) = \beta_0 \beta_1^t$. Odhadněte jeho parametry.
d) Najděte odhad zisku společnosti v 7. a 8. roce její existence.
e) Zjistěte index determinace a sestrojte graf $\left(y_t, \hat{f}(t)\right)$, t = 1, ..., 6.

Řešení pomocí systému STATISTICA:
Vytvoříme datový soubor se dvěma proměnnými čas a Y a 6 případy.

ad a) Graficky znázorníme průběh této časové řady:
Grafy – Bodové grafy – Proměnné čas, Y – OK – vypneme proložení – OK.



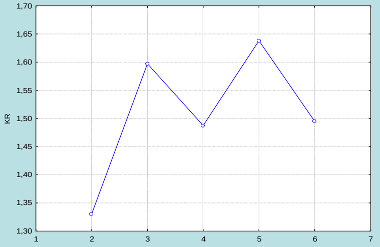
ad b) Výpočet koeficientů růstu:
Statistiky – Pokročilé lineární/nelineární modely – Časové řady/predikce – Proměnné Y – OK – OK (transformace, autokorelace, kříž. korelace, grafy) – Posun – Posun řad vzad - OK (transformovat vybrané řady) – návrat do transformace proměnných – Uložit proměnné.
Ve výstupní tabulce máme proměnné Y a Y_1:

	1 Y	2 Y_1
0		112,000
1	112,000	149,000
2	149,000	238,000
3	238,000	354,000
4	354,000	580,000
5	580,000	867,000
6	867,000	
7		

Za proměnnou Y_1 přidáme proměnnou KR a do jejího Dlouhého jména napíšeme =v2/v1.

	1 Y	2 Y_1	3 KR
0		112,000	
1	112,000	149,000	1,330357
2	149,000	238,000	1,597315
3	238,000	354,000	1,487395
4	354,000	580,000	1,638418
5	580,000	867,000	1,494828
6	867,000		
7			

Vytvoření grafu koeficientů růstu:
Klikneme pravým tlačítkem na název proměnné KR – Grafy bloku dat – Spojnicový graf: celé sloupce



Vidíme, že koeficienty růstu jsou přibližně konstantní.

ad c) Parametry modelu $f(t) = \beta_0 \beta_1^{\,t}$ odhadneme pomocí nelineární regrese:
Statistiky – Pokročilé lineární/nelineární odhady - Nelineární odhady - Vlastní regrese(MNČ) - OK. Do odhadované funkce vepíšeme y=beta0*beta1^cas - OK
Otevře se okno s názvem Odhad nelineárního modelu metodou nejmenších čtverců. Na liště Základní výsledky máme na výběr mezi Levenbergovou - Marquardtovou a Gaussovou - Newtonovou iterační metodou - jednu z nich zvolíme.
Na liště Detailní výsledky máme v případě problémů možnost změnit počet iterací, požadovanou přesnost a počáteční hodnoty parametrů.
Pokračujeme OK - otevře se okno Výsledky. Na liště Základ zvolíme Souhrn: Odhady parametru a získáme tabulku:

Model je: y=beta0*beta1^cas (zisk_spolecnosti.sta)
Záv.prom.:Y
Hladina spolehlivosti:95.0% (alfa =0.050)

	Odhad	Standard chyba	t-hodn. sv = 4	p-hodn.	Dol. sp. Mez	Hor. sp. Mez
beta0	65,35424	3,575415	18,27879	0,000053	55,42730	75,28119
beta1	1,53973	0,015590	98,76297	0,000000	1,49644	1,58301

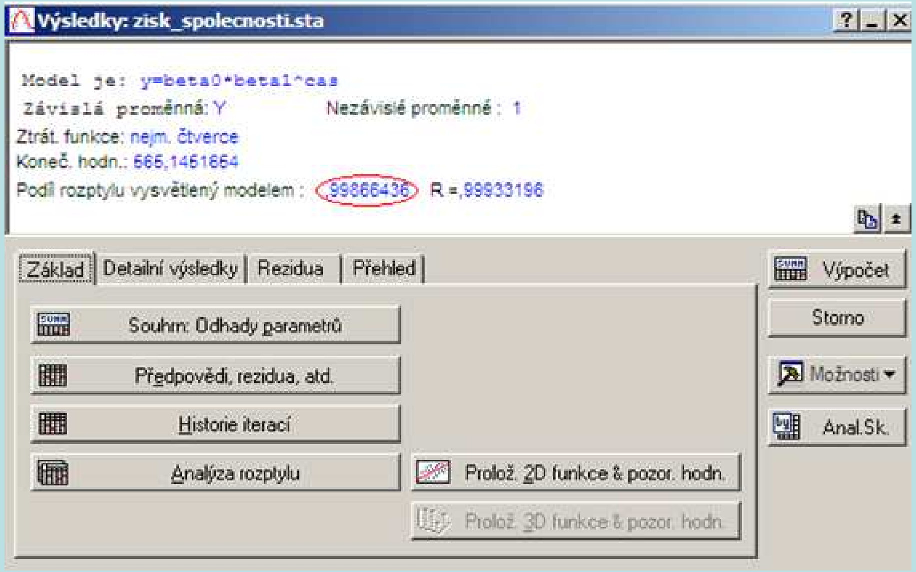
Odhadnutý model má tvar: $y = 65{,}35424 \cdot 1{,}53973^{cas}$
Oba parametry jsou významné na hladině významnosti 0,05.

ad d) Odhad zisku společnosti v 7. a 8. roce existence: Ve výstupní tabulce regrese odstaníme všechny proměnné kromě proměnné Odhad a sloupec parametrů transponujeme (Data – Transponovat – Soubor – OK)
Přidáme dvě nové proměnné Y7, Y8 a do Dlouhého jména proměnné Y7 napíšeme =v1*v2^7 (resp. do Dlouhého jména proměnné Y8 napíšeme =v1*v2^8).

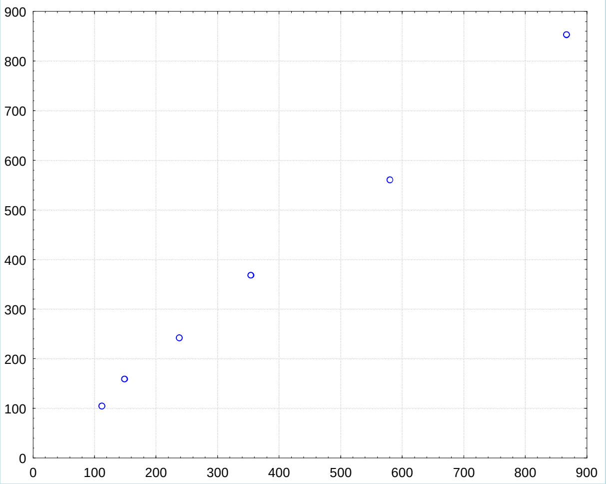
	1 beta0	2 beta1	3 y7	4 y8
Odhad	65,3542445	1,53972973	1340,866	2064,571

Předpověď zisku v 7. roce existence společnosti je tedy 1340,866 tisíc dolarů a v 8. roce 2064,571 tisíc dolarů.

ad e) Index determinace je $ID^2 = 0{,}9987$, jak je uvedeno v záhlaví tabulky regresní analýzy.



Graf závislosti predikovaných hodnot na hodnotách časové řady vytvoříme tak, že na liště Rezidua vybereme Pozorování vs. Předpovědi.



Jak index determinace, tak graf $\left(y_t, \hat{f}(t)\right)$ svědčí o tom, že model byl zvolen správně.

Můžeme též nakreslit dvourozměrný tečkový diagram s odhadnutou regresní křivkou:
Na liště Základ vybereme Prolož. 2D funkce & pozor. hodn.