

6. Lineární regresní modely

6.1 Jednoduchá regrese a validace

6.2 Testy hypotéz v lineární regresi

6.3 Kritika dat v regresním tripletu

6.4 Multikolinearita a polynomy

6.5 Kritika modelu v regresním tripletu

6.6 Kritika metody v regresním tripletu

6.7 Lineární a nelineární kalibrace

7. Korelační modely

Způsoby zadávání svých dat do programu ADSTAT

1. Psát svoje data přímo do ADSTATU:

1.1 Otevřete si okno Lineární regrese a po stisknutí F3 a Enter se otevřou demonstrační data.

1.2 Klávesou Delete nebo Backspace se obsah starých dat v okně smaže.

1.3 Nadepíšete si v prvním řádku třeba dva sloupce x a y a pro přechod na další řádek stisknete Enter. Nyní píšete číselné hodnoty pro x a y, oddělené vždy minimálně jednou mezerou mezerníkem. Na další řádek přejdete stisknutím Enter. Takto napíšete dva sloupce svých dat

1.4 Nyní si svoje v okně zapsaná data uložíte stisknutím klávesy F2 a musíte soubor dat pojmenovat heslem třeba Kocour.txt. Klávesou F4 odejdete z editačního okna dat a textové nadpisy zatím zůstaly.

1.5 Klávesou Esc odejdete z dat a začnete výpočet. Nadpisy sloupců se nepřenesou do výpočtu, program písmena automaticky odstraní. Nadpisy slouží jenom uživateli k popisu dat. Vaše data jsou uložena v ADSTATU v podprogramu LINREG a můžete je kdykoliv vyvolat po F3 zadáním Kocour.txt.



1.1 Defaultní Demo data



1.2 Smazání starých dat



1.3 Nadpisy a svoje data



1.4 Prohlížecké okno má nadpisy



1.5 Návrat totéž ukáže bez nadpisu

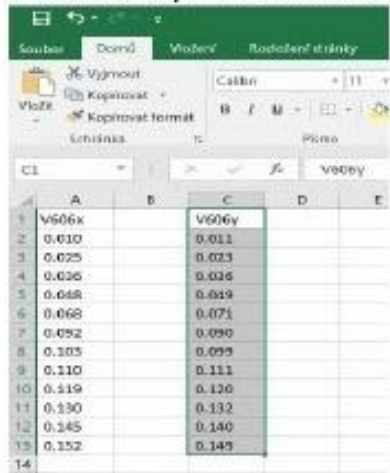
2. Zadávat svoje data do ADSTATU pomocí EXCELU:

2.1 Svoje data si připravíme v Excelu tak, že mezi sloupci x a y necháme 1 sloupec prázdný a máme Kocour.xlsx.

2.2a,b Svoje data uložíme v Excelu v ASCII kódu tak, že jim dáme koncovku .prn nebo .txt, např. Kocour.prn nebo Kocour.txt. Zde jsem užil V606.txt.

2.3 Ve Windows se v programu Lister nebo v jiném programu Total Commander snadno soubor dat přejmenuje z Kocour.prn na Kocour.txt.

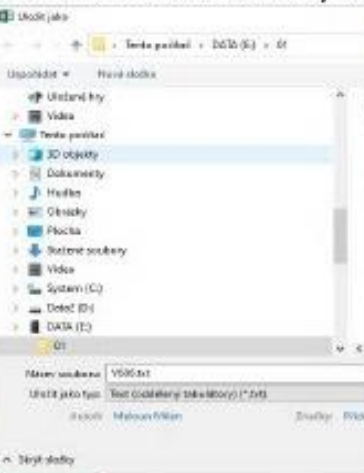
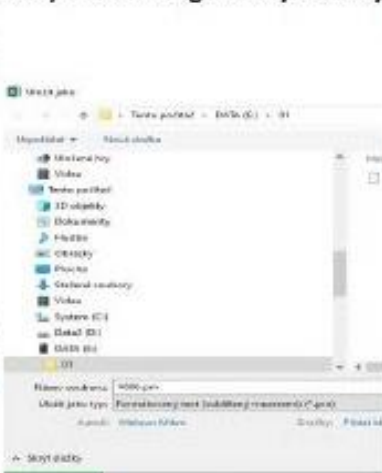
2.4 Soubor svých dat Kocour.txt musíte nyní překopírovat do knihovny lineární regrese c:/AD125/LINREG mezi ostatní soubory dat z Kompedia.



2.1 Svoje data uložit s koncovkou *.xlsx.



2.2a s koncovkou *.prn



2.2b s koncovkou *.txt

Soubor	Upravit	Možnosti	Kódování	Nápověda
V606x				
0.010			0.011	
0.025			0.023	
0.036			0.036	
0.048			0.049	
0.068			0.071	
0.092			0.090	
0.103			0.099	
0.110			0.111	
0.119			0.120	
0.130			0.132	
0.145			0.140	
0.152			0.149	

2.3 přejmenování souboru

REGRESNÍ DIAGNOSTIKA

Vyšetřuje **regresní triplet**, což představuje

- ◆ **Kritiku dat** (zkoumá kvalitu dat pro navržený model)
- ◆ **Kritiku modelu** (zkoumá kvalitu modelu pro daná data)
- ◆ **Kritiku metody odhadu** (prověřuje splnění všech předpokladů požadovaných metodou MNČ)

Kritika modelu

Vyšetření významnosti parametrů

Metoda:

Parciální regresní grafy

Parciální reziduální grafy

Studentův t -test významnosti parametru β

FORMULACE REGRESNÍHO MODELU

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1j} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2j} & \cdots & x_{2m} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{i1} & x_{i2} & \cdots & x_{ij} & \cdots & x_{im} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nj} & \cdots & x_{nm} \end{bmatrix} \cdot \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_j \\ \vdots \\ \beta_m \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_i \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

y
závisle
proměnná

X
nezávisle proměnná

β
regresní
parametry

ε
náhodná
chyba

$$\mathbf{y} = \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

TYPY REGRESNÍHO MODELU

Regresní model předpokládá, že **nezávislá proměnná** (nebo proměnné) je **nenáhodná** (tj. pevně zvolena či nastavena např. experimentátorem) a **závislá proměnná** je **náhodná** (měřená čili zatížená šumem, náhodnou chybou).

Tento předpoklad nebývá v praxi vždy splněn (často jsou totiž obě nebo všechny veličiny naměřené, potom nazýváme tento model **korelačním**).

Rozeznáváme:

- ◆ regresní modely **lineární** – mají **lineární postavení parametrů**
- ◆ regresní modely **nelineární** – mají **nelineární postavení parametrů**

PODSTATA REGRESNÍ ANALÝZY

Podstatou regrese je:

- ◆ Stanovit **nejlepší regresní model** (čili určit matematickou rovnici, která bude popisovat závislost y na x),
- ◆ Stanovit **parametry modelu** (tj. vyčíslit \mathbf{b} nejlepší odhady parametrů β),
- ◆ Stanovit **statistickou významnost modelu** (určit, zda nalezený model přispěje ke zpřesnění odhadu závisle proměnné oproti použití pouhého průměru),
- ◆ Výsledky regresního modelu **interpretovat** čili vysvětlit z hlediska zadání.

KRITÉRIA PRO HLEDÁNÍ A ROZLIŠENÍ NEJLEPŠÍHO REGRESNÍHO MODELU

Střední kvadratická chyba predikce (MEP)

$$\text{MEP} = \frac{1}{n} \sum_{i=1}^n \frac{e_i^2}{(1 - H_{ii})^2}$$

e_i^2 čtverec reziduí modelu
 H_{ii} i -tý diagonální prvek
projekční matice H

Akaikovo informační kritérium (AIC)

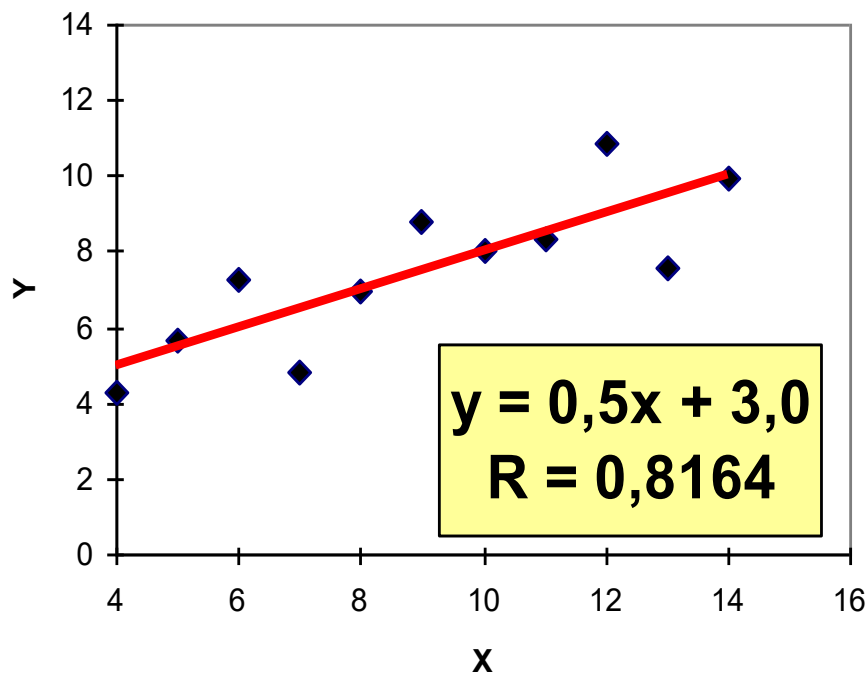
$$\text{AIC} = n \cdot \ln \left(\frac{\text{RSC}}{n} \right) + 2m$$

RSC reziduální součet čtverců
 m počet parametrů

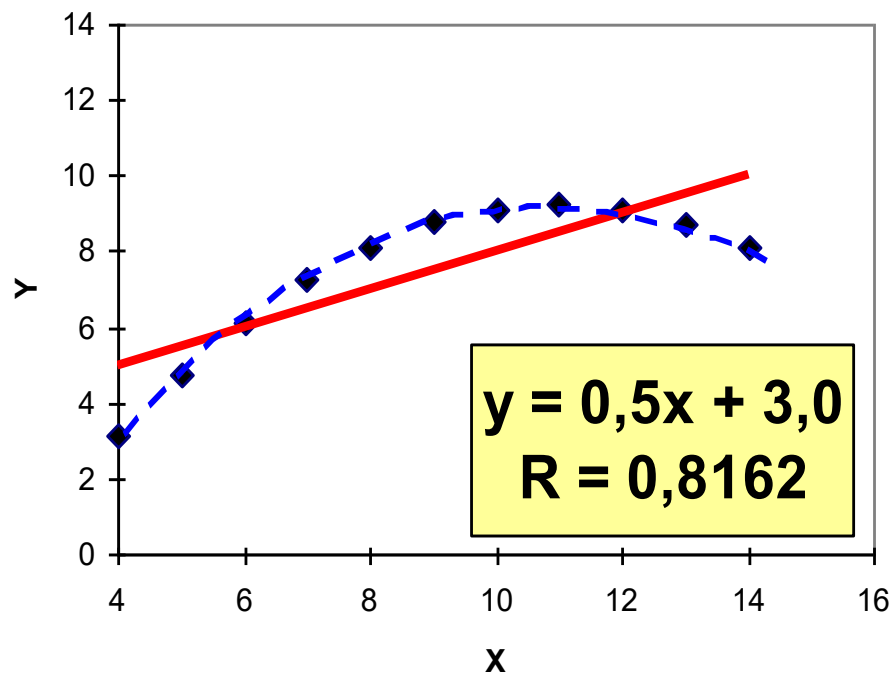
Pravidlo: Čím je AIC (nebo MEP) menší, tím je model vhodnější.

Postačí R a D [%] k nalezení nejlepšího modelu? Anscombe ukázal ve své úloze, že nikoliv!

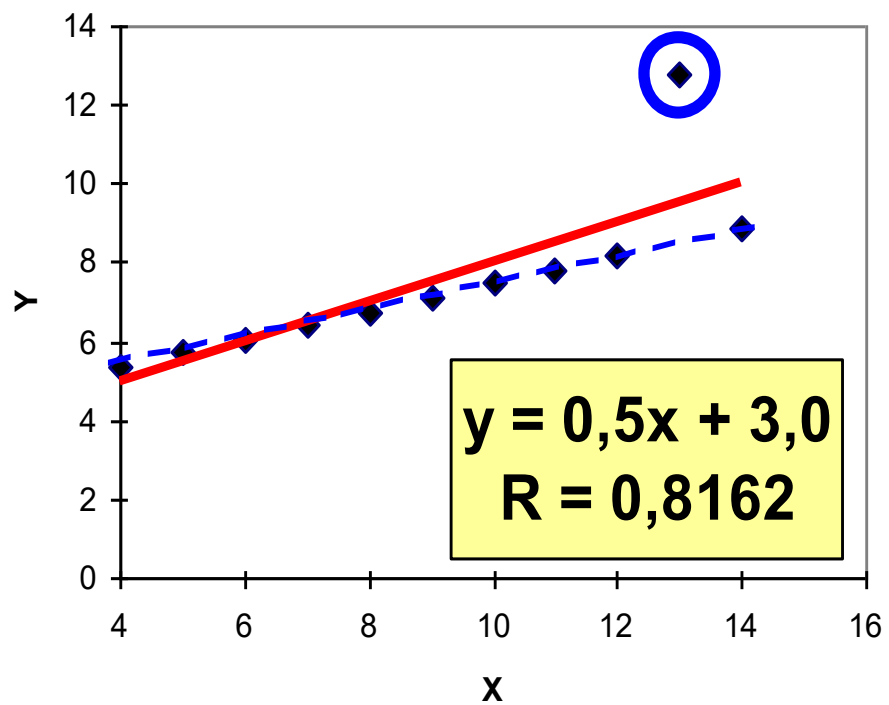
Výběr A



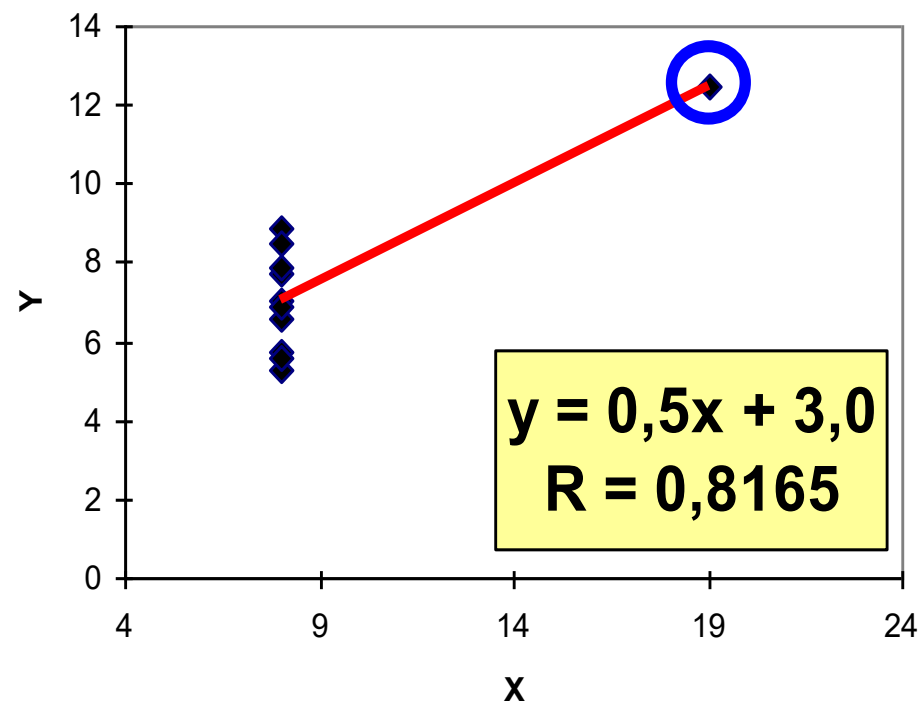
Výběr B



Výběř C



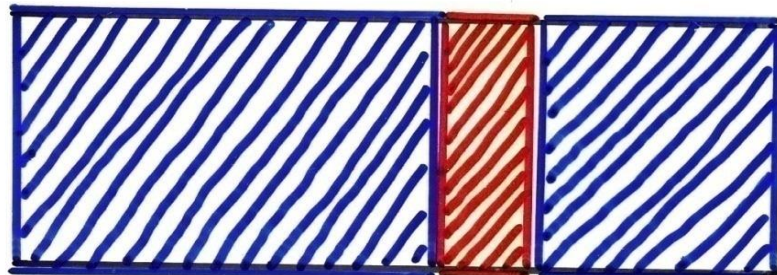
Výběř D



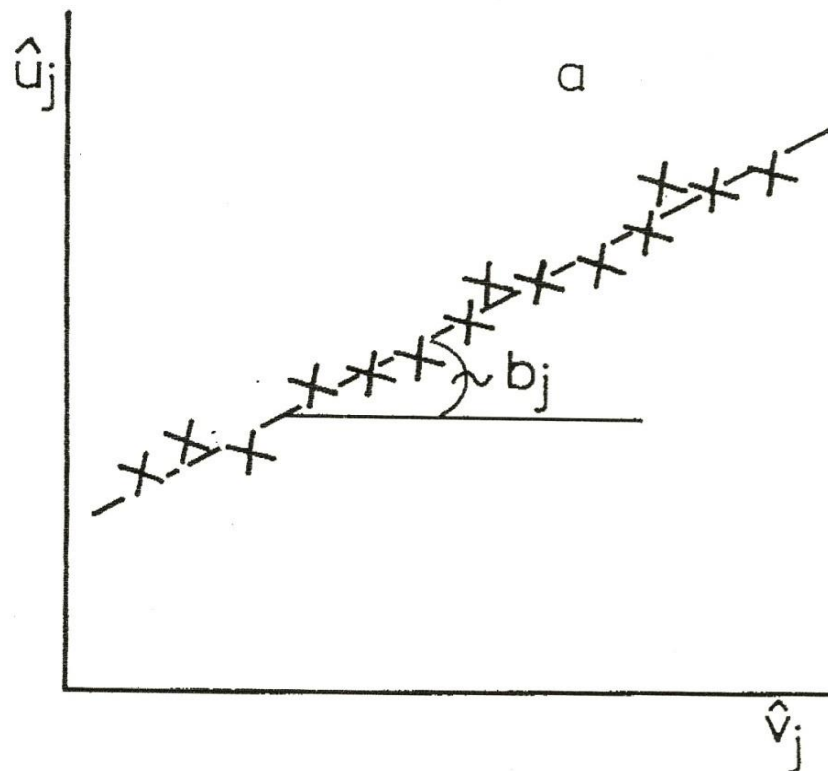
Belseyho parciální regresní grafy

(partial regression leverage plots)

$$\hat{y} = X_{(j)}\beta^* + x_j \cdot c + \varepsilon$$



$X_{(j)}$ jsou konstanty
 x_j

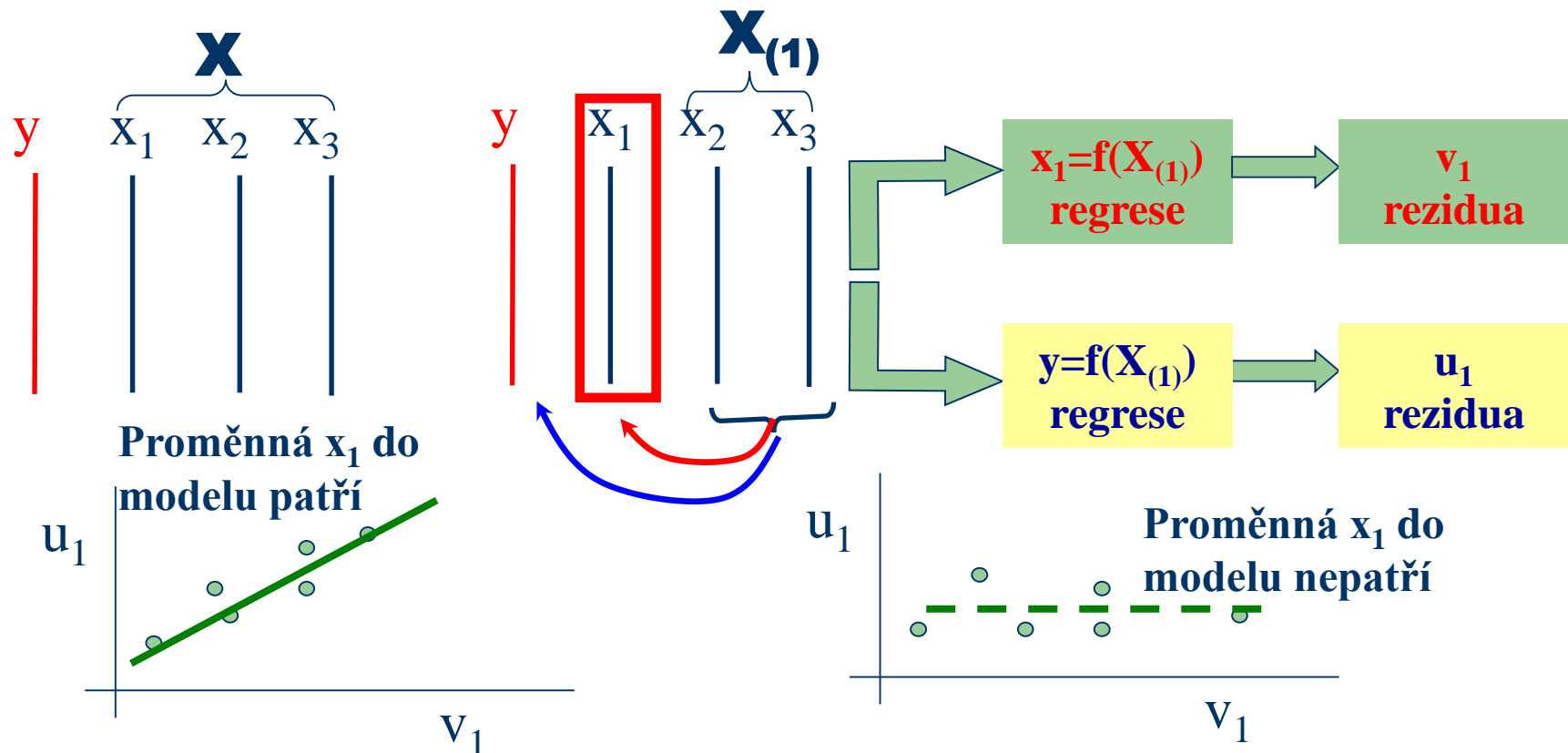


umožňují:

1. Posouzení kvality navrženého regresního modelu,
2. Indikují přítomnost vlivných bodů,
3. Nesplnění předpokladů klasické MNC,
4. Vyjadřují závislost mezi y a zvolenou proměnnou x_j při statisticky neměnném vlivu ostatních $X_{(j)}$

Grafický výklad parciálního regresního grafu

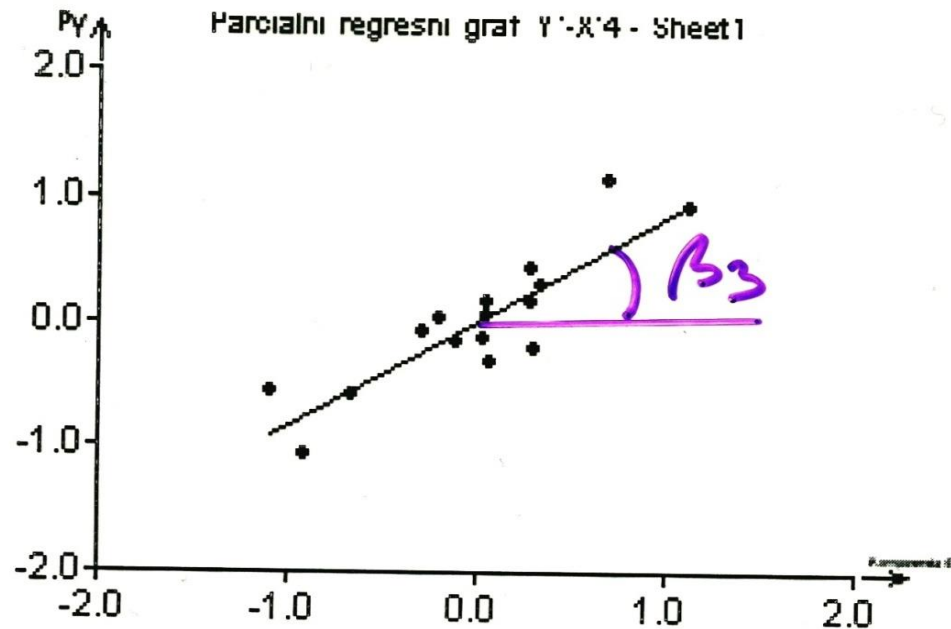
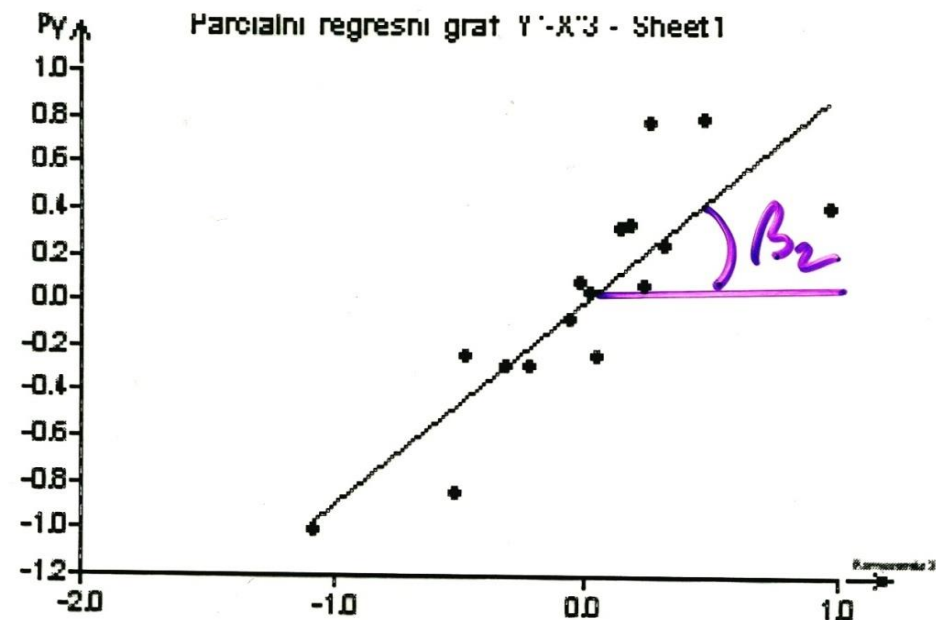
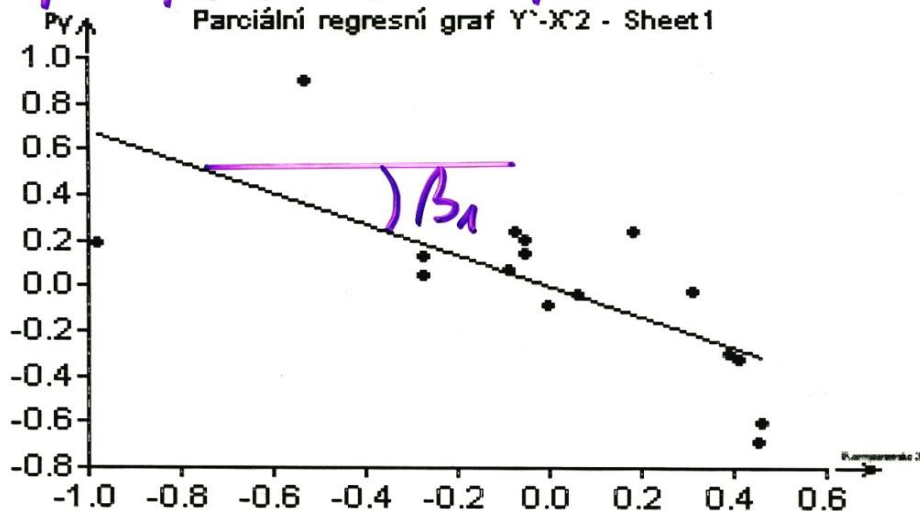
Příklad: Zajímá nás, zda všechny proměnné x_{1-3} jsou v modelu oprávněně. Postup je zde nyní vysvětlen pro proměnnou x_1 , zda x_1 patří do modelu či ne.



Ukázky parciálních regresních grafů u vícenásobného regresního modelu pro $m = 3$ u úlohy M619, tj. Cd v otrubách x_1 , stonku x_2 a kořeni x_3 pšenice

Kritika modelu: parciální regresní graf

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$



Vzorová úloha M619 na výstavbu lineárního regresního modelu pomocí parciálního regresního grafu

Jak vysvětlit výsledky úlohy a napsat semestrální práci je
v *Kompendiu*, str. 692

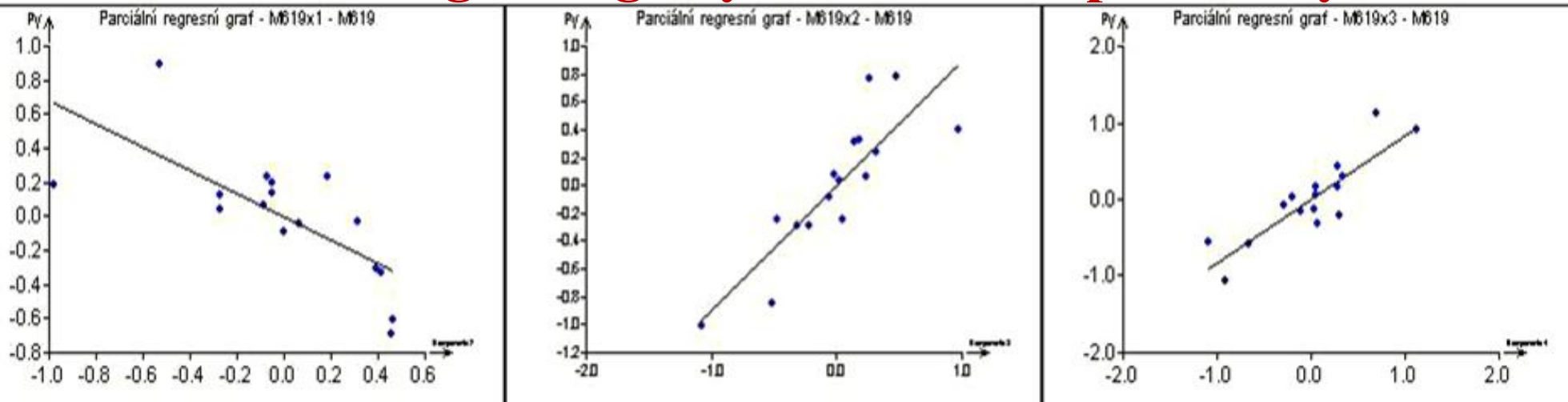
Návod k vysvětlení a napsání semestrální práce k zadané úloze najdete ve cvičebnici *M. Meloun, J. Militký: Kompendium statistického zpracování dat, Karolinum Praha 1992, ISBN 978-80-246-2196-8*, ve vzorovém příkladu na str. 692 – 705 a obecný výklad outputu v lineární regresi na str. 582 - 583. V ADSTATu používejte na přenos výstupů a obrázků **PrintScreen** a upravte pak negativ na černobílý pozitiv ve MS Windows.

Úloha M619. *Vliv tří parametrů na obsah kadmia v potravinářské pšenici*
Obsah kadmia v zrně y [mg/l] v závislosti na obsahu kadmia v otrubách x_1 [mg/l], ve stonku s listy x_2 [mg/l] a v kořenovém systému x_3 [mg/l].
Vyšetřete regresní triplet (data, model, metoda) a nalezněte lineární regresní model.

Výstavba lineárního regresního modelu:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

Parciální regresní grafy tří nezávisle proměnných



Porovnání výsledků z více numerických regresních programů

(3) ODHADY PARAMETRŮ A TESTY VÝZNAMNOSTI: **ADSTAT 1.25**

Parametr	Odhad	Směrodatná odchylka	Test $H_0: B[j] = 0$ vs. $H_A: B[j] \neq 0$	t-kriterium	hypoteza H_0 je	Hlad. výz.
B[0]	-7.2666E-02	1.3791E-01	-5.2692E-01	Akceptována	0.608	
B[1]	-6.8505E-01	1.9165E-01	-3.5746E+00	Zamítnuta	0.004	
B[2]	8.9619E-01	1.6072E-01	5.5761E+00	Zamítnuta	0.000	
B[3]	8.3769E-01	1.3322E-01	6.2879E+00	Zamítnuta	0.000	

Odhady parametrů

Proměnná

Proměnná	Odhad	Sm. od.	Závěr	Pravděp.	Spodní	Horní
Abs	-0.073	0.138	Nevýzn.	0.608	-0.373	0.228
M619x1	-0.685	0.192	Význ.	0.004	-1.103	-0.267
M619x2	0.896	0.161	Význ.	0.000	0.546	1.246
M619x3	0.838	0.133	Význ.	0.000	0.547	1.128

QC-EXPERT 3.1

NCSS2007

Regression Equation Section

Independent Variable	Regression Coefficient	Standard Error	T-Value (Ho: B=0)	Prob Level	Decision (5%)	Power (5%)
Intercept	-7.266543E-02	0.137906	-0.5269	0.607852	Accept Ho	0.077473
M619x1	-0.6850502	0.1916463	-3.5746	0.003820	Reject Ho	0.906070
M619x2	0.8961921	0.1607203	5.5761	0.000121	Reject Ho	0.999127
M619x3	0.8376914	0.1332221	6.2879	0.000040	Reject Ho	0.999919
R-Squared	0.997161					

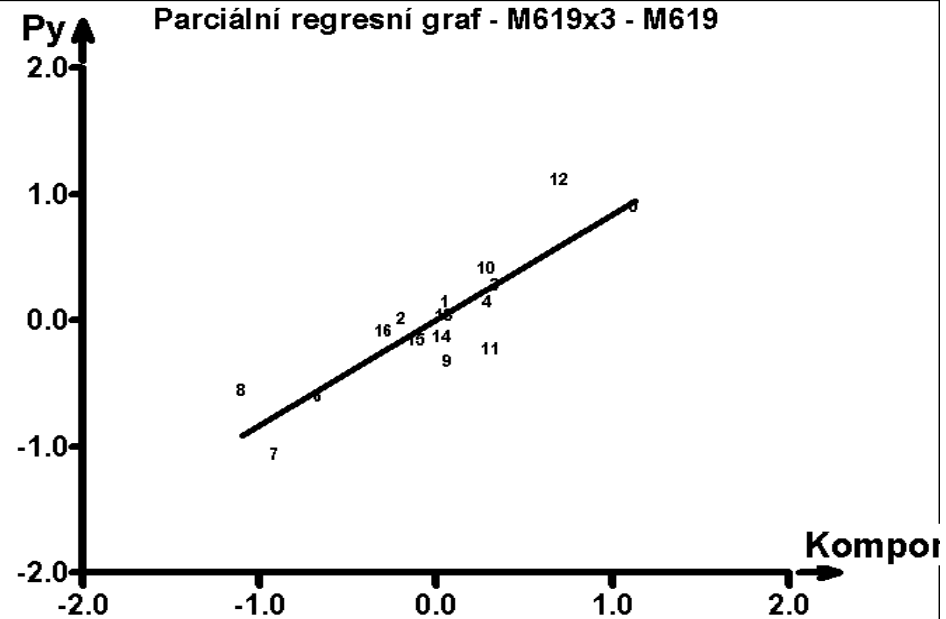
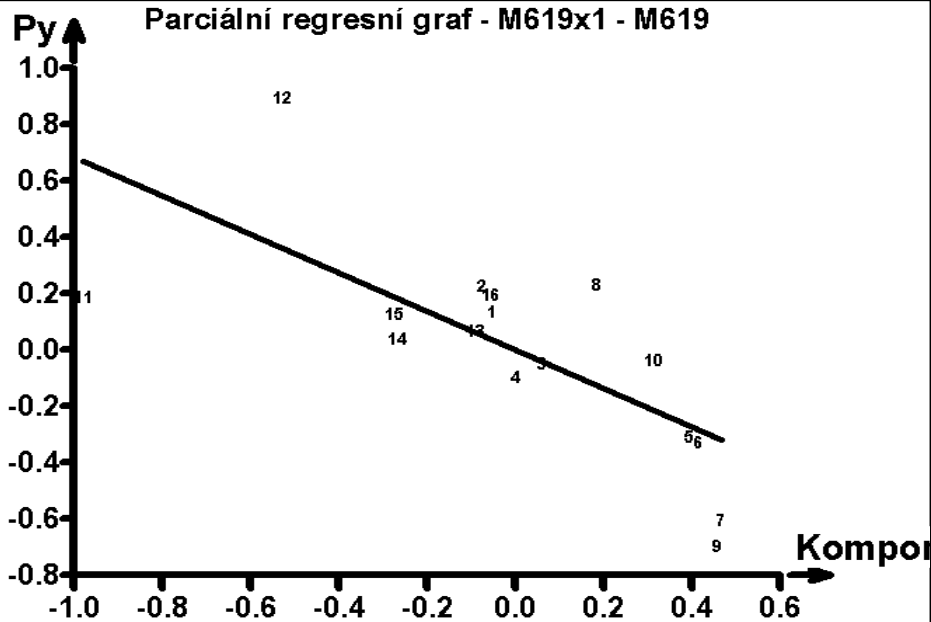
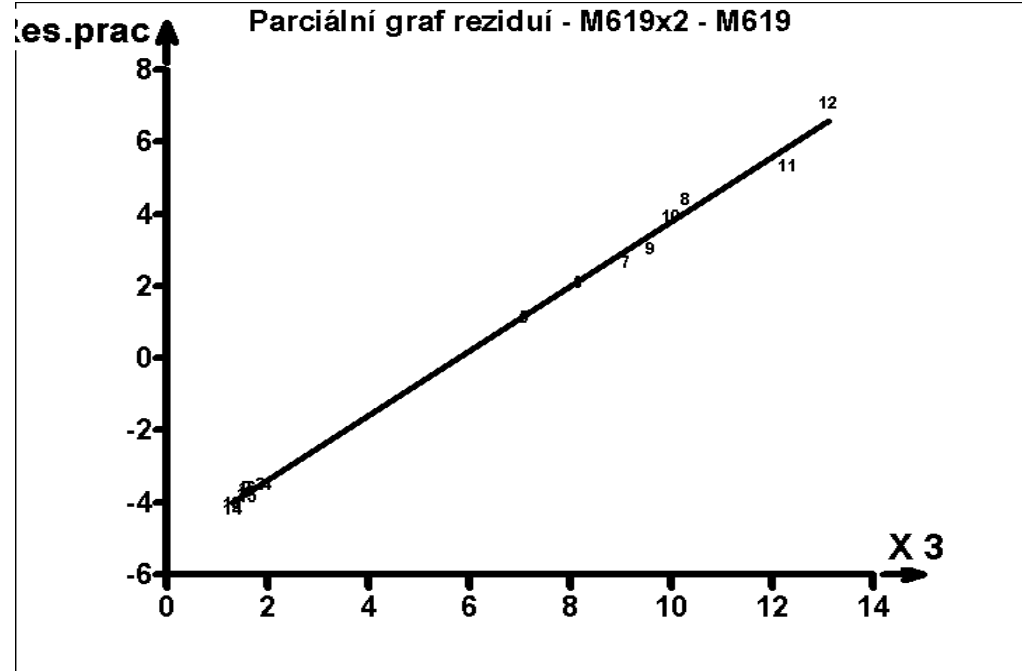
Model: $-7.266543E-02 - 0.6850502 * M619x1 + 0.8961921 * M619x2 + 0.8376914 * M619x3$

ADSTAT 1.25: statistická kritéria věrohodnosti regr. modelu

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R	: 9.9858E-01
Koeficient determinace, R^2	: 9.9716E-01
Predikovaný korelační koeficient, R_p^2	: 9.9527E-01
Střední kvadratická chyba predikce, MEP	: 2.1014E-01
Akaikeho informační kritérium, AIC	:-3.6180E+01

Parciální regresní grafy pro tři nezávisle proměnné mají vesměs nenulovou směrnici.



Závěr: Nejlepší lineární regresní model musí vždy obsahovat důkazy čili kritéria věrohodnosti D, MEP, AIC a $s(e)$

D. Data bez 8, 11, 12 a opravený model vedou k závěru

$$y = -0.86(0.37)x_1 + 0.95(0.25)x_2 + 0.92(0.13)x_3$$

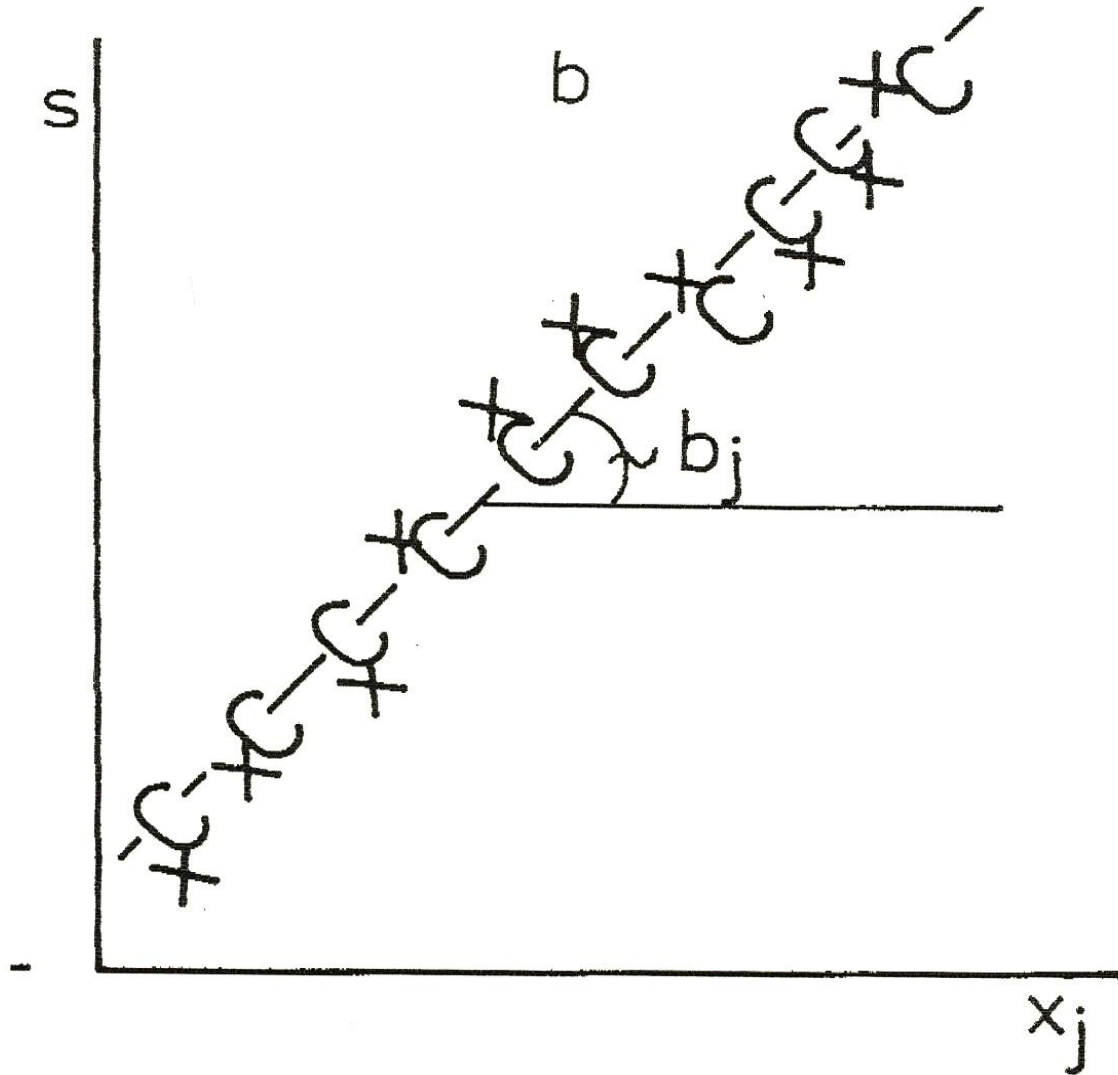
(LS: $t_{1-0.05/2}(13-3) = 2.228$, $D = 99.83\%$, $MEP = 0.05101$, $AIC = -43.55$, $s(e) = 0.170$)

β_0	0.000	---	---	---	---
β_1	-0.855	0.372	-2.3002	Významný	0.044
β_2	0.954	0.251	3.8038	Významný	0.003
β_3	0.916	0.132	6.9263	Významný	0.000

Model: $y = -0.86(0.37)x_1 + 0.95(0.25)x_2 + 0.92(0.13)x_3$

Parciální reziduální grafy

(grafy "komponenta + reziduum")

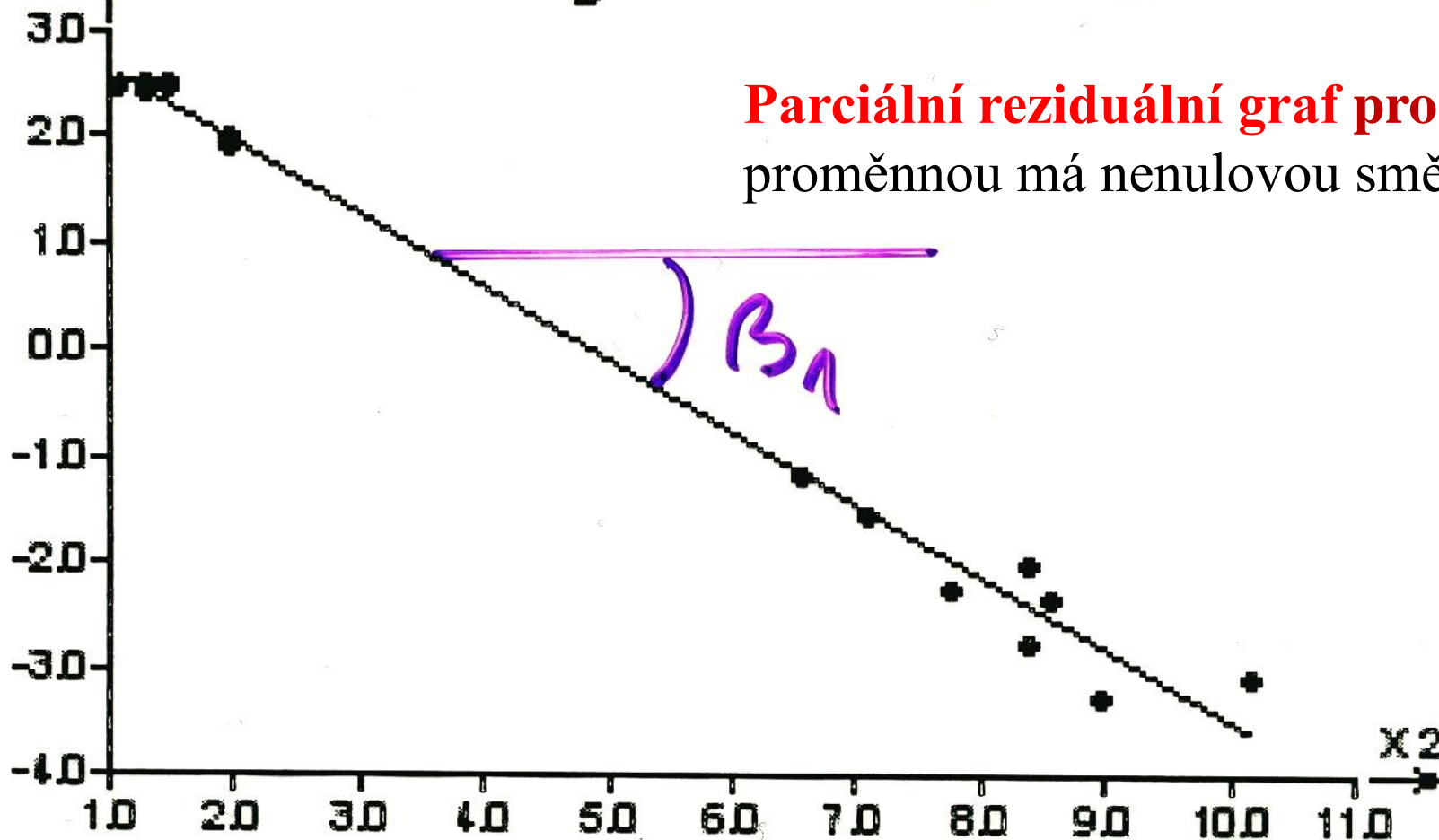


Kritika modelu: parciální reziduální graf

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

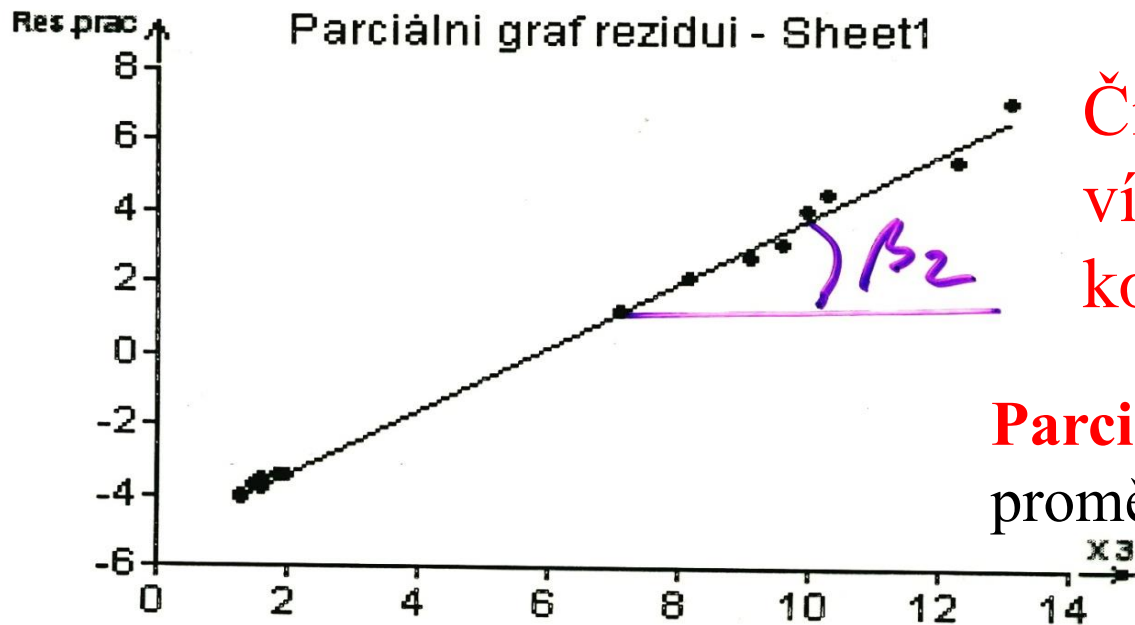
Res. graf

Parciální graf rezidui - Sheet1



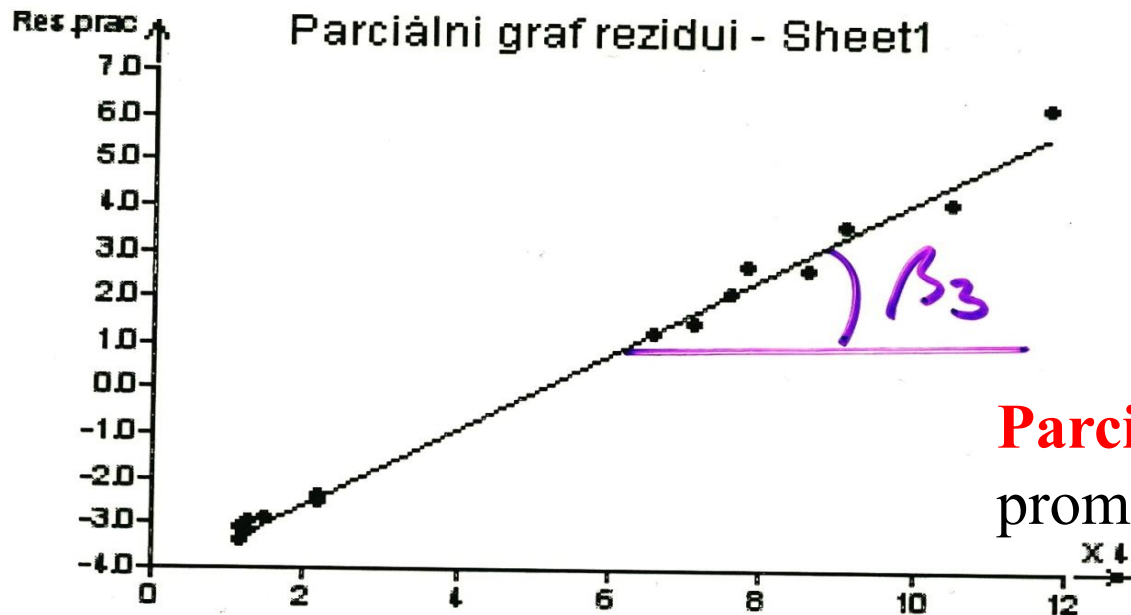
Parciální reziduální graf pro x_1
proměnnou má nenulovou směrnici.

Čím více Cd v otrubách, tím méně Cd v zrně (negativní korelace).



Čím více Cd ve stonku, tím více Cd v zrna (pozitivní korelace).

Parciální reziduální graf pro x2 proměnnou má nenulovou směrnici.



Čím více Cd v kořeni, tím více Cd v zrna (pozitivní korelace).

Parciální reziduální graf pro x3 proměnnou má nenulovou směrnici.

Závěr: Nalezený nejlepší lineární regresní model musí obsahovat kritéria věrohodnosti

D. Data bez 8, 11, 12 a opravený model vedou k závěru

$$y = -0.86(0.37)x_1 + 0.95(0.25)x_2 + 0.92(0.13)x_3$$

(LS: $t_{1-0.05/2}(13-3) = 2.228$, $D = 99.83\%$, $MEP = 0.05101$, $AIC = -43.55$, $s(e) = 0.170$)

β_0	0.000	---	---	---	---
β_1	-0.855	0.372	-2.3002	Významný	0.044
β_2	0.954	0.251	3.8038	Významný	0.003
β_3	0.916	0.132	6.9263	Významný	0.000

Model: $y = -0.86(0.37)x_1 + 0.95(0.25)x_2 + 0.92(0.13)x_3$

OBEČNÝ POSTUP VÝSTAVBY REGRESNÍHO MODELU

Jak vysvětlit výsledky úlohy a napsat semestrální práci je
v *Kompendiu*, str. 582 a 692

Návod k vysvětlení výsledků a napsání semestrální práce k zadané úloze najdete ve cvičebnici *M. Meloun, J. Militký: Kompendium statistického zpracování dat, Karolinum Praha 1992, ISBN 978-80-246-2196-8*, ve vzorovém příkladu na str. 583 – 593 a 692 - 705 a obecný výklad outputu v lineární regresi je na str. 582 – 583. V ADSTATu používejte na přenos výstupů a obrázků **PrintScreen** a upravte pak negativ na černobílý pozitiv ve MS Windows.

1. Kvalita nalezených odhadů parametrů

a) Podle **intervalů spolehlivosti** (čím menší interval spolehlivosti, tím lépe)

$$\beta_j = b_j \pm \sqrt{C_{mm} \cdot m \cdot s^2 \cdot F_{1-\alpha; m; n-m}}$$

b) Podle **rozptylů parametrů**, kde pro kvalitní odhad musí platit (Sillénovo empirické pravidlo)

$$2 \cdot \sqrt{D(b_j)} < |b_j|$$

(3) ODHADY PARAMETRŮ A TESTY VÝZNAMNOSTI:

Parametr	Odhad	Směrodatná odchylna	Test H0: B[j] = 0 vs. HA: B[j] <> 0	t-kriterium	hypoteza H0 je	Hlad. výz.
B[0]	-7.2666E-02	1.3791E-01	-5.2692E-01	Akceptována	0.608	
B[1]	-6.8505E-01	1.9165E-01	-3.5746E+00	Zamítnuta	0.004	
B[2]	8.9619E-01	1.6072E-01	5.5761E+00	Zamítnuta	0.000	
B[3]	8.3769E-01	1.3322E-01	6.2879E+00	Zamítnuta	0.000	

2. Kvalita dosažené těsnosti proložení

- a) Podle **reziduálního rozptylu $s(y)$** .
- b) Podle **regresního rabatu D** (= koeficient determinace v %: čím více se blíží 100 %, tím lepší je proložení).

3. Vhodnost navrženého modelu

Akaikovo informační kritérium AIC (čím je menší nebo zápornější, tím vhodnější je navržený model).

Střední kvadratická chyba predikce MEP (čím je MEP menší, tím je predikční schopnost navrženého modelu lepší).

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R	: 9.9858E-01
Koeficient determinace, R^2	: 9.9716E-01
Predikovaný korelační koeficient, R_p^2	: 9.9527E-01
Střední kvadratická chyba predikce, MEP	: 2.1014E-01
Akaikeho informační kritérium, AIC	:-3.6180E+01

4. Predikční schopnost modelu

Střední kvadratická chyba predikce *MEP* (čím je *MEP* menší, tím je predikční schopnost navrženého modelu lepší).

5. Kvalita experimentálních dat

- a) Na základě analýzy rozličných druhů **reziduí**.
- b) Na základě **Indikace vlivných bodů** (tj. odlehlé body a extrémny).

(Odhlehlé body: Jackknife rezidua, standardizovaná rezidua, normovaná rezidua, predikovaná rezidua, rekurzivní rezidua,

Extrémy: Cookova vzdálenost, diagonální prvky projekční matice a věrohodnostní vzdálenosti).

(5) ANALÝZA KLASICKÝCH REZIDUÍ:

Bod	Meřená hodnota	Predikovaná hodnota	Směrodatná odchylka	Klasické reziduum	Relativní reziduum
i	yexp[i]	yvyp[i]	s(yvyp[i])	e[i]	er[i]
1	1.6000E+00	1.5006E+00	1.0203E-01	9.9415E-02	6.2135E+00
2	1.6000E+00	1.4227E+00	1.0587E-01	1.7733E-01	1.1083E+01
3	2.1000E+00	2.1029E+00	1.1045E-01	-2.9205E-03	-1.3907E-01
4	2.1000E+00	2.1925E+00	1.0412E-01	-9.2540E-02	-4.4067E+00
5	8.1000E+00	8.1354E+00	2.0871E-01	-3.5421E-02	-4.3729E-01
6	7.9000E+00	7.9410E+00	1.3352E-01	-4.1016E-02	-5.1919E-01
7	8.4000E+00	8.6869E+00	1.6249E-01	-2.8690E-01	-3.4155E+00
8	1.0300E+01	9.9377E+00	1.7913E-01	3.6232E-01	3.5176E+00
9	9.6000E+00	9.9805E+00	1.3555E-01	-3.8050E-01	-3.9636E+00
10	1.0800E+01	1.0621E+01	1.3061E-01	1.7918E-01	1.6591E+00
11	1.3100E+01	1.3581E+01	2.2832E-01	-4.8080E-01	-3.6703E+00
12	1.5100E+01	1.4565E+01	1.8609E-01	5.3530E-01	3.5450E+00
13	1.3000E+00	1.2908E+00	1.0514E-01	9.1821E-03	7.0632E-01
14	1.2000E+00	1.3441E+00	1.1635E-01	-1.4406E-01	-1.2005E+01
15	1.5000E+00	1.5597E+00	1.1723E-01	-5.9675E-02	-3.9784E+00
16	1.5000E+00	1.3389E+00	1.0987E-01	1.6110E-01	1.0740E+01
Rezidualní součet čtverců, RSC:				1.0114E+00	
Průměr absolutních hodnot reziduí, Me:				1.9048E-01	
Průměr relativních reziduí, Mer:				4.3750E+00	
Odhad reziduálního rozptylu, $s^2(e)$:				8.4282E-02	
Odhad směrodatné odchylky reziduí, s(e):				2.9031E-01	
Odhad šikmosti reziduí, g1(e):				9.2356E-02	
Odhad špičatosti reziduí, g2(e):				2.8755E+00	

(7) INDIKACE VLIVNÝCH BODŮ:

(* indikuje odlehlý nebo vlivný bod)

Bod	Standard. reziduum	Jackknife reziduum	Predikované reziduum	Diagonální prvky
i	eS[i]	eJ[i]	eP[i]	H[i,i]
1	3.6577E-01	3.5217E-01	1.1343E-01	1.2352E-01
2	6.5601E-01	6.3966E-01	2.0453E-01	1.3298E-01
3	-1.0878E-02	-1.0415E-02	-3.4148E-03	1.4474E-01
4	-3.4147E-01	-3.2854E-01	-1.0620E-01	1.2863E-01
5	-1.7552E-01	-1.6827E-01	-7.3309E-02	5.1683E-01*
6	-1.5911E-01	-1.5250E-01	-5.2020E-02	2.1152E-01
7	-1.1925E+00	-1.2161E+00	-4.1777E-01	3.1326E-01
8	1.5859E+00	1.7078E+00	5.8504E-01	3.8069E-01
9	-1.4821E+00	-1.5700E+00	-4.8658E-01	2.1800E-01
10	6.9109E-01	6.7525E-01	2.2465E-01	2.0240E-01
11	-2.6814E+00	-4.0550E+00*	-1.2604E+00	6.1852E-01*
12	2.4023E+00	3.1923E+00*	9.0863E-01	4.1087E-01
13	3.3931E-02	3.2488E-02	1.0568E-02	1.3115E-01
14	-5.4162E-01	-5.2502E-01	-1.7163E-01	1.6062E-01
15	-2.2469E-01	-2.1557E-01	-7.1301E-02	1.6305E-01
16	5.9952E-01	5.8279E-01	1.8803E-01	1.4322E-01

Bod

Věrohodnostní vzdálenosti

i	LD(b)[i]	LD(s ²)[i]	LD(b,s ²)[i]
1	2.5120E-02	2.2351E-02	4.6185E-02
2	8.7766E-02	6.2215E-03	9.1797E-02
3	2.6701E-05	3.2606E-02	3.2632E-02
4	2.2933E-02	2.3569E-02	4.5295E-02
5	4.3879E-02	3.0095E-02	7.1395E-02
6	9.0525E-03	3.0538E-02	3.9045E-02
7	8.4237E-01	3.0991E-02	9.5089E-01
8	1.9389E+00	2.4682E-01	2.6917E+00
9	7.9637E-01	1.5878E-01	1.0958E+00
10	1.6078E-01	4.5425E-03	1.6232E-01
11	1.0861E+01*	7.8276E+00*	4.4183E+01*
12	4.6276E+00	3.4386E+00	1.3130E+01*
13	2.3173E-04	3.2520E-02	3.2738E-02
14	7.4671E-02	1.2499E-02	8.4425E-02
15	1.3108E-02	2.8534E-02	4.0880E-02
16	7.9906E-02	9.2132E-03	8.6631E-02

Indikace vlivných dat

A. Analýza reziduí

Index	Standard.	Jackknife	Predik.	Diag(Hii)	Diag(H*ii)	Cook. vzd.
1.000	0.366	0.352	0.113	0.124	0.133	0.013
2.000	0.656	0.640	0.205	0.133	0.164	0.025
3.000	-0.011	-0.010	-0.003	0.145	0.145	0.000
4.000	-0.341	-0.329	-0.106	0.129	0.137	-0.013
5.000	-0.176	-0.168	-0.073	0.517	0.518	-0.047
6.000	-0.159	-0.152	-0.052	0.212	0.213	-0.011
7.000	-1.193	-1.216	-0.418	0.313	0.395	-0.136
8.000	1.586	1.708	0.585	0.381	0.510	0.244
9.000	-1.482	-1.570	-0.487	0.218	0.361	-0.103
10.000	0.691	0.675	0.225	0.202	0.234	0.044
11.000	-2.681	-4.055	-1.260	0.619	0.847	-1.087
12.000	2.402	3.192	0.909	0.411	0.694	0.419
13.000	0.034	0.032	0.011	0.131	0.131	0.001
14.000	-0.542	-0.525	-0.172	0.161	0.181	-0.026
15.000	-0.225	-0.216	-0.071	0.163	0.167	-0.011
16.000	0.600	0.583	0.188	0.143	0.169	0.025

B. Analýza vlivu

Index	Atkinson. vzd.	Andrew.-P	Vliv na Y^	Vliv LD(b	Vliv LD(s)	Vliv LD(b,s
1.000	0.229	0.867	0.132	0.025	0.022	0.046
2.000	0.434	0.836	0.251	0.088	0.006	0.092
3.000	0.007	0.855	-0.004	0.000	0.033	0.033
4.000	0.219	0.863	-0.126	0.023	0.024	0.045
5.000	0.301	0.482	-0.174	0.044	0.030	0.071
6.000	0.137	0.787	-0.079	0.009	0.031	0.039
7.000	1.423	0.605	-0.821	0.842	0.031	0.951
8.000	2.319	0.490	1.339	1.939	0.247	2.692
9.000	1.436	0.639	-0.829	0.796	0.159	1.096
10.000	0.589	0.766	0.340	0.161	0.005	0.162
11.000	8.943	0.153	-5.163	10.861	7.828	44.183
12.000	4.618	0.306	2.666	4.628	3.439	13.130
13.000	0.022	0.869	0.013	0.000	0.033	0.033
14.000	0.398	0.819	-0.230	0.075	0.012	0.084
15.000	0.165	0.833	-0.095	0.013	0.029	0.041
16.000	0.413	0.831	0.238	0.080	0.009	0.087

6. Testy v regresním tripletu Data + Model + Metoda pro Kritiku (zvolené matematické) metody

- 6.1 Fisher-Snedecorův test celkové regrese,
- 6.2 Scottovo kritérium multikolinearity,
- 6.3 Cook-Weisbergův test heteroskedasticity,
- 6.4 Jarque-Berrův test normality reziduí,
- 6.5 Waldův test autokorelace,
- 6.6 Znaménkový test reziduí.

(6) TESTOVÁNÍ REGRESNÍHO TRIPLETU (DATA + MODEL + METODA):

Fisher-Snedocorův test významnosti regrese, F: 1.4050E+03

Tabulkový kvantil, $F(1-\alpha, m-1, n-m)$: 3.4903E+00

Závěr: Navržený model je významný.

Spočtená hladina významnosti: 0.000

Scottovo kritérium multikolinearity, M: 9.6106E-01

Závěr: Navržený model není korektní.

Cook-Weisbergův test heteroskedasticity, Sf: 5.3739E+00

Tabulkový kvantil, $\chi^2(1-\alpha, 1)$: 3.8415E+00

Závěr: Rezidua vykazují heteroskedasticitu.

Spočtená hladina významnosti: 0.020

Jarque-Berraův test normality reziduí, L(e): 3.3085E-02

Tabulkový kvantil, $\chi^2(1-\alpha, 2)$: 5.9915E+00

Závěr: Normalita je prokázána.

Spočtená hladina významnosti: 0.984

Waldův test autokorelace, Wa: 1.0320E+01

Tabulkový kvantil, $\chi^2(1-\alpha, 1)$: 3.8415E+00

Závěr: Rezidua jsou autokorelována.

Spočtená hladina významnosti: 0.001

Znamékový test, Dt: -1.9739E-01

Tabulkový kvantil, $N(1-\alpha/2)$: 1.6449E+00

Závěr: Rezidua nevykazují trend.

Spočtená hladina významnosti: 0.422

Úlohy na výstavbu lineárního regresního modelu dle regresního tripletu

Kritika modelu

Jak vysvětlit výsledky úlohy a napsat semestrální práci je
v *Kompendiu*, str. 582 a 692

Návod k vysvětlení výsledků a napsání semestrální práce k zadané úloze najdete ve cvičebnici *M. Meloun, J. Militký: Kompendium statistického zpracování dat, Karolinum Praha 1992, ISBN 978-80-246-2196-8*, ve vzorovém příkladu na str. 583 – 593 a 692 - 705 a obecný výklad outputu v lineární regresi je na str. 582 – 583. V ADSTATu používejte na přenos výstupů a obrázků **PrintScreen** a upravte pak negativ na černobílý pozitiv ve MS Windows.

Úloha M6.06 *Vliv čtyř faktorů na koncentraci amoniakálního dusíku*

Zadání: Je vyšetřován vliv teploty x_1 , pH x_2 , koncentrace celkového dusíku x_3 a koncentrace rozpuštěného kyslíku x_4 na koncentraci amoniakálního dusíku y v odtoku z dosazovací nádrže.

Úkoly:

- (1) Postavte vícerozměrný lineární regresní model a vyšetřete regresní triplet.
- (2) Pomocí parciálních reziduálních grafů vyšetřete statistickou významnost jednotlivých parametrů.
- (3) Jsou v datech vlivné body a je nutné odstranit vybočující hodnoty?
- (4) Jak řešíme úlohu v případě porušení předpokladů MNČ, a to především při nalezené heteroskedasticitě v datech, autokorelaci a nenormalitě chyb?

Data: Teplota x_1 , pH x_2 , koncentrace celkového dusíku x_3 , koncentrace rozpuštěného kyslíku x_4 , koncentrace amoniakálního dusíku y :

x_1	x_2	x_3	x_4	y
21	7.2	35	0.1	18.5
...
22.5	7.5	38.5	0.1	25.5

DOSBox 0.74, Cpu speed: 3000 cycles, Frameskip 0, Program: MLREG

F1=nápověda F2=uložení F3=čtení F4=numerický mód F5=zoom
 Řádek 1 Sloupec 5 Insert NUM M606.1

Textový editor

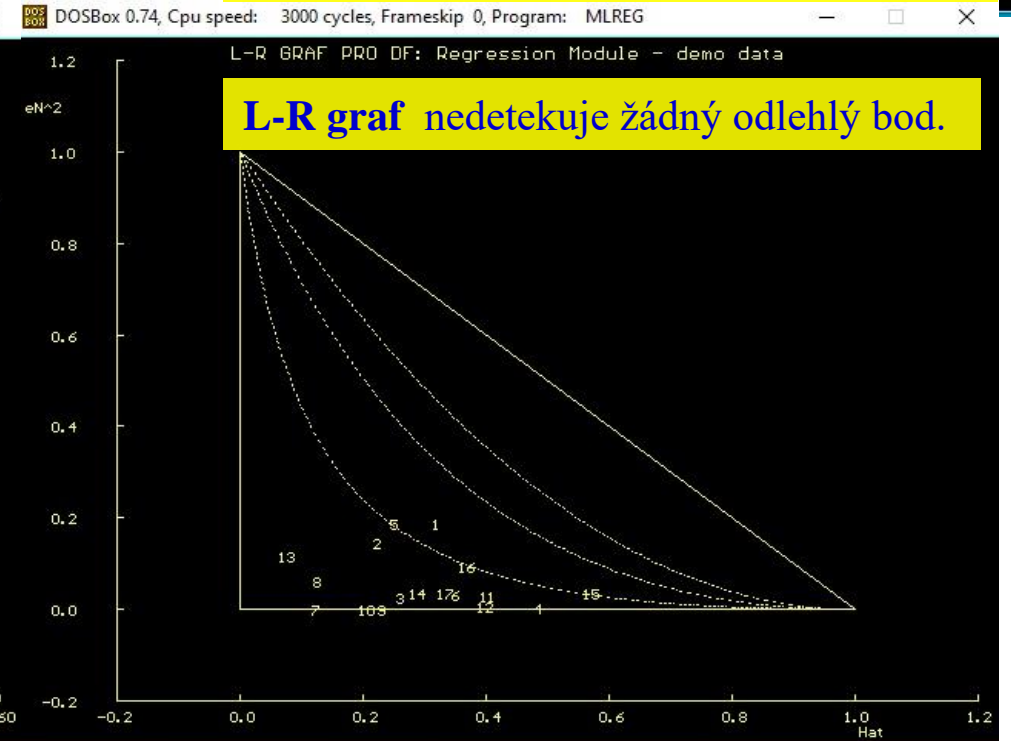
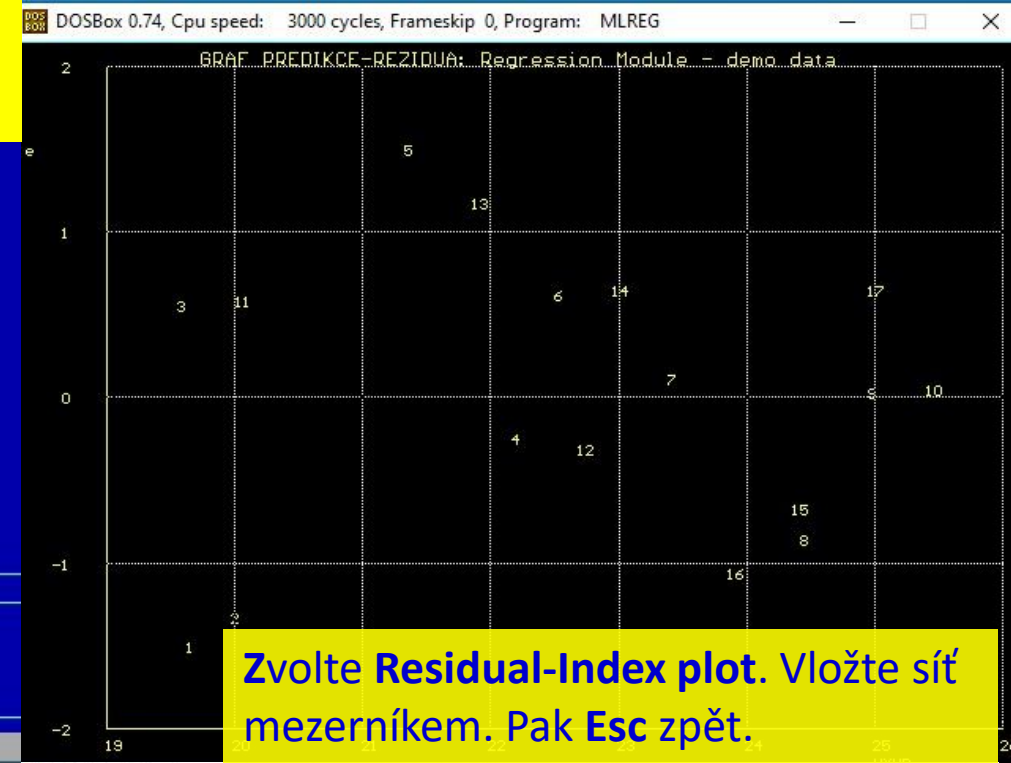
M606x1 M606x2 M606x3 M606x4 M606y

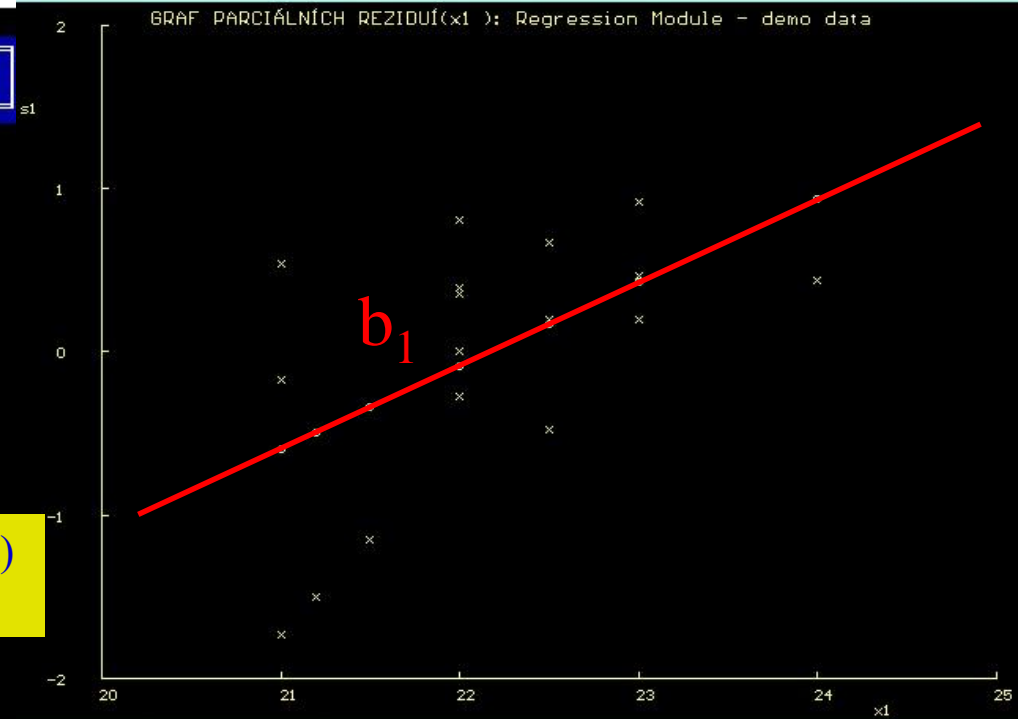
21.00000	7.20000	35.00000	0.10000	18.50000
21.20000				
21.00000				
22.00000				
21.00000	7.30000	39.00000	0.11000	22.50000
22.00000	7.30000	36.00000	0.13000	23.00000
22.00000	7.40000	40.00000	0.11000	23.50000
22.50000	7.45000	40.10000	0.11000	23.80000
22.50000	7.50000	39.50000	0.10000	25.00000
23.00000	7.50000	41.00000	0.11000	25.50000
22.00000	7.20000	38.00000	0.10000	20.50000
23.00000	7.25000	38.50000	0.15000	22.50000
22.00000	7.30000	38.00000	0.11000	22.80000
23.00000	7.30000	37.00000	0.13000	23.50000

Čtení ze souboru: M606.txt

DC

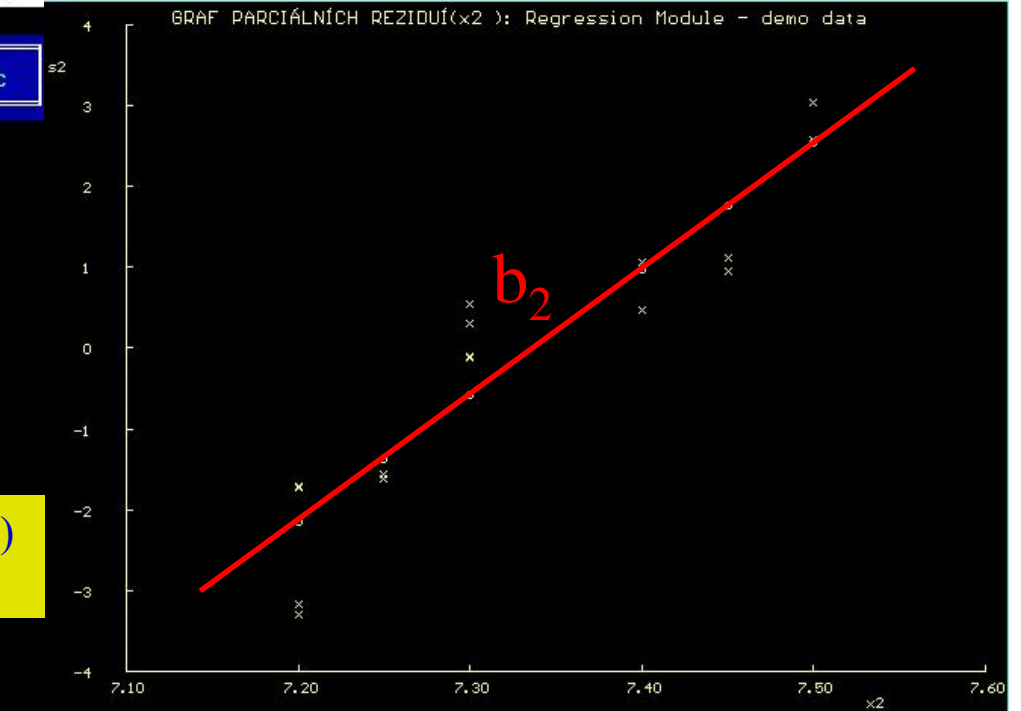
Otevřít Data, pak stisk F3, pak smazat demo.dat, napsat M606 a ukončit Esc a A.





Parciální reziduální graf testuje významnost (= šikmost) směrnice b_1 a rozptýlení bodů okolo směrnice.

Zobrazení grafů komponent a reziduí



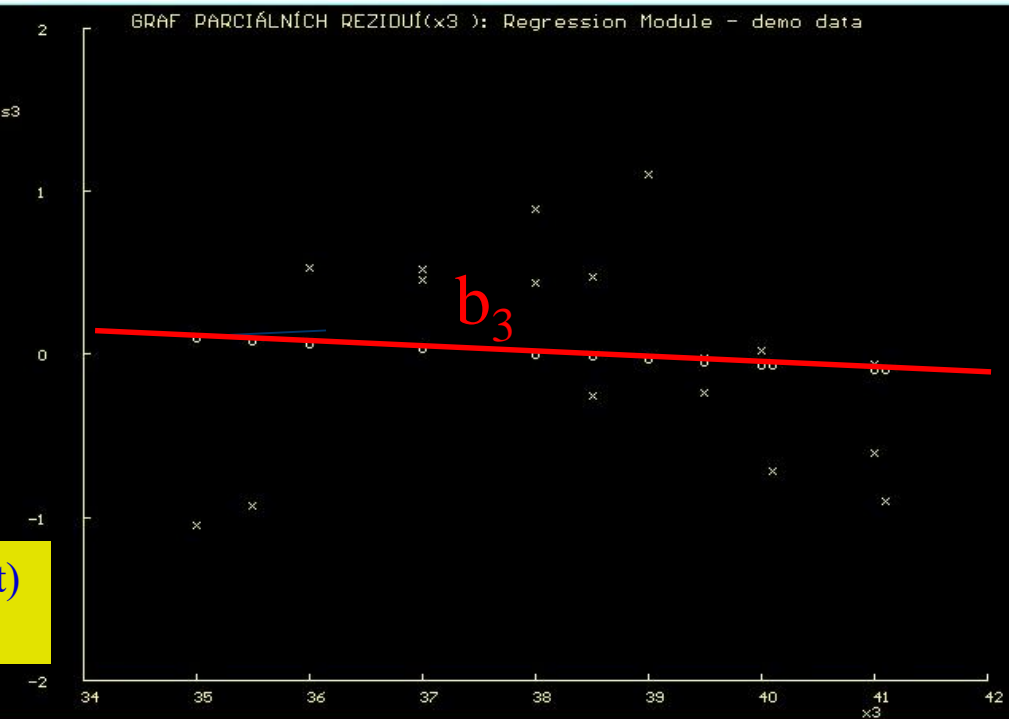
Parciální reziduální graf testuje významnost (= šikmost) směrnice b_2 a rozptýlení bodů okolo směrnice.

Zobrazení grafů komponent a reziduí

REGRESNÍ MODUL : Regresní diagnostika

Data	Metoda	Zadání	Výpočet	Výsledky	Grafy	Konec
------	--------	--------	---------	----------	-------	-------

- Graf regresní křivky
- Rezidua vs. predikce
- Residual-index plot
- Rankitové (Q-Q) grafy
- Parciální rezidua vs X[1]
- Rezid. komponenty vs X[03]
- Grafy vlivných bodů
- Autokorelace
- Heteroskedasticita



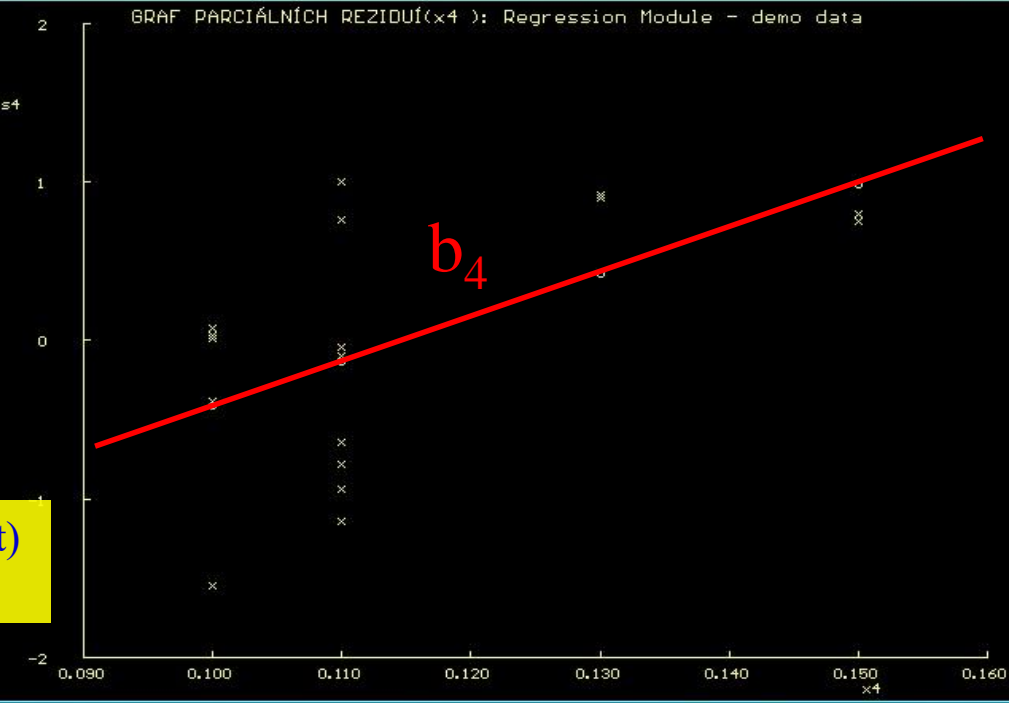
Parciální reziduální graf testuje významnost (= šikmost) směrnice b_3 a rozptýlení bodů okolo směrnice.

Zobrazení grafů komponent a reziduí

REGRESNÍ MODUL : Regresní diagnostika

Data	Metoda	Zadání	Výpočet	Výsledky	Grafy	Konec
------	--------	--------	---------	----------	-------	-------

- Graf regresní křivky
- Rezidua vs. predikce
- Residual-index plot
- Rankitové (Q-Q) grafy
- Parciální rezidua vs X[1]
- Rezid. komponenty vs X[04]
- Grafy vlivných bodů
- Autokorelace
- Heteroskedasticita



Parciální reziduální graf testuje významnost (= šikmost) směrnice b_4 a rozptýlení bodů okolo směrnice.

Zobrazení grafů komponent a reziduí

DOSBox 0.74, Cpu speed: 3000 cycles, Frameskip 0, Program: MLREG

F1=nápořveda F2=uloženi F3=čteni F4=numerický mód F5=zoom ESC=konec

Počer řádků: 17 Počer sloupců: 5 Uodorovně M606.txt

Numerický editor

	1	2	3	4
1	0	7.200000	3.500000 E+01	1.000000 E-01
5	0	7.300000	3.900000 E+01	1.100000 E-01
6	0	7.300000	3.600000 E+01	1.300000 E-01
7	0	7.400000	4.000000 E+01	1.100000 E-01
8	0	7.450000	4.010000 E+01	1.100000 E-01
9	0	7.500000	3.950000 E+01	1.000000 E-01
10	0	7.500000	4.100000 E+01	1.100000 E-01
11	0	7.200000	3.800000 E+01	1.000000 E-01
12	0	7.250000	3.850000 E+01	1.500000 E-01
13	0	7.300000	3.800000 E+01	1.100000 E-01
14	0	7.300000	3.700000 E+01	1.300000 E-01
15	0	7.300000	3.700000 E+01	1.100000 E-01
16	0	7.300000	3.700000 E+01	1.100000 E-01

Transformace hodnot v bloku (zadej funkci s parametrem x)

x=0

X=zrušení/obnovení
MEZERA=změna směru
F8=transformace
F10=transpozice

Proměnnou x odstraníte klikem na S, F8, napišete x*0, Enter, a klávesou Delete postupně smažete všechny nuly ve sloupci.

DOSBox 0.74, Cpu speed: 3000 cycles, Frameskip 0, Program: MLREG

F1=nápořveda F2=uloženi F3=čteni F4=numerický mód F5=zoom ESC=konec

Řádek 17 Sloupec 5 Insert NUM M606.txt

7.200000	1.8
7.200000	1.9
7.200000	2.0
7.250000	2.2
7.300000	2.2
7.300000	2.3
7.400000	2.3
7.450000	2.3
7.500000	2.5
7.500000	2.5
7.200000	2.0
7.250000	2.2
7.300000	2.2
7.300000	2.3
7.400000	2.3
7.450000	2.3
7.500000	2.3

Po odstranění x_1, x_3, x_4 zůstanou pouze 1 nezávisle proměnná x_2 a úsek b_0 a také závisle proměnná y .

DOSBox 0.74, Cpu speed: 3000 cycles, Frameskip 0, Program: MLREG

U V S L E D K Y

(3) ODHADY PARAMETRŮ A TESTY ÚZNAMNOSTI:

Parametr	Odhad	Směrodatná odchylka	Test $H_0: B[j] = 0$ vs. $H_A: B[j] <> 0$	t-kriterium	hypoteza H_0 je	Hlad. úzj.
B[0]	-9.5960E+01	1.5792E+01	-6.0764E+00		Zamítnuta	0.000
B[1]	1.6166E+01	2.1527E+00	7.5098E+00		Zamítnuta	0.000

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R	: 8.8877E-01
Koeficient determinace, R^2	: 7.8991E-01
Predikovaný korelační koeficient, R_p^2	: 8.5518E-01
Střední kvadratická chyba predikce, MEP	: 1.0919E+00
Akaikeho informační kritérium, AIC	: 1.3139E+00

(5) ANALÝZA KLASICKÝCH REZIDUÍ:

Bod	Meřená hodnota	Predikovaná hodnota	Směrodatná odchylka	Klasické reziduum	Relativní reziduum
1	7.2000E+01	7.2125E+01	3.0100E-01	7.5555E-01	3.4331E+00

Porovnejte nyní diagnostiky R, R^2 , MEP a AIC proti původnímu modelu a také novou těsnost proložení.

DOSBox 0.74, Cpu speed: 3000 cycles, Frameskip 0, Program: MLREG

U V S L E D K Y

(6) TESTOVÁNÍ REGRESNÍHO TRIPLETU (DATA + MODEL + METODA):

Fisher-Snedocorův test významnosti regrese, F	: 5.6397E+01
Tabulkový kvantil, $F(1-\alpha, m-1, n-m)$: 4.5431E+00
Závěr: Navržený model je přijat jako významný.	
Spočtená hladina významnosti	: 0.000
Scottovo kritérium multikolinearity, M	: -1.2598E-16
Závěr: Navržený model je korektní.	
Cook-Weisbergův test heteroskedasticity, Sf	: 5.6250E+01
Tabulkový kvantil, $\chi^2(1-\alpha, 1)$: 3.8415E+00
Závěr: Rezidua vykazují heteroskedasticitu.	
Spočtená hladina významnosti	: 0.000
Jarque-Berraův test normality reziduí, L(e)	: 8.5625E-01
Tabulkový kvantil, $\chi^2(1-\alpha, 2)$: 5.9915E+00
Závěr: Normalita je přijata.	
Spočtená hladina významnosti	: 0.652

Porovnejte nyní diagnostiky regresního tripletu u nového modelu proti diagnostikám původního modelu.

F1=nápořveda F2=uloženi F3=čteni F4=numerický mód F5=zoom ESC=konec

Rádek	1	Sloupec	79	Insert	NUM	M606.txt
M606x1	M606x2	M606x3	M606x4			
2.100000 E+01	7.200000		1.000000 E-01			
2.120000 E+01	7.200000		1.100000 E-01			
2.100000 E+01	7.200000		1.000000 E-01			
2.200000 E+01	7.250000		1.500000 E-01			
2.100000 E+01	7.300000		1.100000 E-01			
2.200000 E+01	7.300000		1.300000 E-01			
2.200000 E+01	7.400000		1.100000 E-01			
2.250000 E+01	7.450000		1.100000 E-01			
2.250000 E+01	7.500000		1.000000 E-01			
2.300000 E+01	7.500000		1.100000 E-01			
2.200000 E+01	7.200000		1.000000 E-01			
2.300000 E+01	7.250000		1.500000 E-01			
2.200000 E+01	7.300000		1.100000 E-01			
2.300000 E+01	7.300000		1.300000 E-01			
2.400000 E+01	7.400000		1.100000 E-01			
2.150000 E+01	7.450000		1.100000 E-01			
2.250000 E+01	7.500000		1.000000 E-01			

Porovnejte nyní diagnostiky R, R², MEP a AIC proti původnímu, když odstraníte pouze x3 a také těsnost proložení u nové hypotézy modelu.

U V Y S L E D K Y

(3) ODHADY PARAMETRŮ A TESTY VÝZNAMNOSTI:

Parametr	Odhad	Směrodatná odchylka	Test H0: B[j] = 0 vs. HA: B[j] <> 0	t-kriterium	hypoteza H0 je	Hlad. výz.
BI 01	-1.0408E+02	1.3027E+01	-7.9895E+00		Zamítnuta	0.000
BI 11	5.0498E-01	2.8259E-01	1.7870E+00		Akceptována	0.097
BI 21	1.5324E+01	2.0680E+00	7.4101E+00		Zamítnuta	0.000
BI 31	2.7318E+01	1.3229E+01	2.0650E+00		Akceptována	0.059

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R : 9.4842E-01
 Koeficient determinace, R² : 8.9951E-01
 Predikovaný korelační koeficient, Rp² : 9.1125E-01
 Střední kvadratická chyba predikce, MEP : 6.8939E-01
 Akaikeho informační kritérium, AIC : -7.2229E+00

(5) ANALÝZA KLASICKÝCH REZIDUÍ:

Bod	Meřená hodnota	Predikovaná hodnota	Směrodatná odchylka	Klasické reziduum	Relativní reziduum
i	yxpil	yvypil	s(yvypil)	elil	erilil
1	1.8500E+01	1.9594E+01	3.6159E-01	-1.0943E+00	-5.9149E+00
2	1.9000E+01	1.9968E+01	3.0625E-01	-9.6844E-01	-5.0970E+00

Napověda-F1 Řádek: 83 - 105 Celkem: 225 Délka: 12650

U V Y S L E D K Y

3	2.0000E+01	1.9594E+01	3.6159E-01	4.0575E-01	2.0287E+00
4	2.2000E+01	2.2231E+01	4.6949E-01	-2.3137E-01	-1.0517E+00
5	2.2500E+01	2.1400E+01	3.1734E-01	1.1001E+00	4.8895E+00
6	2.3000E+01	2.2451E+01	2.6857E-01	5.4879E-01	2.3860E+00
7	2.3500E+01	2.3437E+01	2.2553E-01	6.2746E-02	2.6700E-01
8	2.3800E+01	2.4456E+01	2.5536E-01	-6.5595E-01	-2.7561E+00
9	2.5000E+01	2.4949E+01	3.3179E-01	5.1025E-02	2.0410E-01
10	2.5500E+01	2.5475E+01	3.2517E-01	2.5349E-02	9.9407E-02
11	2.0500E+01	2.0099E+01	4.1894E-01	4.0076E-01	1.9549E+00
12	2.2500E+01	2.2736E+01	4.6032E-01	-2.3635E-01	-1.0505E+00
13	2.2800E+01	2.1905E+01	1.9906E-01	8.9516E-01	3.9261E+00
14	2.3500E+01	2.2956E+01	3.1562E-01	5.4381E-01	2.3141E+00
15	2.3900E+01	2.4447E+01	5.1703E-01	-5.4722E-01	-2.2896E+00
16	2.3100E+01	2.3951E+01	3.8966E-01	-8.5096E-01	-3.6838E+00
17	2.5500E+01	2.4949E+01	3.3179E-01	5.5102E-01	2.1609E+00

Rezidualní součet čtverců, RSC : 6.9431E+00
 Průměr absolutních hodnot reziduí, Me : 5.3936E-01
 Průměr relativních reziduí, Mer : 2.4750E+00
 Odhad reziduálního rozptylu, s²(e) : 5.3408E-01
 Odhad směrodatné odchylky reziduí, s(e) : 7.3081E-01
 Odhad šikmosti reziduí, g1(e) : -1.4258E-01
 Odhad špičatosti reziduí, g2(e) : 1.9568E+00

Napověda-F1 Řádek: 106 - 128 Celkem: 225 Délka: 12650

U V Y S L E D K Y

(6) TESTOVÁNÍ REGRESNÍHO TRIPLETU (DATA + MODEL + METODA):

Fisher-Snedocorův test významnosti regrese, F : 3.8788E+01
 Tabulkový kvantil, F(1-alpha, m-1, n-m) : 3.4105E+00
 Závěr: Navržený model je přijat jako významný.
 Spočtená hladina významnosti : 0.000

Scottovo kritérium multikolinearity, M : 3.0211E-01
 Závěr: Navržený model je korektní.

Cook-Weisbergův test heteroskedasticity, Sf : 2.6297E+01
 Tabulkový kvantil, Chi²(1-alpha, 1) : 3.8415E+00
 Závěr: Rezidua vykazují heteroskedasticitu.
 Spočtená hladina významnosti : 0.000

Jarque-Berraův test normality reziduí, L(e) : 8.2851E-01
 Tabulkový kvantil, Chi²(1-alpha, 2) : 5.9915E+00
 Závěr: Normalita je přijata.
 Spočtená hladina významnosti : 0.661

Waldův test autokorelace, Wa : 2.3363E-01
 Tabulkový kvantil, Chi²(1-alpha, 1) : 3.8415E+00
 Závěr: Rezidua nejsou autokorelována.
 Spočtená hladina významnosti : 0.629

Napověda-F1 Řádek: 131 - 153 Celkem: 225 Délka: 12650

Závěry a vysvětlení výstupu úlohy M606

Úkoly

Odpovědi, závěry a vysvětlení

(1) Postavte vícerozměrný lineární regresní model a vyšetřete regresní triplet.

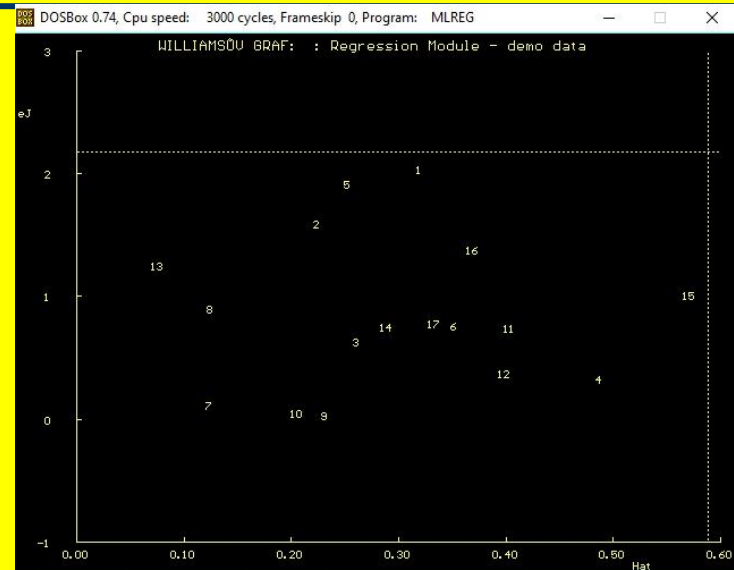
Výše uvedené výstupy **A, B, C** ukazují hledání regresního modelu dle statistické významnosti jednotlivých parametrů

(2) Pomocí parciálních reziduálních grafů vyšetřete statistickou významnost jednotlivých parametrů.

Výše uvedené parciální reziduální grafy **b₁, b₂, b₃, b₄** ukazují sklon směrnice v souladu se statistickou významností parametru **b**.

(3) Jsou v datech vlivné body a je nutné odstranit vybočující hodnoty?

Williamsův graf v Kritice dat ukazuje, že v datech nejsou odlehlé body.



(4) Jak řešíme úlohu v případě porušení předpokladů MNČ?

V ADSTATu není možné, musíme zvolit QCEXPRT (ADSTAT pod Windows).

Úloha M6.55 *Vliv škodlivin ovzduší na koncentraci ozonu v ionosféře*

Zadání: Na stanovišti byly automatickými analyzátory po dobu 24 hodin proměřovány škodliviny a faktory ovzduší.

Úkoly:

- (1) Navrhněte vhodný regresní model pro závislost koncentrace ozonu v ionosféře na jednotlivých sledovaných proměnných.
- (2) Jsou v datech vlivné a odlehlé body?
- (3) Který z parametrů je statisticky významný?
- (4) Vyšetřením regresního tripletu proved'te kritiku dat a pomocí parciálních reziduálních grafů a Studentova t-testu vyšetřete statistickou významnost všech parametrů modelu.

Data: x_1 značí obsah SO_2 [$\mu\text{g}/\text{m}^3$], x_2 obsah CO [$\mu\text{g}/\text{m}^3$], x_3 obsah NO_2 [$\mu\text{g}/\text{m}^3$], x_4 obsah benzenu [$\mu\text{g}/\text{m}^3$], x_5 množství prachu PM_{10} [$\mu\text{g}/\text{m}^3$], x_6 směr větru [stupeň], x_7 rychlost větru [m/s], x_8 teplota [$^{\circ}\text{C}$], x_9 relativní vlhkost [%], y značí koncentraci ozonu O_3 [$\mu\text{g}/\text{m}^3$].

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	y
17	314	12	2.2	73	26	1.8	9.2	70	38
...
31	496	25	5.4	37	31	0.7	10	59	29

DOSBox 0.74, Cpu speed: 3000 cycles, Frameskip 0, Program: MLREG

F1=nápořveda F2=uloženi F3=čteni F4=numerický mód F5=zo
 Počet řádků: 24 Počet sloupců: 10 Uodorovně M655

DC

Numerický editor

	1	2	3	4
1	1.700000 E+01	3.140000 E+02	1.200000 E+01	2.200000
2	1.800000 E+01	3.580000 E+02	1.100000 E+01	2.030000
3	1.800000 E+01	3.410000 E+02	8.000000	1.870000
4	1.700000 E+01	3.010000 E+02	8.000000	6.050000 E+01
5	1.600000 E+01	3.790000 E+02	9.000000	5.000000 E-02
6	1.400000 E+01	3.270000 E+02	9.000000	5.000000 E-02
7	1.300000 E+01	3.050000 E+02	9.000000	2.070000
8	1.400000 E+01	5.360000 E+02	2.100000 E+01	1.050000
9	1.400000 E+01	5.340000 E+02	1.400000 E+01	2.210000
10	1.300000 E+01	5.420000 E+02	1.300000 E+01	2.960000
11	1.300000 E+01	4.480000 E+02	1.100000 E+01	2.710000
12	1.300000 E+01	3.540000 E+02	1.000000 E+01	3.280000
13	1.300000 E+01	3.950000 E+02	1.000000 E+01	2.380000

z
b
o

Otevíř Data, pak stisk F3, pak smazat demo.dat, napsat M655 a ukončit Esc a A.

DOSBox 0.74, Cpu speed: 3000 cycles, Frameskip 0, Program: MLREG

U Y S L E D K Y

(3) ODHADY PARAMETRŮ A TESTY VYZNAMNOSTI:

Parametr	Odhad	Směrodatná odchyľka	Test H0: B[j] = 0 vs. H1: B[j] <> 0	t-kriterium	hypoteza H0 je	Hlad. výz.
BI 01	-9.0599E+00	1.0091E+01	-8.9781E-01		Akceptována	0.384
BI 11	-3.2739E-01	2.9564E-01	-1.1074E+00		Akceptována	0.287
BI 21	1.4961E-02	1.1975E-02	1.2493E+00		Akceptována	0.232
BI 31	-4.0356E-02	2.5359E-01	-1.5914E-01		Akceptována	0.876
BI 41	3.4275E-02	4.3110E-02	7.9504E-01		Akceptována	0.440
BI 51	-1.4229E-01	4.2768E-02	-3.3270E+00		Zamítuta	0.005
BI 61	-4.0334E-02	1.2442E-02	-3.2418E+00		Zamítuta	0.006
BI 71	6.9853E+00	1.1337E+00	6.1617E+00		Zamítuta	0.000
BI 81	2.0146E+00	1.7805E-01	1.1315E+01		Zamítuta	0.000
BI 91	3.8246E-01	1.3881E-01	2.7552E+00		Zamítuta	0.015

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R : 9.9348E-01
 Koeficient determinace, R^2 : 9.8700E-01
 Predikovaný korelační koeficient, Rp^2 : 7.1424E-01
 Střední kvadratická chyba predikce, MEP : 1.1842E+02
 Akaikeho informační kritérium, AIC : 4.7491E+01

(5) ANALÝZA KLASICKÝCH REZIDUÍ:

Napověda-F1 Řádek: 164 - 186 Celkem: 340 Délka: 20963

DOSBox 0.74, Cpu speed: 3000 cycles, Frameskip 0, Program: MLREG

U Y S L E D K Y

10	1.5000E+01	1.4395E+01	1.4774E+00	6.0549E-01	4.0366E+00
11	9.0000E+00	8.9430E+00	1.9644E+00	5.7011E-02	6.3345E-01
12	7.0000E+00	5.5808E+00	1.6513E+00	1.4192E+00	2.0274E+01
13	5.0000E+00	7.2462E+00	1.4264E+00	-2.2462E+00	-4.4924E+01
14	8.0000E+00	5.6785E+00	8.9574E-01	2.3215E+00	2.9019E+01
15	8.0000E+00	6.7891E+00	1.0441E+00	1.2109E+00	1.5136E+01
16	4.0000E+00	5.7949E+00	1.1035E+00	-1.7949E+00	-4.4872E+01
17	5.0000E+00	3.8067E+00	1.0895E+00	1.1933E+00	2.3867E+01
18	7.0000E+00	8.9456E+00	1.3347E+00	-1.9456E+00	-2.7795E+01
19	5.0000E+00	2.9177E+00	1.6080E+00	2.0823E+00	4.1646E+01
20	4.0000E+00	7.1360E+00	1.6375E+00	-3.1360E+00	-7.8400E+01
21	6.0000E+00	3.5115E+00	1.5984E+00	2.4885E+00	4.1475E+01
22	1.0000E+01	1.0884E+01	1.9902E+00	-8.8405E-01	-8.8405E+00
23	2.0000E+01	2.1433E+01	1.3145E+00	-1.4335E+00	-7.1674E+00
24	2.9000E+01	2.8473E+01	2.0221E+00	5.2700E-01	1.8172E+00

Rezidualní součet čtverců, RSC : 7.5454E+01
 Průměr absolutních hodnot reziduí, Me : 1.5166E+00
 Průměr relativních reziduí, Mer : 1.7929E+01
 Odhad reziduálního rozptylu, s^2(e) : 5.3896E+00
 Odhad směrodatné odchyľky reziduí, s(e) : 2.3215E+00
 Odhad šikmosti reziduí, g1(e) : 9.0478E-02
 Odhad špičatosti reziduí, g2(e) : 1.9837E+00

Napověda-F1 Řádek: 200 - 222 Celkem: 340 Délka: 20963

DOSBox 0.74, Cpu speed: 3000 cycles, Frameskip 0, Program: MLREG

U Y S L E D K Y

(6) TESTOVÁNÍ REGRESNÍHO TRIPLETU (DATA + MODEL + METODA):

Fisher-Snedocorův test významnosti regrese, F : 1.1806E+02
 Tabulkový kvantil, F(1-alpha, m-1, n-m) : 2.6458E+00
 Závěr: Navržený model je přijat jako významný.
 Spočtená hladina významnosti : 0.000

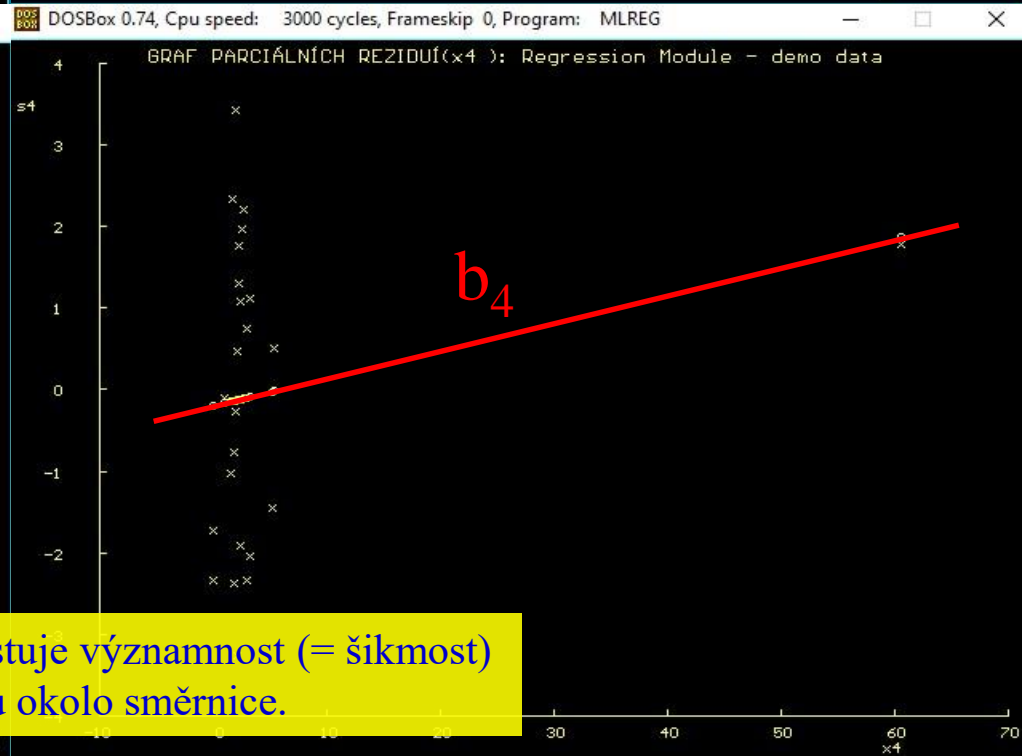
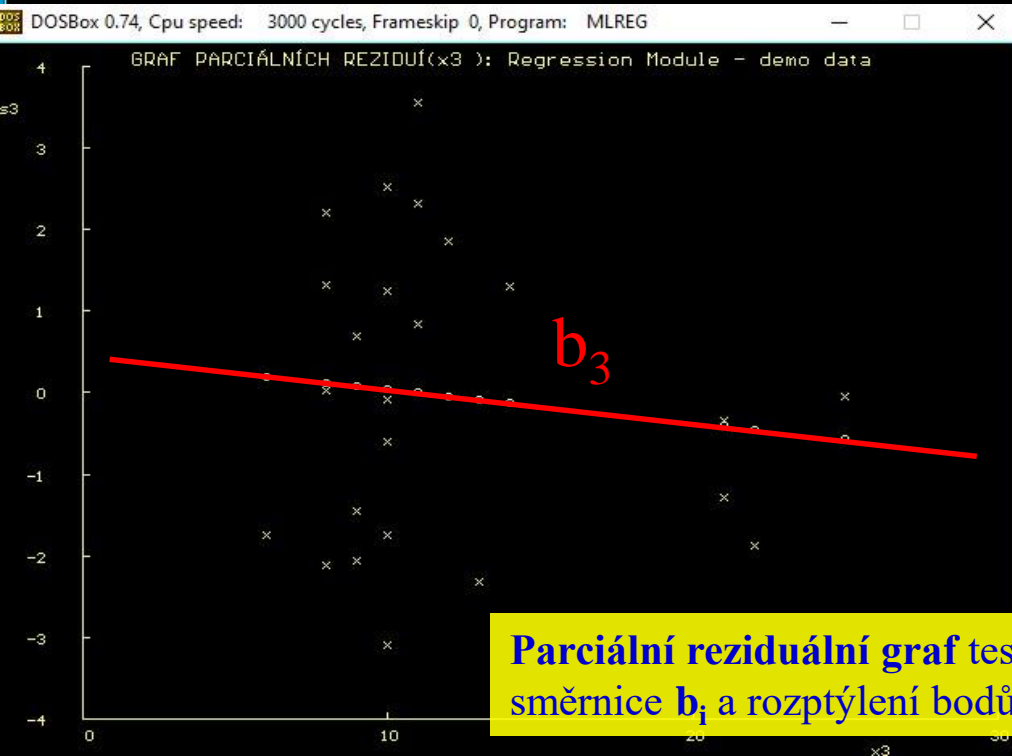
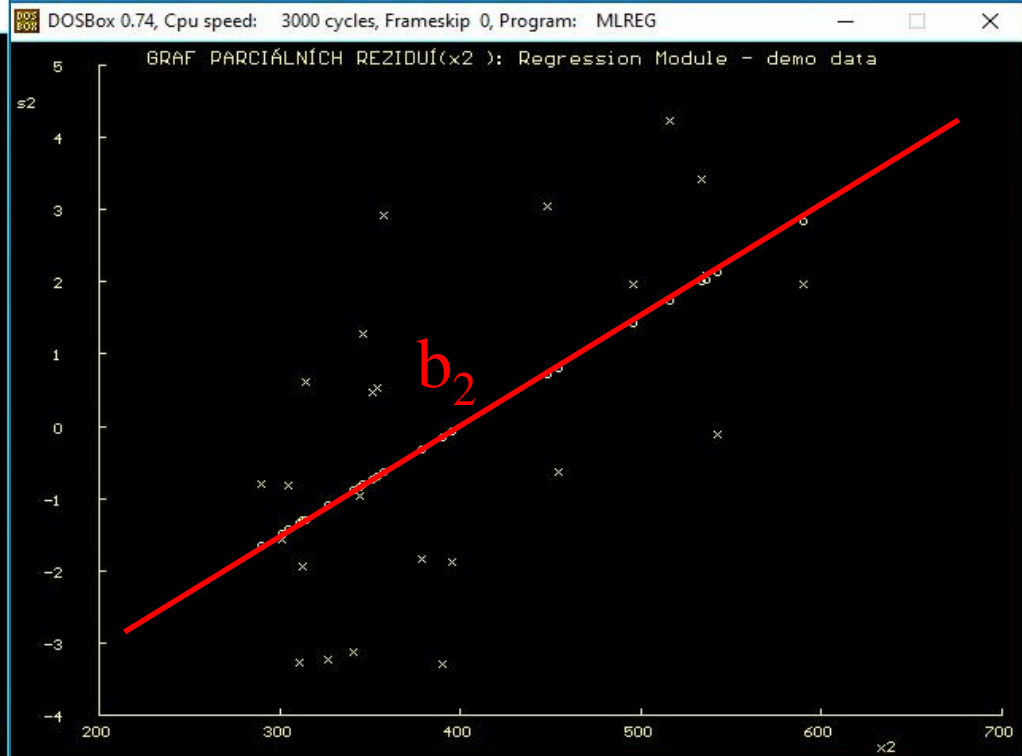
Scottovo kritérium multikolinearity, M : 6.8504E-01
 Závěr: Navržený model není korektní.

Cook-Weisbergův test heteroskedasticity, Sf : 8.0338E+00
 Tabulkový kvantil, Chi^2(1-alpha, 1) : 3.8415E+00
 Závěr: Rezidua vykazují heteroskedasticitu.
 Spočtená hladina významnosti : 0.005

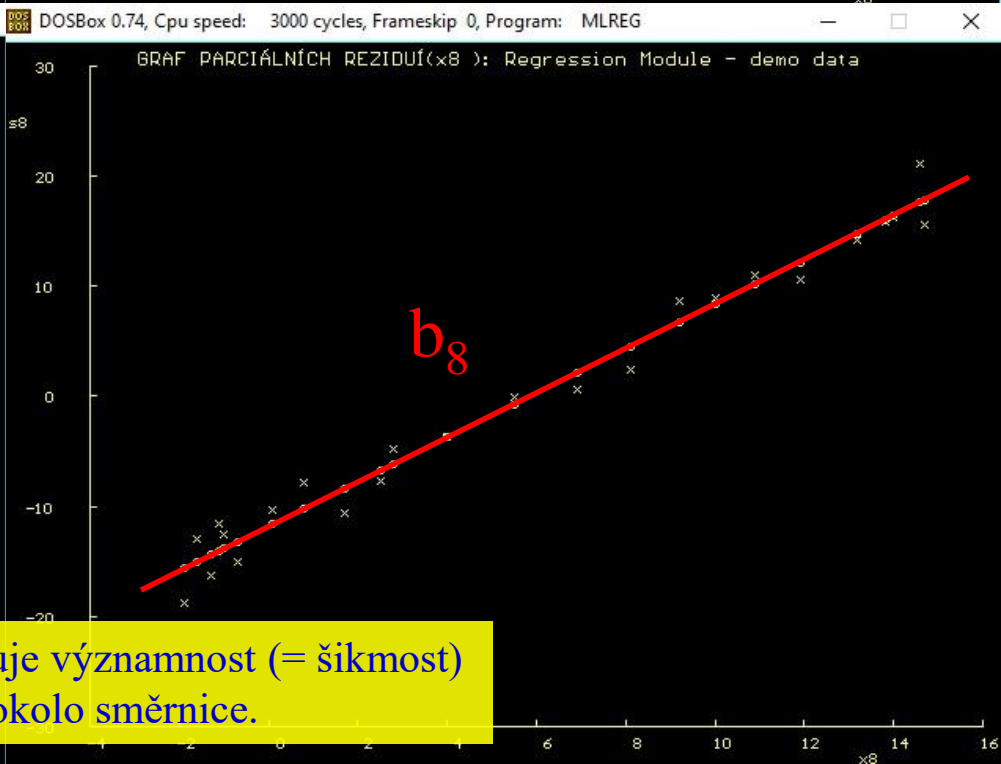
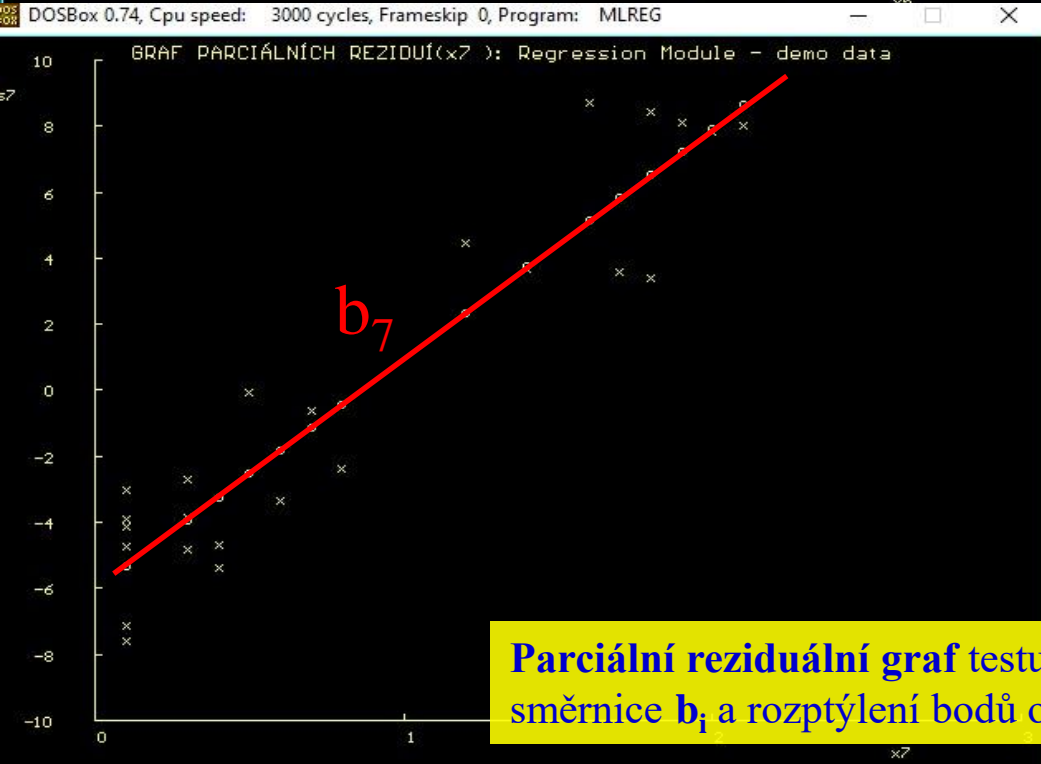
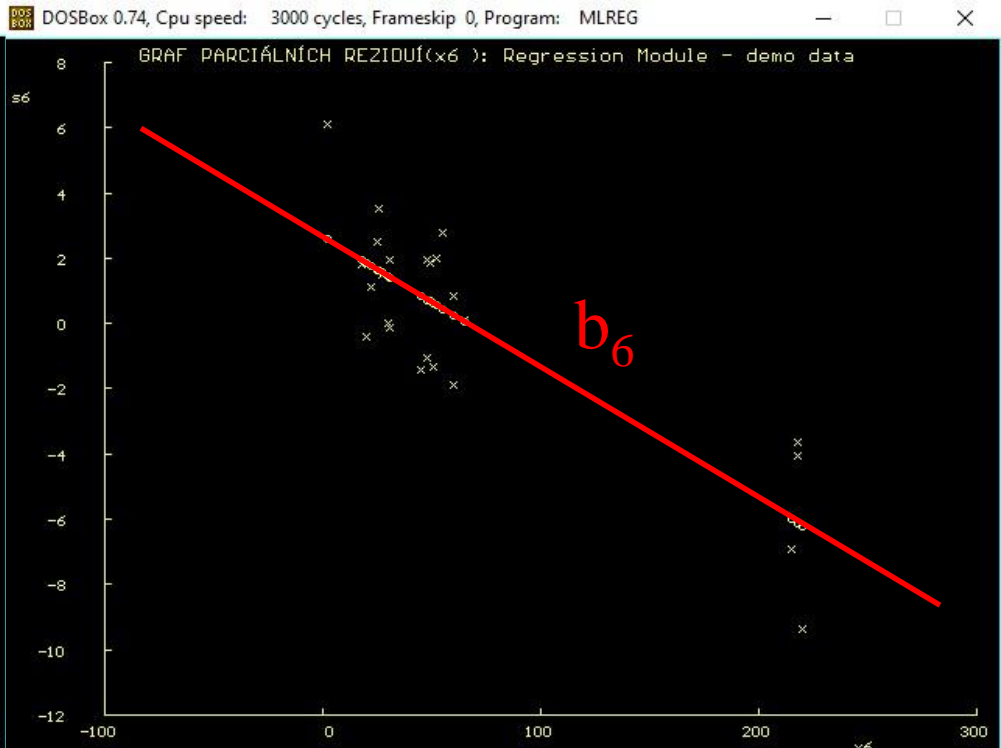
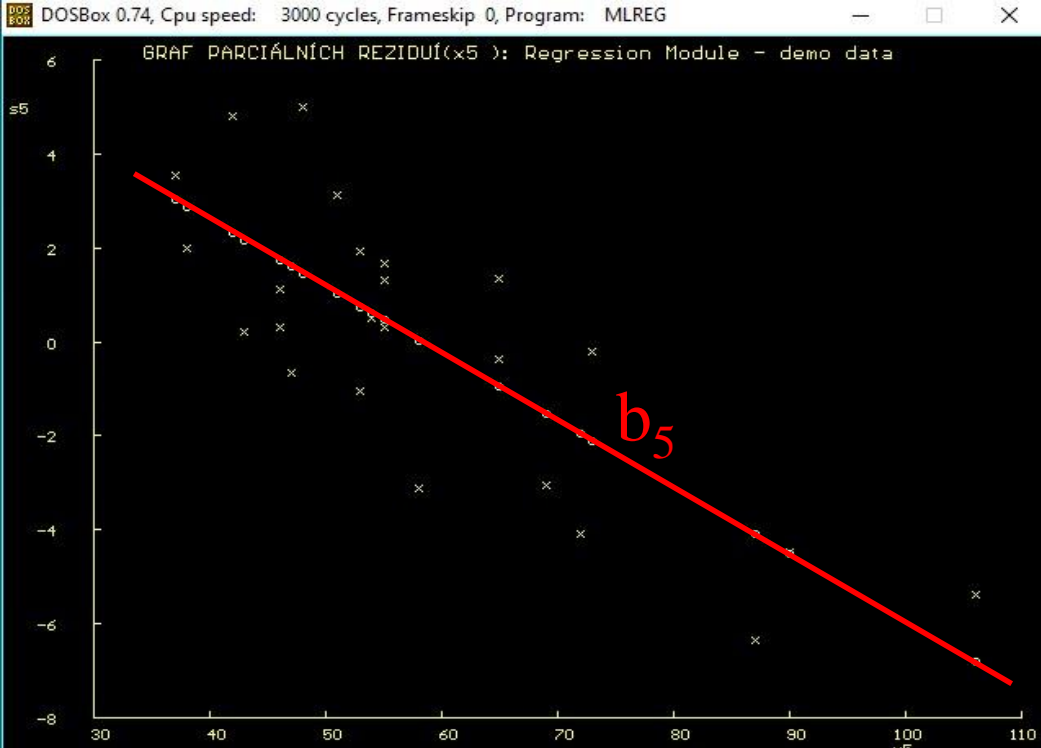
Jarque-Berraův test normality reziduí, L(e) : 1.0656E+00
 Tabulkový kvantil, Chi^2(1-alpha, 2) : 5.9915E+00
 Závěr: Normalita je přijata.
 Spočtená hladina významnosti : 0.587

Waldův test autokorelace, Wa : 7.7695E+00
 Tabulkový kvantil, Chi^2(1-alpha, 1) : 3.8415E+00
 Závěr: Rezidua jsou autokorelována.
 Spočtená hladina významnosti : 0.005

Napověda-F1 Řádek: 225 - 247 Celkem: 340 Délka: 20963



Parciální reziduální graf testuje významnost (= šikmost) směrnice b_i a rozptýlení bodů okolo směrnice.



Parciální reziduální graf testuje významnost (= šikmost) směrnice b_i a rozptýlení bodů okolo směrnice.

F1=nápověda F2=uložení F3=čtení F4=numerický mód F5=zoom ESC=konec
 Počet řádků: 24 Počet sloupců: 10 Vodorovně M655.txt

Numerický editor

	1	2	3	4
1	0	3.140000 E+02	1.200000 E+01	2.200000
Transformace hodnot v bloku (zadej funkci s parametrem x)				
x=0				
5	0	3.580000 E+02	1.100000 E+01	2.030000
6	0	3.410000 E+02	8.000000	1.870000
7	0	3.010000 E+02	8.000000	6.050000 E+01
8	0	3.790000 E+02	9.000000	5.000000 E-02
9	0	3.270000 E+02	9.000000	5.000000 E-02
10	0	3.050000 E+02	9.000000	2.070000
11	0	5.360000 E+02	2.100000 E+01	1.050000
12	0	5.340000 E+02	1.400000 E+01	2.210000
13	0	5.420000 E+02	1.300000 E+01	2.960000
14	0	4.480000 E+02	1.100000 E+01	2.710000
15	0	3.540000 E+02	1.000000 E+01	3.280000
16	0	3.950000 E+02	1.000000 E+01	2.380000

Proměnnou x odstraní klikem na S, F8, napište x*0, Enter, a klávesou Delete postupně smažte všechny nuly ve sloupci.

U V Y S L E D K Y

(3) ODHADY PARAMETRŮ A TESTY VÝZNAMNOSTI:

Parametr	Odhad	Směrodatná odchylka	Test H0: $B[j] = 0$ vs. HA: $B[j] <> 0$	t-kriterium	hypoteza H0 je	Hlad. výz.
B[0]	0.0000E+00					
B[1]	-9.9328E-02	2.8971E-02	-3.4285E+00		Zamítnuta	0.003
B[2]	-3.0278E-02	9.3315E-03	-3.2447E+00		Zamítnuta	0.004
B[3]	6.4155E+00	8.9495E-01	7.1686E+00		Zamítnuta	0.000
B[4]	1.8733E+00	1.2787E-01	1.4649E+01		Zamítnuta	0.000
B[5]	2.0153E-01	3.7917E-02	5.3150E+00		Zamítnuta	0.000

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

- Vícenásobný korelační koeficient, R : 9.9155E-01
- Koeficient determinace, R² : 9.8318E-01
- Predikovaný korelační koeficient, R_p² : 9.8467E-01
- Střední kvadratická chyba predikce, MEP : 7.3540E+00
- Akaikeho informační kritérium, AIC : 4.3669E+01

(5) Porovnejte nyní diagnostiky R, R², MEP a AIC proti původnímu, když odstraní pouze x₃ a také těsnost proložení u nové hypotézy modelu.

U V Y S L E D K Y

10	1.5000E+01	1.4965E+01	7.8758E-01	3.5079E-02	2.3386E-01
11	9.0000E+00	9.3928E+00	1.0970E+00	-3.9283E-01	-4.3648E+00
12	7.0000E+00	5.6738E+00	1.3795E+00	1.3262E+00	1.8945E+01
13	5.0000E+00	5.9140E+00	9.6092E-01	-9.1397E-01	-1.8279E+01
14	8.0000E+00	5.5059E+00	6.9655E-01	2.4941E+00	3.1176E+01
15	8.0000E+00	7.0860E+00	8.1843E-01	9.1399E-01	1.1425E+01
16	4.0000E+00	5.1076E+00	9.8689E-01	-1.1076E+00	-2.7689E+01
17	5.0000E+00	4.1122E+00	9.4259E-01	8.8777E-01	1.7755E+01
18	7.0000E+00	8.9739E+00	1.1926E+00	-1.9739E+00	-2.8198E+01
19	5.0000E+00	4.7240E+00	1.2074E+00	2.7600E-01	5.5201E+00
20	4.0000E+00	7.8601E+00	1.5216E+00	-3.8601E+00	-9.6502E+01
21	6.0000E+00	2.4628E+00	1.1859E+00	3.5332E+00	5.8887E+01
22	1.0000E+01	9.0601E+00	1.4839E+00	9.7988E-01	9.7988E+01
23	2.0000E+01	2.2308E+01	9.1338E-01	-2.3077E+00	-1.1539E+01
24	2.9000E+01	3.0500E+01	1.0362E+00	-1.5001E+00	-5.1726E+00

Rezidualní součet čtverců, RSC : 9.7613E+01
 Průměr absolutních hodnot reziduí, Me : 1.6512E+00
 Průměr relativních reziduí, Mer : 1.6677E+01
 Odhad reziduálního rozptylu, s²(e) : 5.1370E+00
 Odhad směrodatné odchylky reziduí, s(e) : 2.2665E+00
 Odhad šikmosti reziduí, g1(e) : -1.0518E-02
 Odhad špičatosti reziduí, g2(e) : 2.2330E+00

Napověda-F1 Řádek: 142 - 164 Celkem: 282 Délka: 16838

U V Y S L E D K Y

(6) TESTOVÁNÍ REGRESNÍHO TRIPLETU (DATA + MODEL + METODA):

- Fisher-Snedocorův test významnosti regrese, F : 2.7759E+02
- Tabulkový kvantil, F(1-alpha, m-1, n-m) : 2.8951E+00
- Závěr: Navržený model je přijat jako významný.
- Spočtená hladina významnosti : 0.000
- Scottovo kritérium multikolinearity, M : 6.2859E-01
- Závěr: Navržený model není korektní.
- Cook-Weisbergův test heteroskedasticity, Sf : 2.8918E+01
- Tabulkový kvantil, Chi²(1-alpha, 1) : 3.8415E+00
- Závěr: Rezidua vykazují heteroskedasticitu.
- Spočtená hladina významnosti : 0.000
- Jarque-Berraův test normality reziduí, L(e) : 5.8877E-01
- Tabulkový kvantil, Chi²(1-alpha, 2) : 5.9915E+00
- Závěr: Normalita je přijata.
- Spočtená hladina významnosti : 0.745
- Waldův test autokorelace, Wa : 1.5479E+00
- Tabulkový kvantil, Chi²(1-alpha, 1) : 3.8415E+00
- Závěr: Rezidua nejsou autokorelována.
- Spočtená hladina významnosti : 0.213

Napověda-F1 Řádek: 167 - 189 Celkem: 282 Délka: 16838

Závěry a vysvětlení výstupu úlohy M655

Úkoly

Odpovědi, závěry a vysvětlení

(1) Navrhněte vhodný regresní model pro závislost koncentrace ozonu v ionosféře na jednotlivých sledovaných proměnných.

(3) ODHADY PARAMETRŮ A TESTY VÝZNAMNOSTI:

Parametr	Odhad	Směrodatná odchylka	Test $H_0: B[j] = 0$ vs. $H_A: B[j] \neq 0$	t-kriterium	hypoteza H_0 je	Hlad. výz.
BI 01	-9.0599E+00	1.0091E+01	-8.9781E-01	0.384	Akceptována	0.384
BI 11	-3.2739E-01	2.9564E-01	-1.1074E+00	0.287	Akceptována	0.287
BI 21	1.4961E-02	1.1975E-02	1.2493E+00	0.232	Akceptována	0.232
BI 31	-4.0356E-02	2.5359E-01	-1.5914E-01	0.876	Akceptována	0.876
BI 41	3.4275E-02	4.3110E-02	7.9504E-01	0.440	Akceptována	0.440
BI 51	-1.4229E-01	4.2768E-02	-3.3270E+00	0.005	Zamítnuta	0.005
BI 61	-4.0334E-02	1.2442E-02	-3.2418E+00	0.006	Zamítnuta	0.006
BI 71	6.9853E+00	1.1337E+00	6.1617E+00	0.000	Zamítnuta	0.000
BI 81	2.0146E+00	1.7805E-01	1.1315E+01	0.000	Zamítnuta	0.000
BI 91	3.8246E-01	1.3881E-01	2.7552E+00	0.015	Zamítnuta	0.015

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R	: 9.9348E-01
Koeficient determinace, R^2	: 9.8700E-01
Predikovaný korelační koeficient, R_p^2	: 7.1424E-01
Střední kvadratická chyba predikce, MEP	: 1.1842E+02
Akaikeho informační kritérium, AIC	: 4.7491E+01

(3) ODHADY PARAMETRŮ A TESTY VÝZNAMNOSTI:

Parametr	Odhad	Směrodatná odchylka	Test $H_0: B[j] = 0$ vs. $H_A: B[j] \neq 0$	t-kriterium	hypoteza H_0 je	Hlad. výz.
BI 01	0.0000E+00	-----	-----	-----	-----	-----
BI 11	-9.9328E-02	2.8971E-02	-3.4285E+00	0.003	Zamítnuta	0.003
BI 21	-3.0278E-02	9.3315E-03	-3.2447E+00	0.004	Zamítnuta	0.004
BI 31	6.4155E+00	8.9495E-01	7.1686E+00	0.000	Zamítnuta	0.000
BI 41	1.8733E+00	1.2787E-01	1.4649E+01	0.000	Zamítnuta	0.000
BI 51	2.0153E-01	3.7917E-02	5.3150E+00	0.000	Zamítnuta	0.000

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R	: 9.9155E-01
Koeficient determinace, R^2	: 9.8318E-01
Predikovaný korelační koeficient, R_p^2	: 9.8467E-01
Střední kvadratická chyba predikce, MEP	: 7.3540E+00
Akaikeho informační kritérium, AIC	: 4.3669E+01

(3) Který z parametrů je statisticky významný?

Výše uvedené parciální reziduální grafy $b_1, b_2, b_3, b_4, b_5, b_6, b_7, b_8$, ukazují sklon směrnice v souladu se statistickou významností každého parametru b .

Úloha M6.04 Vliv čtyř faktorů stanovení hořčíku na absorbanci AAS – předpoklady MNČ

Zadání: Při stanovení hořčíku ve vodě metodou AAS byla sledována závislost absorbance y při $\lambda = 202.5$ nm na výšce hořáku x_1 , nasávaném množství vzorku x_2 , průtoku vzduchu x_3 a průtoku acetylenu x_4 . Předpokládejme platnost vícerozměrného lineárního regresního modelu ve tvaru $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4$.

Úkoly:

- (1) Postavte regresní model metodou regresního tripletu.
- (2) Pomocí parciálních reziduálních grafů diskutujte významnost regresních parametrů modelu.
- (3) Proveďte Cookův-Weisbergův test heteroskedasticity, autokorelační test a Jarque-Berrův test normality v datech.

Data: Výška hořáku x_1 , množství vzorku x_2 , průtok vzduchu x_3 , průtok acetylenu x_4 , měřená absorbance y :

x_1	x_2	x_3	x_4	y
10.4	4.4	12.3	37	0.25
...
9.6	3.9	11.2	31.9	0.275

DOSBox 0.74, Cpu speed: 3000 cycles, Frameskip 0, Program: MLREG

F1=nápoředa F2=uloženi F3=čteni F4=numerický mód F5=zoom

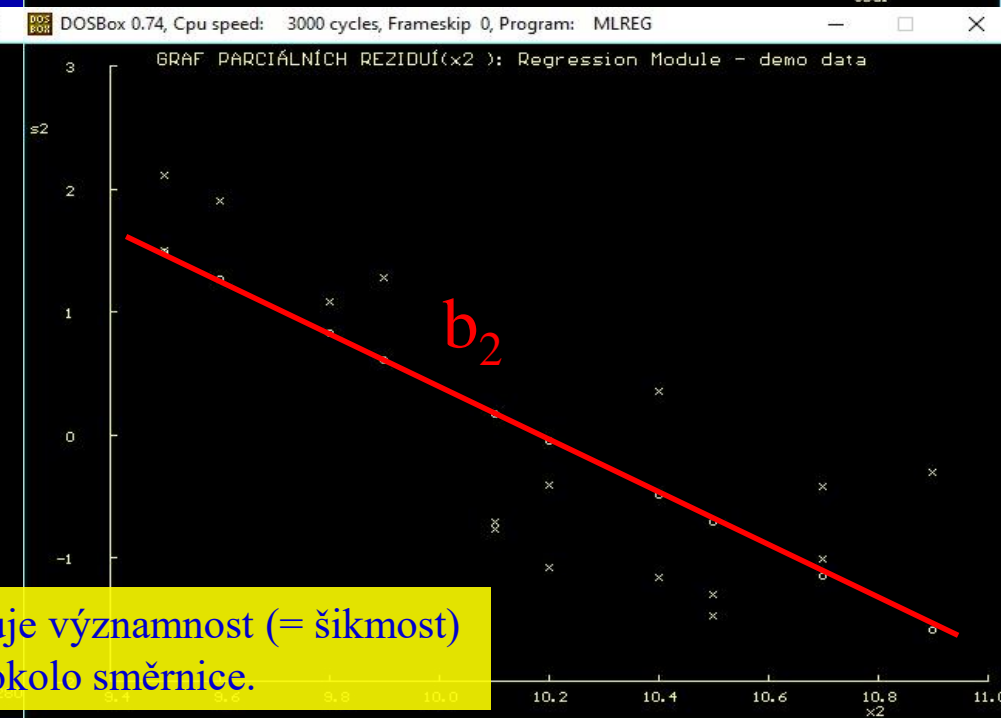
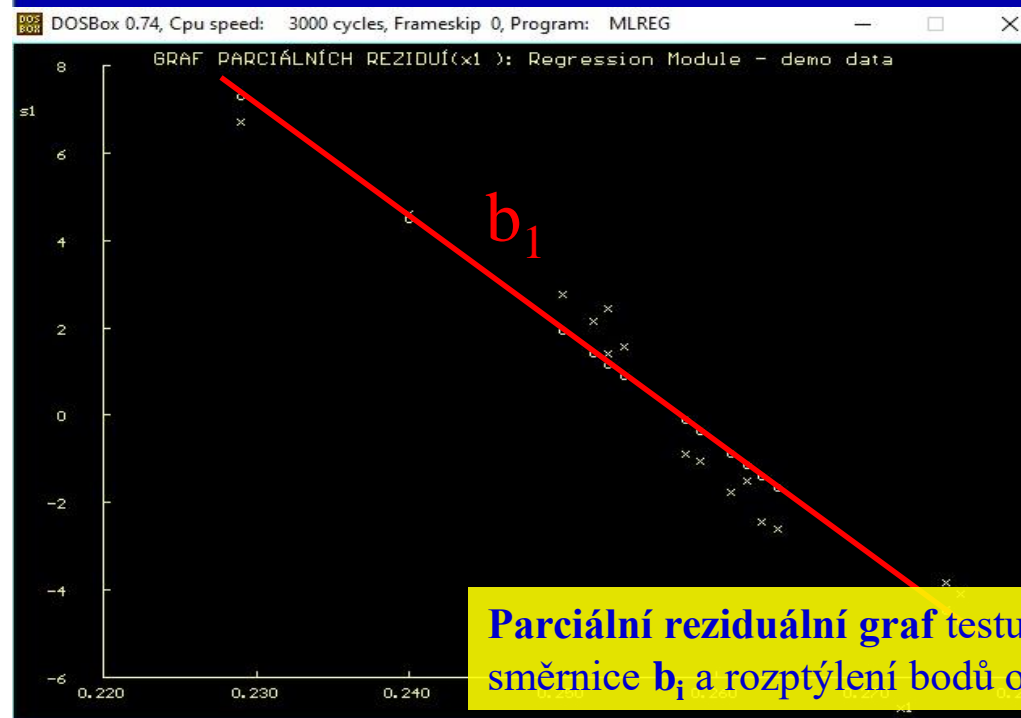
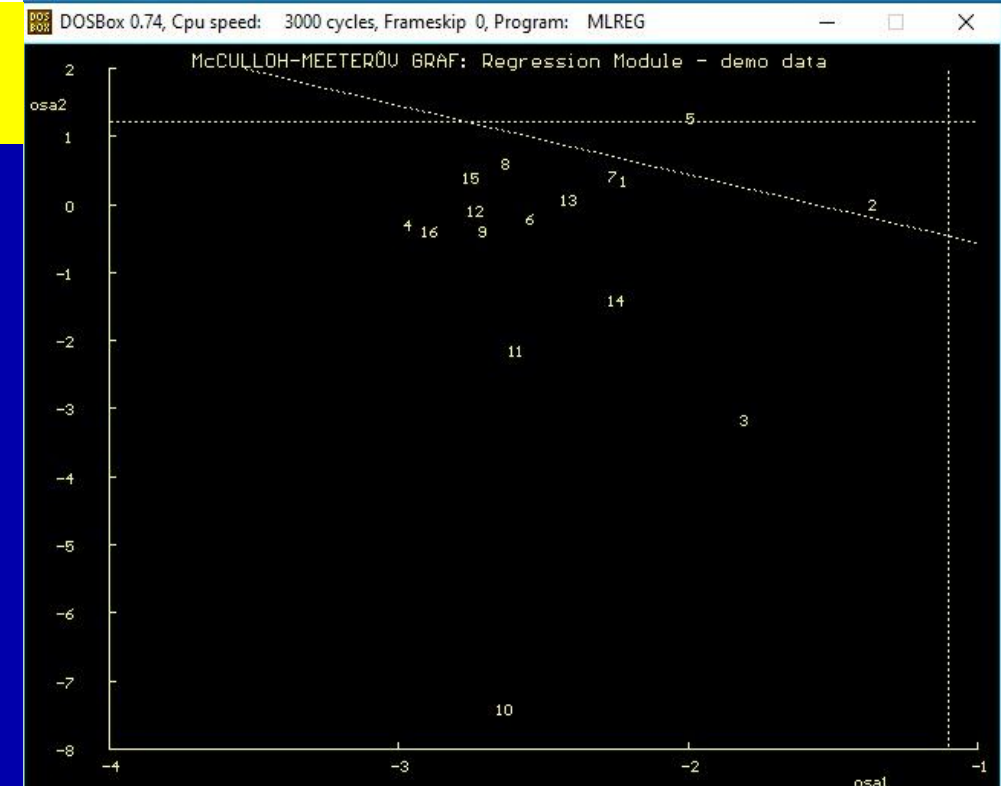
Řádek 1 Sloupec 5 Insert NUM M604.1

M604x1 M604x2 M604x3 M604x4 M604y

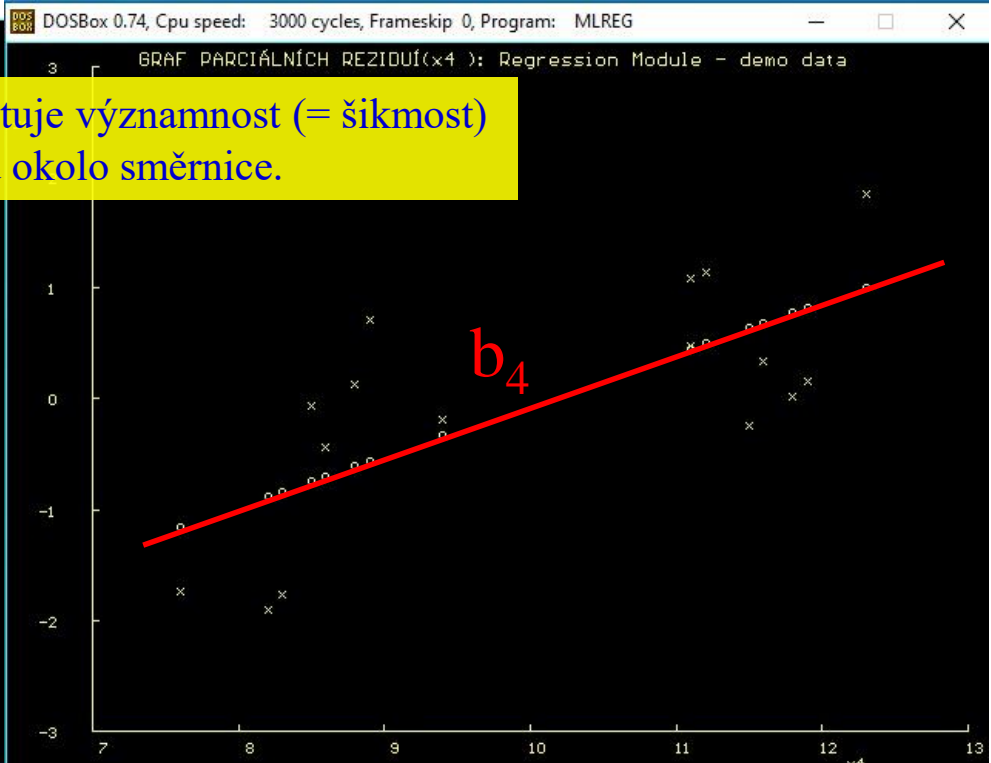
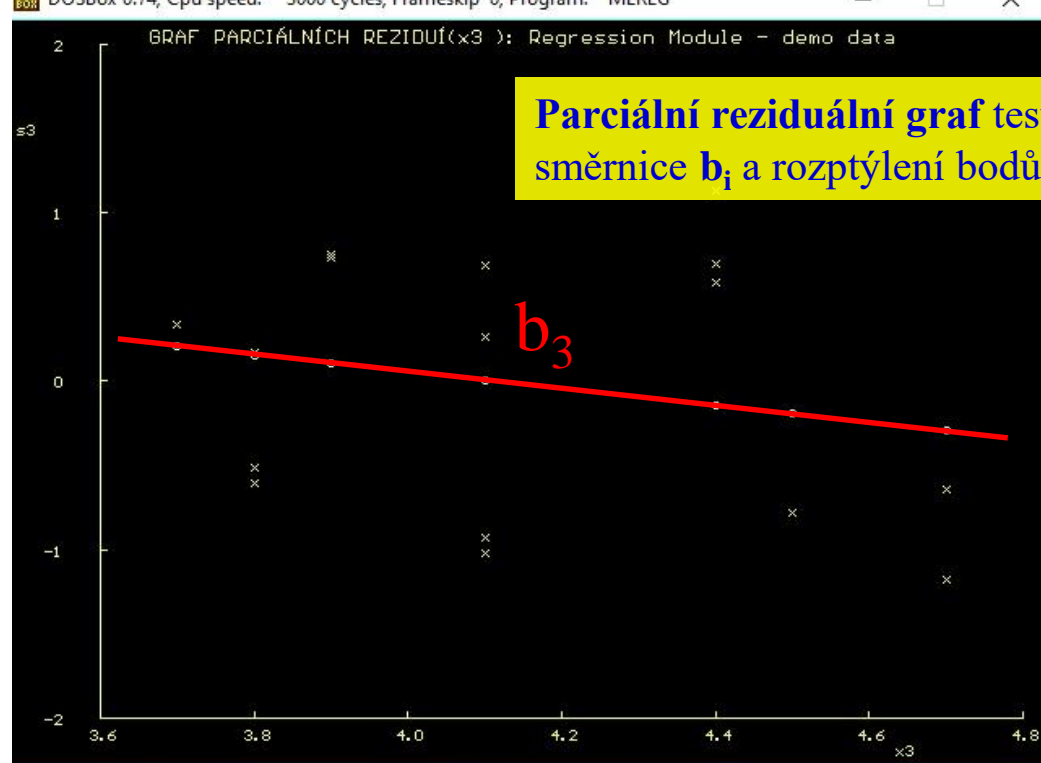
0.25000	10.40000	4.40000	12.30000	37.00000
0.22900	10.50000	4.50000	7.60000	38.50000
0.24000				
0.25400				
0.25300				
0.25900	10.40000	3.80000	11.90000	33.30000
0.26100	10.10000	4.70000	11.50000	32.60000
0.26300	10.20000	4.10000	8.20000	30.50000
0.27600	9.50000	3.90000	11.10000	31.80000
0.27500	9.50000	3.80000	11.10000	31.50000
0.25300	9.80000	4.10000	8.60000	35.40000
0.25200	10.70000	4.40000	8.80000	34.10000
0.25800	10.50000	3.80000	11.80000	33.20000
0.26200	10.20000	4.70000	11.60000	32.70000
0.26400	10.10000	4.10000	8.30000	30.60000
0.27500	9.60000	3.90000	11.20000	31.90000

čteni ze souboru: M604.txt

DC



Parciální reziduální graf testuje významnost (= šikmost) směrnice b_i a rozptýlení bodů okolo směrnice.



DOSBox 0.74, Cpu speed: 3000 cycles, Frameskip 0, Program: MLREG

DOSBox 0.74, Cpu speed: 3000 cycles, Frameskip 0, Program: MLREG

V Ý S L E D K Y

(3) ODHADY PARAMETRŮ A TESTY VÝZNAMNOSTI:

Parametr	Odhad	Směrodatná odchylka	Test H0: B[j] = 0 vs. HA: B[j] <> 0	t-kriterium	hypoteza H0 je	Hlad. výz.
BI 01	1.1939E+02	1.3927E+01	8.5723E+00	8.5723E+00	Zamítnuta	0.000
BI 11	-2.5568E+02	2.9382E+01	-8.7019E+00	-8.7019E+00	Zamítnuta	0.000
BI 21	-2.1913E+00	7.5640E-01	-2.8970E+00	-2.8970E+00	Zamítnuta	0.015
BI 31	-5.0099E-01	7.2389E-01	-6.9209E-01	-6.9209E-01	Akceptována	0.503
BI 41	4.6189E-01	1.5922E-01	2.9010E+00	2.9010E+00	Zamítnuta	0.014

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R	: 9.4982E-01
Koeficient determinace, R ²	: 9.0215E-01
Predikovaný korelační koeficient, Rp ²	: 8.8334E-01
Střední kvadratická chyba predikce, MEP	: 1.1813E+00
Akaikeho informační kritérium, AIC	: -2.7528E-01

(5) ANALÝZA KLASICKÝCH REZIDUÍ:

Bod	Meřená hodnota	Predikovaná hodnota	Směrodatná odchylka	Klasické reziduum	Relativní reziduum
i	yexp[i]	yvyp[i]	s(yvyp[i])	el[i]	er[i]
1	3.7000E+01	3.6155E+01	5.1835E-01	8.4513E-01	2.2841E+00

Napověda-F1 Řádek: 92 - 114 Celkem: 231 Délka: 12981

V Ý S L E D K Y

Odhad směrodatné odchylky reziduí, s(e)	: 8.7480E-01
Odhad šikmosti reziduí, g1(e)	: 5.8601E-02
Odhad špičatosti reziduí, g2(e)	: 1.6121E+00

(6) TESTOVÁNÍ REGRESNÍHO TRIPLETU (DATA + MODEL + METODA):

Fisher-Snedocorův test významnosti regrese, F	: 2.5354E+01
Tabulkový kvantil, F(1-alpha, m-1, n-m)	: 3.3567E+00
Závěr: Navržený model je přijat jako významný.	
Spočtená hladina významnosti	: 0.000
Scottovo kritérium multikolinearity, M	: 4.3235E-02
Závěr: Navržený model je korektní.	
Cook-Weisbergův test heteroskedasticity, Sf	: 7.2842E-01
Tabulkový kvantil, Chi ² (1-alpha, 1)	: 3.8415E+00
Závěr: Rezidua vykazují homoskedasticitu.	
Spočtená hladina významnosti	: 0.393
Jarque-Berraův test normality reziduí, L(e)	: 1.2934E+00
Tabulkový kvantil, Chi ² (1-alpha, 2)	: 5.9915E+00
Závěr: Normalita je přijata.	
Spočtená hladina významnosti	: 0.524

Napověda-F1 Řádek: 135 - 157 Celkem: 231 Délka: 12981

DOSBox 0.74, Cpu speed: 3000 cycles, Frameskip 0, Program: MLREG

F1=nápověda F2=uložení F3=čtení F4=numerický mód F5=zoom ESC=konec

Řádek	Sloupec	Insert	NUM	M604.txt	
2.500000	E-01	1.040000	E+01	1.230000 E+01	3
2.290000	E-01	1.050000	E+01	7.600000	3
2.400000	E-01	1.070000	E+01	9.400000	3
2.540000	E-01	9.900000		8.500000	3
2.590000	E-01	1.040000	E+01	1.190000 E+01	3
2.610000	E-01	1.010000	E+01	1.150000 E+01	3
2.630000	E-01	1.020000	E+01	8.200000	3
2.760000	E-01	9.500000		1.110000 E+01	3
2.750000	E-01	9.500000		1.110000 E+01	3
2.530000	E-01	9.800000		8.600000	3
2.520000	E-01	1.070000	E+01	8.800000	3
2.580000	E-01	1.050000	E+01	1.180000 E+01	3
2.620000	E-01	1.020000	E+01	1.160000 E+01	3
2.640000	E-01	1.010000	E+01	8.300000	3.
2.750000	E-01	9.600000		1.120000 E+01	3

DOSBox 0.74, Cpu speed: 3000 cycles, Frameskip 0, Program: MLREG

U Ý S L E D K Y

(3) ODHADY PARAMETRŮ A TESTY VÝZNAMNOSTI:

Parametr	Odhad	Směrodatná odchylka	Test H0: B[j] = 0 vs. HA: B[j] <> 0	t-kriterium	hypoteza H0 je	Hlad. výz.
BI 01	1.3173E+02	1.3882E+01	9.4897E+00	Zamítnuta		0.000
BI 11	-2.7735E+02	2.8184E+01	-9.8407E+00	Zamítnuta		0.000
BI 21	-3.1589E+00	7.9927E-01	-3.9522E+00	Zamítnuta		0.002
BI 31	5.5076E-01	1.4631E-01	3.7645E+00	Zamítnuta		0.003

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

- Vícenásobný korelační koeficient, R : 9.6190E-01
- Koeficient determinace, R² : 9.2525E-01
- Predikovaný korelační koeficient, Rp² : 9.1798E-01
- Střední kvadratická chyba predikce, MEP : 9.0135E-01
- Akaikeho informační kritérium, AIC : -4.7181E+00

(5) ANALÝZA KLASICKÝCH REZIDUÍ:

Bod	Meřená hodnota	Predikovaná hodnota	Směrodatná odchylka	Klasické reziduum	Relativní reziduum
i	yexp[i]	yvyp[i]	s(yvyp[i])	er[i]	er[i]
1	3.7000E+01	3.6320E+01	4.4100E-01	6.8038E-01	1.8389E+00
2	3.8500E+01	3.9239E+01	5.6304E-01	-7.3944E-01	-1.9206E+00

Nápověda-F1 Řádek: 77 - 99 Celkem: 211 Délka: 11884

Proměnnou x odstraní klikem na S, F8, napište x*0, Enter, a klávesou Delete postupně smažte všechny nuly ve sloupci.

DOSBox 0.74, Cpu speed: 3000 cycles, Frameskip 0, Program: MLREG

U Ý S L E D K Y

1	3.7000E+01	3.6320E+01	4.4100E-01	6.8038E-01	1.8389E+00
2	3.8500E+01	3.9239E+01	5.6304E-01	-7.3944E-01	-1.9206E+00
3	3.7200E+01	3.6548E+01	3.6340E-01	6.5178E-01	1.7521E+00
4	3.5300E+01	3.4697E+01	3.4768E-01	6.0323E-01	1.7089E+00
5	3.3300E+01	3.3603E+01	3.4063E-01	-3.0319E-01	-9.1047E-01
6	3.2600E+01	3.3776E+01	2.6430E-01	-1.1758E+00	-3.6069E+00
7	3.2500E+01	3.4088E+01	4.4762E-01	5.8735E-01	1.8870E+00
8	3.2500E+01	3.4088E+01	4.4762E-01	5.8735E-01	1.8870E+00
9	3.2500E+01	3.4088E+01	4.4762E-01	5.8735E-01	1.8870E+00
10	3.2500E+01	3.4088E+01	4.4762E-01	5.8735E-01	1.8870E+00
11	3.2500E+01	3.4088E+01	4.4762E-01	5.8735E-01	1.8870E+00
12	3.2500E+01	3.4088E+01	4.4762E-01	5.8735E-01	1.8870E+00
13	3.2500E+01	3.4088E+01	4.4762E-01	5.8735E-01	1.8870E+00
14	3.0600E+01	3.1181E+01	4.1523E-01	-5.8137E-01	-1.8999E+00
15	3.1900E+01	3.1307E+01	3.5373E-01	5.9281E-01	1.8583E+00

Reziduální součet čtverců, RSC : 6.4249E+00
Průměr absolutních hodnot reziduí, Me : 5.7371E-01
Průměr relativních reziduí, Mer : 1.7015E+00
Odhad reziduálního rozptylu, s²(e) : 5.8408E-01
Odhad směrodatné odchylky reziduí, s(e) : 7.6425E-01
Odhad šikmosti reziduí, g1(e) : 6.7171E-02
Odhad špičatosti reziduí, g2(e) : 1.9658E+00

Nápověda-F1 Řádek: 98 - 120 Celkem: 211 Délka: 11884

Porovnejte nyní diagnostiky R, R², MEP a AIC proti původnímu, když odstraní pouze x3 a také těsnost proložení u nové hypotézy modelu.

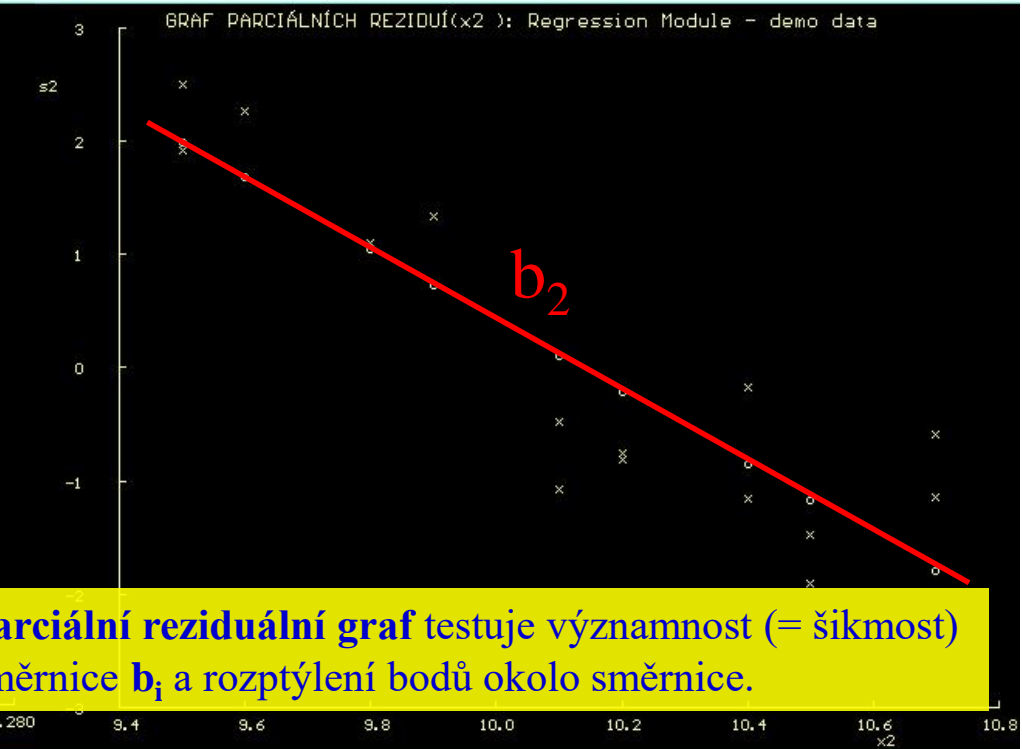
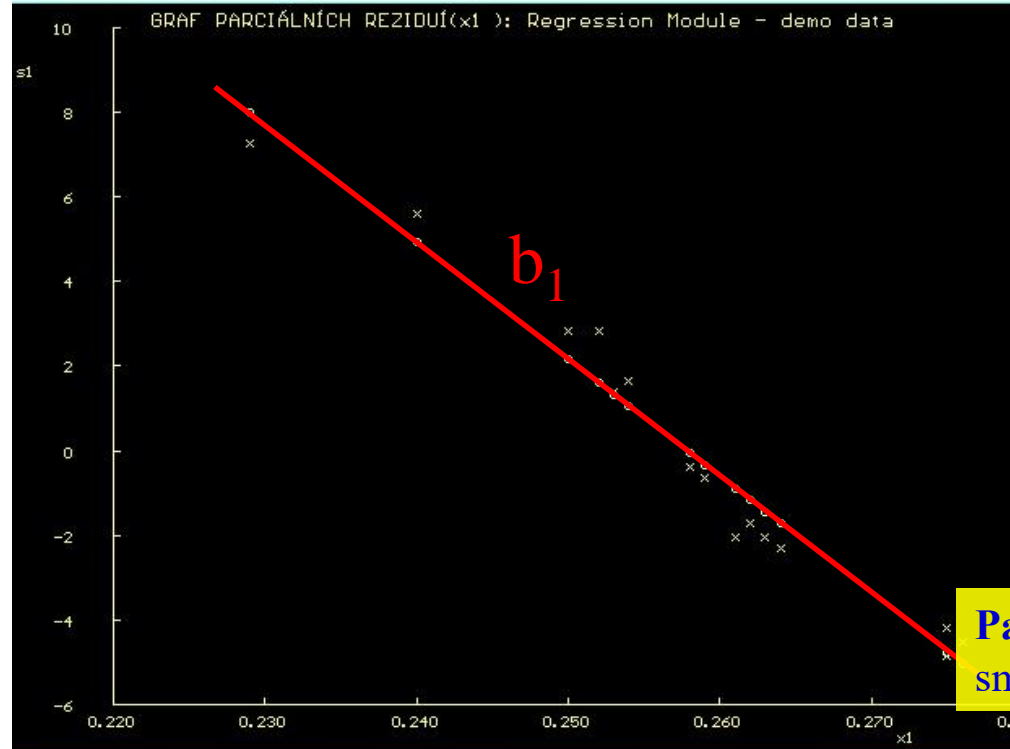
DOSBox 0.74, Cpu speed: 3000 cycles, Frameskip 0, Program: MLREG

U Ý S L E D K Y

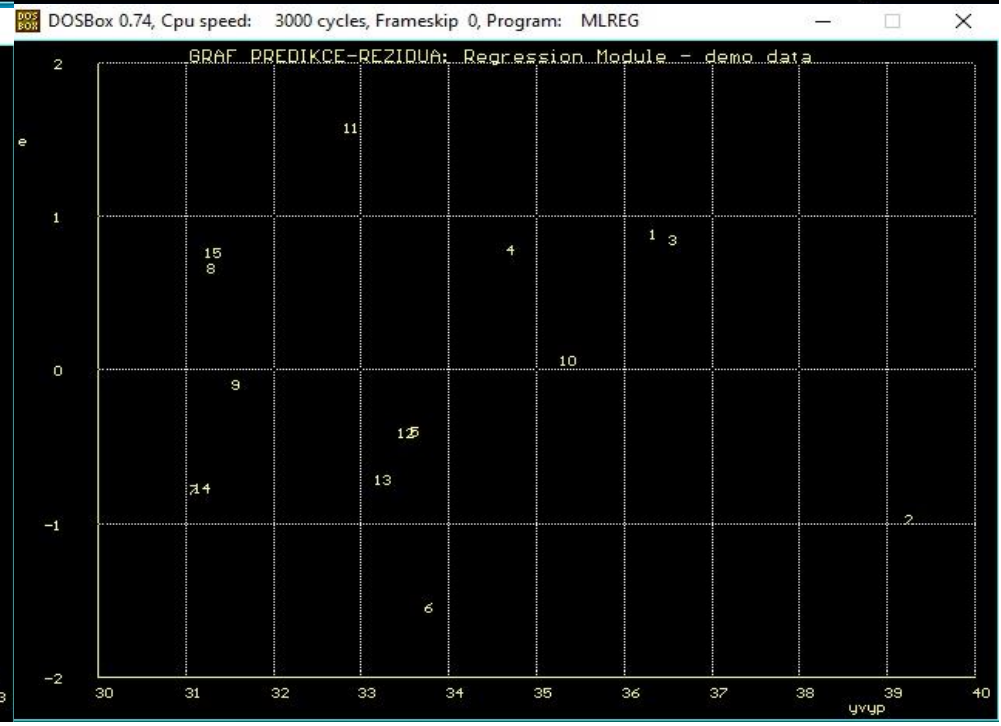
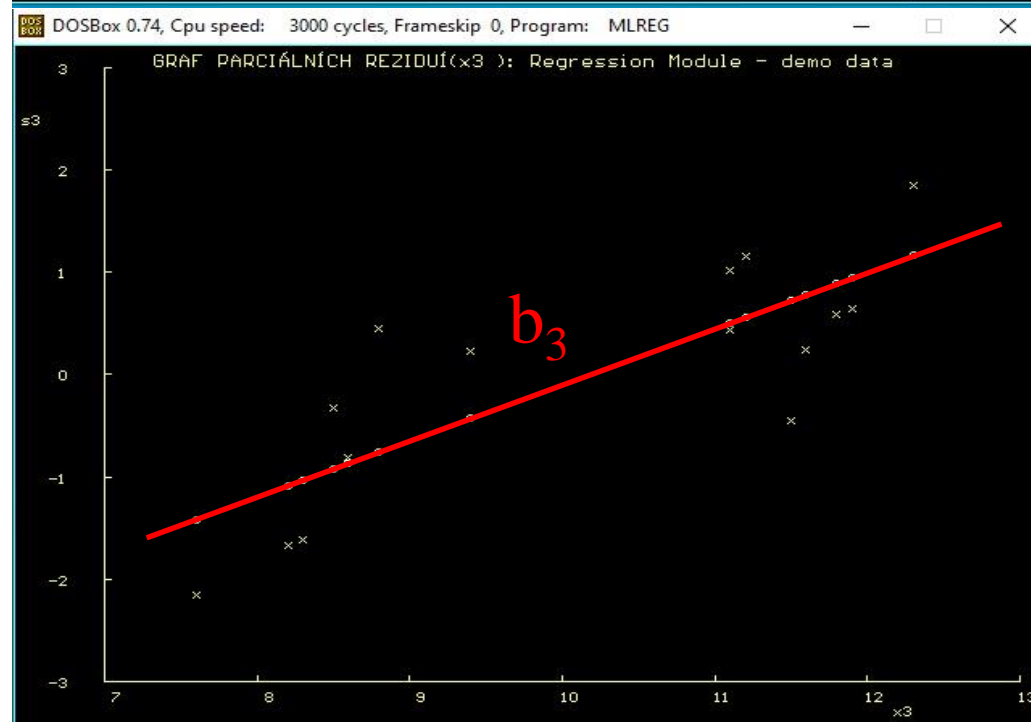
(6) TESTOVÁNÍ REGRESNÍHO TRIPLETU (DATA + MODEL + METODA):

- Fisher-Snedocorův test významnosti regrese, F : 4.5384E+01
- Tabulkový kvantil, F(1-alpha, m-1, n-m) : 3.5874E+00
- Závěr: Navržený model je přijat jako významný.
- Spočtená hladina významnosti : 0.000
- Scottovo kritérium multikolinearity, M : 3.6238E-02
- Závěr: Navržený model je korektní.
- Cook-Weisbergův test heteroskedasticity, Sf : 1.3545E+01
- Tabulkový kvantil, Chi²(1-alpha, 1) : 3.8415E+00
- Závěr: Rezidua vykazují heteroskedasticitu.
- Spočtená hladina významnosti : 0.000
- Jarque-Berraův test normality reziduí, L(e) : 6.7972E-01
- Tabulkový kvantil, Chi²(1-alpha, 2) : 5.9915E+00
- Závěr: Normalita je přijata.
- Spočtená hladina významnosti : 0.712
- Waldův test autokorelace, Wa : 2.3212E-02
- Tabulkový kvantil, Chi²(1-alpha, 1) : 3.8415E+00
- Závěr: Rezidua nejsou autokorelována.
- Spočtená hladina významnosti : 0.879

Nápověda-F1 Řádek: 123 - 145 Celkem: 211 Délka: 11884



Parciální reziduální graf testuje významnost (= šikmost) směrnice b_i a rozptýlení bodů okolo směrnice.



Závěry a vysvětlení výstupu úlohy M604

Úkoly

Odovědi, závěry a vysvětlení

(1) Postavte regresní model metodou regresního tripletu.

(3) ODHADY PARAMETRŮ A TESTY VÝZNAMNOSTI:

Parametr	Odhad	Směrodatná odchylnka	Test H0: $B_{[j]} = 0$ vs. $H_A: B_{[j]} <> 0$	t-kriterium	hypoteza H0 je	Hlad. úzj.
BI 01	1.1939E+02	1.3927E+01	8.5723E+00	Zamítuta		0.000
BI 11	-2.5568E+02	2.9382E+01	-8.7019E+00	Zamítuta		0.000
BI 21	-2.1913E+00	7.5640E-01	-2.8970E+00	Zamítuta		0.015
BI 31	-5.0099E-01	7.2389E-01	-6.9209E-01	Akceptována		0.503
BI 41	4.6189E-01	1.5922E-01	2.9010E+00	Zamítuta		0.014

(3) ODHADY PARAMETRŮ A TESTY VÝZNAMNOSTI:

Parametr	Odhad	Směrodatná odchylnka	Test H0: $B_{[j]} = 0$ vs. $H_A: B_{[j]} <> 0$	t-kriterium	hypoteza H0 je	Hlad. úzj.
BI 01	1.3173E+02	1.3882E+01	9.4897E+00	Zamítuta		0.000
BI 11	-2.7735E+02	2.8184E+01	-9.8407E+00	Zamítuta		0.000
BI 21	-3.1589E+00	7.9927E-01	-3.9522E+00	Zamítuta		0.002
BI 31	5.5076E-01	1.4631E-01	3.7645E+00	Zamítuta		0.003

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R	: 9.4982E-01
Koeficient determinace, R^2	: 9.0215E-01
Predikovaný korelační koeficient, R_p^2	: 8.8334E-01
Střední kvadratická chyba predikce, MEP	: 1.1813E+00
Akaikeho informační kritérium, AIC	: -2.7528E-01

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R	: 9.6190E-01
Koeficient determinace, R^2	: 9.2525E-01
Predikovaný korelační koeficient, R_p^2	: 9.1798E-01
Střední kvadratická chyba predikce, MEP	: 9.0135E-01
Akaikeho informační kritérium, AIC	: -4.7181E+00

(2) Pomocí parciálních reziduálních grafů diskutujte významnost regresních parametrů modelu.

Výše uvedené parciální reziduální grafy b_1, b_2, b_3, b_4 ukazují sklon směrnice v souladu se statistickou významností každého parametru b .

(3) Proveďte Cookův-Weisbergův test heteroskedasticity, autokorelační test a Jarque-Berrův test normality v datech.

DOSBox 0.74, Cpu speed: 3000 cycles, Frameskip 0, Program: MLREG

U V Ě S L E D K Y

(6) TESTOVÁNÍ REGRESNÍHO TRIPLETU (DATA + MODEL + METODA):

Fisher-Snedecorův test významnosti regrese, F	: 4.5384E+01
Tabulkový kvantil, $F(1-\alpha, m-1, n-m)$: 3.5874E+00
Závěr: Navržený model je přijat jako významný.	
Spočtená hladina významnosti	: 0.000
Scottovo kritérium multikolinearity, M	: 3.6238E-02
Závěr: Navržený model je korektní.	
Cook-Weisbergův test heteroskedasticity, Sf	: 1.3545E+01
Tabulkový kvantil, $\chi^2(1-\alpha, 1)$: 3.8415E+00
Závěr: Rezidua vykazují heteroskedasticitu.	
Spočtená hladina významnosti	: 0.000
Jarque-Berraův test normality reziduí, L(e)	: 6.7972E-01
Tabulkový kvantil, $\chi^2(1-\alpha, 2)$: 5.9915E+00
Závěr: Normalita je přijata.	
Spočtená hladina významnosti	: 0.712
Waldův test autokorelace, Wa	: 2.3212E-02
Tabulkový kvantil, $\chi^2(1-\alpha, 1)$: 3.8415E+00
Závěr: Rezidua nejsou autokorelována.	
Spočtená hladina významnosti	: 0.879

Napověda-F1 Řádek: 123 - 145 Celkem: 211 Délka: 11884