

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

1. Otevřete ADSTAT červená titulní stránka.
 2. **Enter** a otevře se menu, najedťe dolu na **Lineární regrese** a **Enter**.
 3. V menu **Data** klik na **Enter**, a pak **F3** a nyní smazat **demo.dat** a zadejte data úlohy **P608a** nebo **P608a.txt** a ukončete **Enter**.
 4. Klik na **Esc** a **A**.
 5. **Enter** otevře **Metoda**. Šípkou dolu zvolte **Regresní diagnostika** a pak **Enter**.
 6. **Enter** v menu **Zadání** otevře **Podmínky**. V **Název** napište nadpis úlohy *P608a* a své jméno *Pepík Brown*. Pak 6x **Enter** a nakonec **A**.
 7. **Enter** ve **Volby** a pak 6x **Enter** a **A**.
 8. **Enter** v menu **Výpočet** spustí výpočet.
 9. V menu **Grafy** klik **Enter** otevře roletku s nabídkou řady grafů. Zůstaneme na prvním grafu **Graf regresní křivky** a pak **Enter**. Když nyní klikneme na **N**, musíme zadat rozsahy obou os, tj. začáteční a koncovou hodnotu každé osy. Lépe je přijmout nabídnuté rozsahy kliknutím na **A**. Prohlédněte si graf a ukončete **Esc**.
 10. Najedeme dolu na druhý graf **Rezidua vs. Predikce** a **Enter**. Vidíte mrak reziduí, který informuje o rozptýlení a velikostech reziduí. Klik na **mezerník** klávesnice vloží síť a znovu klik na mezerník síť zase odstraní. Tento graf ukončete **Esc**.
 11. Obdobně si prohlédněte i ostatní grafy v této roletce.
 12. Určitě musíte nyní vyšetřit vlivné body v datech a hlavně nalézt a odstranit odlehlé body, které ničí regresní analýzu. Najedťe proto dolu na **Grafy vlivných bodů** a **Enter**. V otevřené roletce pěti grafů zvolte ten nejlepší, a to je **Williamsův graf**. Body nad horní rovnoběžkou s x-ou osou jsou odlehlé. Zde není žádný. Ukončete 3x **Esc**.
 13. V menu **Výsledky** otevřete roletku klikem na **Enter**. Ťukáním na šipka dolu si prohlédnete celý výstup, zvláště nejdůležitější bloky: Blok (3) **ODHADY PARAMETRŮ A TESTY VÝZNAMNOSTI** jsou odhady úseku a směrnice přímky. Blok (4) **STATISTICKÉ CHARAKTERISTIKY REGRESE** obsahuje diagnostiky těsnosti proložení. Blok (5) **ANALÝZA KLASICKÝCH REZIDUÍ** na své konci pod tabulkou přináší statistické posouzení velikosti reziduí. Blok (6) **TESTOVÁNÍ REGRESNÍHO TRIPLETU** ukazuje na splnění předpokladů zvolené matematické metody. Blok (7) **INDIKACE VLVNÝCH BODŮ** přináší v číselné tabelární podobě vlivné body, které jsou zobrazeny v grafech vlivných bodů. Blok **Výsledky** ukončíme **Esc**.
- ADSTAT ukončíme klikem na **K** a pak **A**. Dále klik na **K** a pak **A**.

Také tento postup na dvou stranách si před analýzou dat v ADSTATu vytiskněte!

Návod k sestavení a napsání semestrální práce k této 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 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.

DC

6.1 JEDNOROZMĚRNÉ LINEÁRNÍ REGRESNÍ MODELY - - Postup analýzy dat v ADSTAT 1.25

Příklad 6.8 Omezení klasické analýzy lineárního modelu dle Anscomba-výklad R a R^2

- 1. Modul:** V červeném menu zvolte **Lineární regrese** a pak klávesou **Enter** otevřete menu **Data**.
- 2. Zadání dat:** Stiskněte klávesu **F3** a smažte *demo.dat* a napište **P608a.txt** a ukončete klávesou **Esc** a **A**.
- 3. Metoda:** V nabídce zvolte **Regresní diagnostika** a **Enter**.
- 4. Zadání:** Ve **Podmínky** nechte vše předvolené a 7x **Enter**. Ve **Volby** nechte vše předvolené a 7x **Enter**.
- 5. Výpočet:** dejte **Enter**.
- 6. Grafy:** dejte **Enter** a prohlédněte si pouze **Graf regresní křivky** a **Residual-Index plot**.
- 7. Výsledky:** dejte **Enter** a z **Výstupu** vyberte níže požadované hodnoty a vyplňte tabulku.

Celý postup nyní opakujte pro data **P608b.txt**, **P608c.txt**, **P608d.txt** k vyplnění tabulky.

Statistika	P608a	P608b	P608c	P608d
Úsek, b_0, s_0				
Směrnice b_1, s_1				
Test významnosti úseku, t_0				
Test významnosti směrnice, t_1				
Test celkové regrese, F_R				
Korelační koeficient, R				
Koeficient determinace, D				
Směrodatná odchylka, $s(y)$				
Trend v reziduích				
Závěr: model je				

Tento postup na jedné stránce si před analýzou dat v ADSTATu vytiskněte!

Návod k sestavení a napsání semestrální práce k této 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. 609 – 618 a obecný výklad outputu v lineární regresi na str. 582 - 583. V ADSTATu používejte na přenos výstupů a obrázků na papír **PrintScreen** a upravte pak negativ na černobílý pozitiv.

DC

6.2 VALIDACE – - Postup analýzy dat v ADSTAT 1.25

Příklad 6.7 Validace nové analytické metody

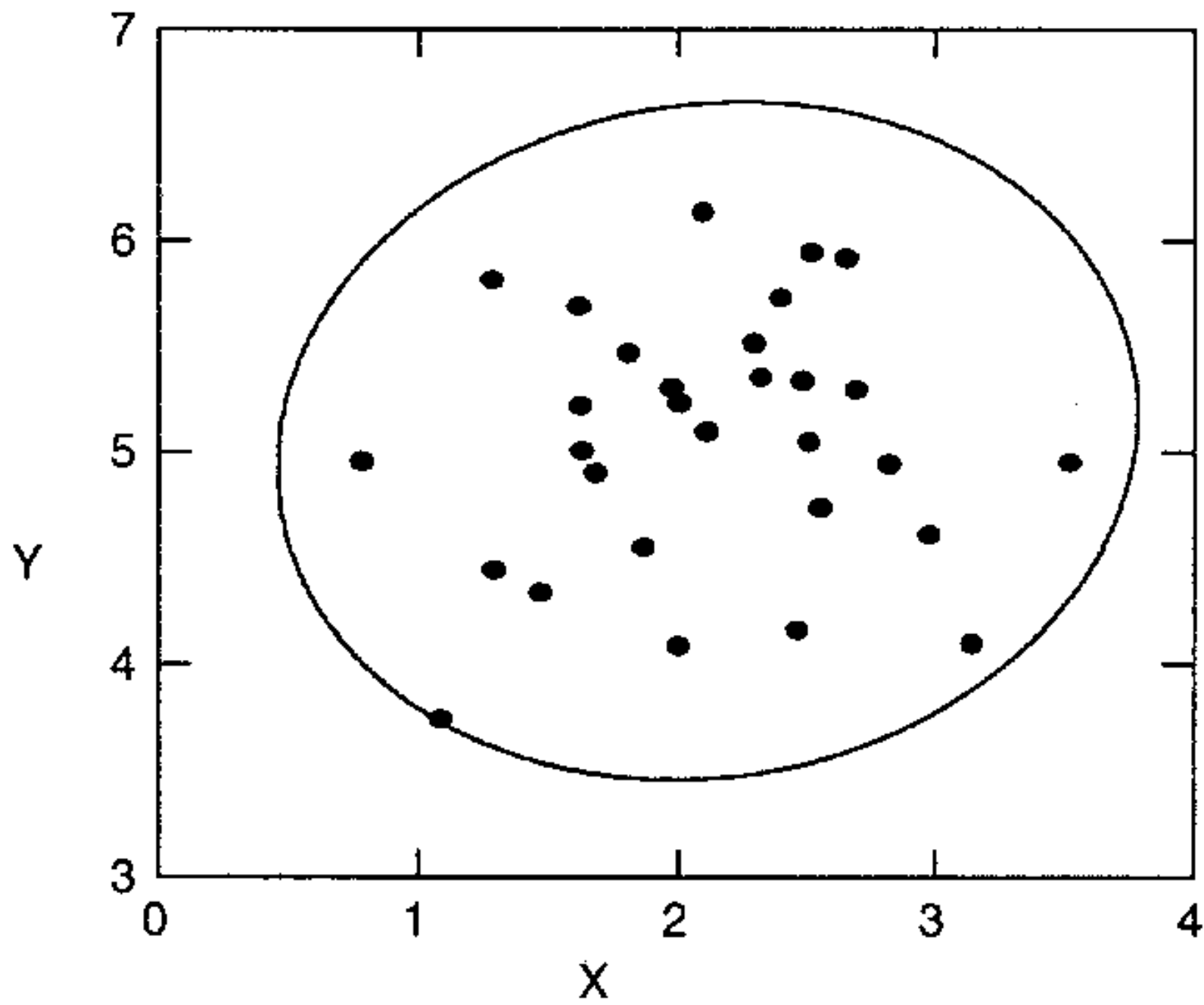
1. **Modul:** V červeném menu zvolte **Lineární regrese** a pak klávesou **Enter** otevřete menu **Data**.
2. **Zadání dat:** Nyní stiskněte klávesu **F3** a smažte *demo.dat* a napište **P607.txt** a ukončete **Esc** a **A**.
3. **Metoda:** V nabídce zvolte **Regresní diagnostika** a **Enter**.
4. **Zadání:** Ve **Podmínky** nechte vše předvolené a 7x **Enter**. Ve **Volby** nechte vše předvolené a 7x **Enter**.
5. **Výpočet:** V menu **Výpočet** dejte **Enter**.
6. **Grafy:** V menu na **Grafy** dejte **Enter** a prohlédněte si nejprve **Graf regresní křivky** a **Residual-Index plot**, u kterého stiskněte mezerník pro síť os. Všimněte si podezřelých odlehlých bodů 18, 20, 24. Rozhodnou grafy odlehlých bodů o vyloučení odlehlých bodů. Dejte **Esc** a zvolte **Grafy vlivných bodů**, které detekují 18, 20, 24 k odstranění.
7. **Výsledky:** dejte **Enter** a z **Výstupu** vyberte a zapište si k pozdějšímu porovnání: (3) **Odhady parametrů** a (4) **Statistické charakteristiky regrese** a (5) **Analýza klasických reziduí** a (6) **Testování regresního tripletu**.

Celý postup nyní opakujte pro data bez odlehlých hodnot

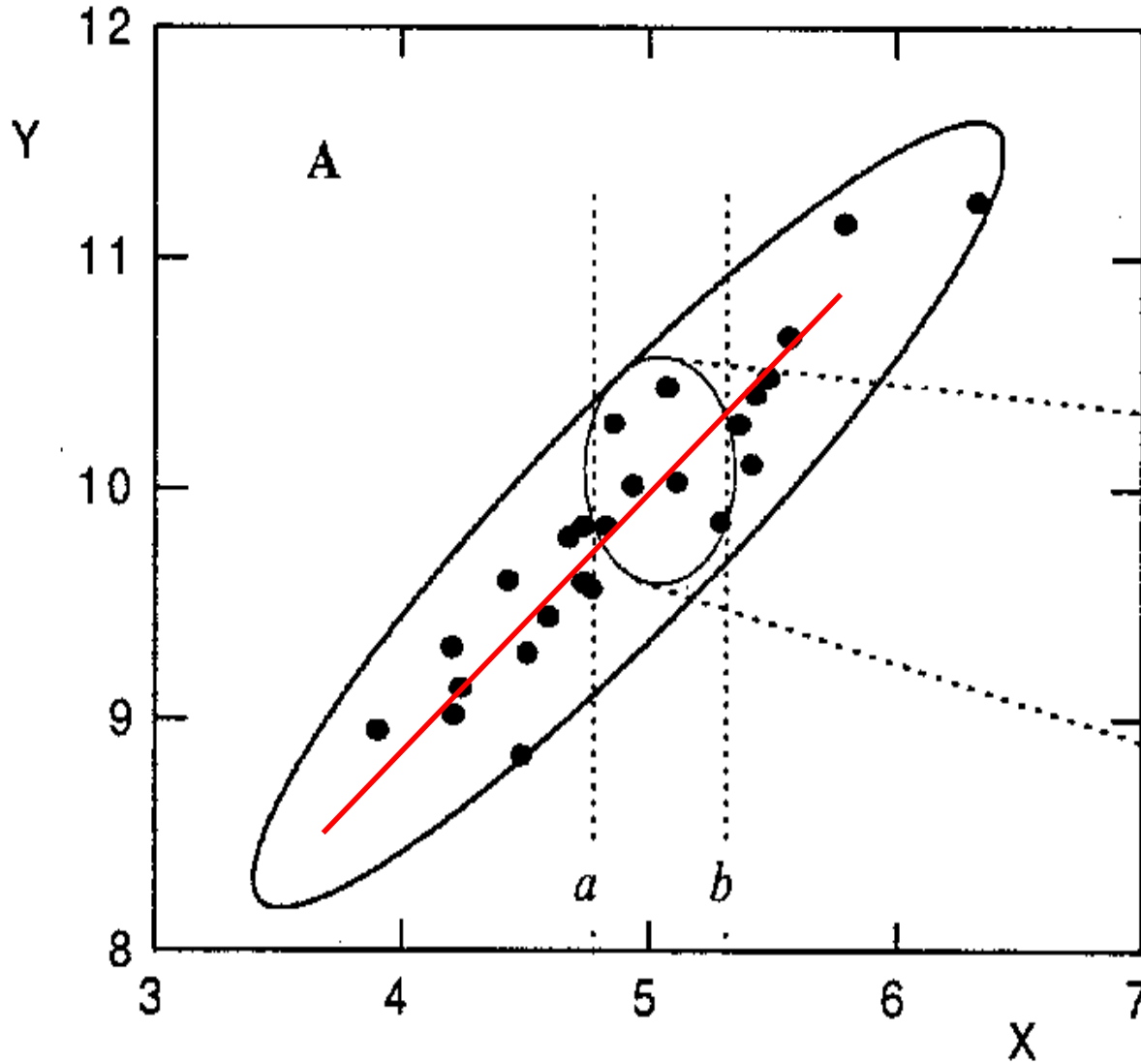
2. **Zadání dat:** Vraťte se nyní do bloku **Data**, pak **Enter** a **F4** a **F5**. Najed'te na řádek 18 a ťukejte **Delete** až po obou číslech zůstane prázdné místo. Totéž opakujte v řádce 20 a 24. Pak stiskněte **Esc** a **A**.
3. **Metoda:** V nabídce zvolte **Regresní diagnostika** a **Enter**.
4. **Zadání:** V **Podmínky** nechte vše předvolené a dejte 7x **Enter**. Ve **Volby** nechte vše předvolené a 7x **Enter**.
5. **Výpočet:** V menu **Výpočet** dejte **Enter**.
6. **Grafy:** V menu na **Grafy** dejte **Enter** a prohlédněte si nejprve **Graf regresní křivky** a **Residual-Index plot**, u kterého stiskněte mezerník síť os.
7. **Výsledky:** dejte **Enter** a z **Výstupu** vyberte a zapište si k porovnání s předešlým výstupem: (3) **Odhady parametrů** a (4) **Statistické charakteristiky regrese** a (5) **Analýza klasických reziduí** a (6) **Testování regresního tripletu**.

Podrobný rozbor analýzy dat a výstavby regresního modelu s výkladem celého výstupu programu ADSTAT je ve cvičebnici Kompendium a začíná na str. 609 a dále.

Nekorelovaná data

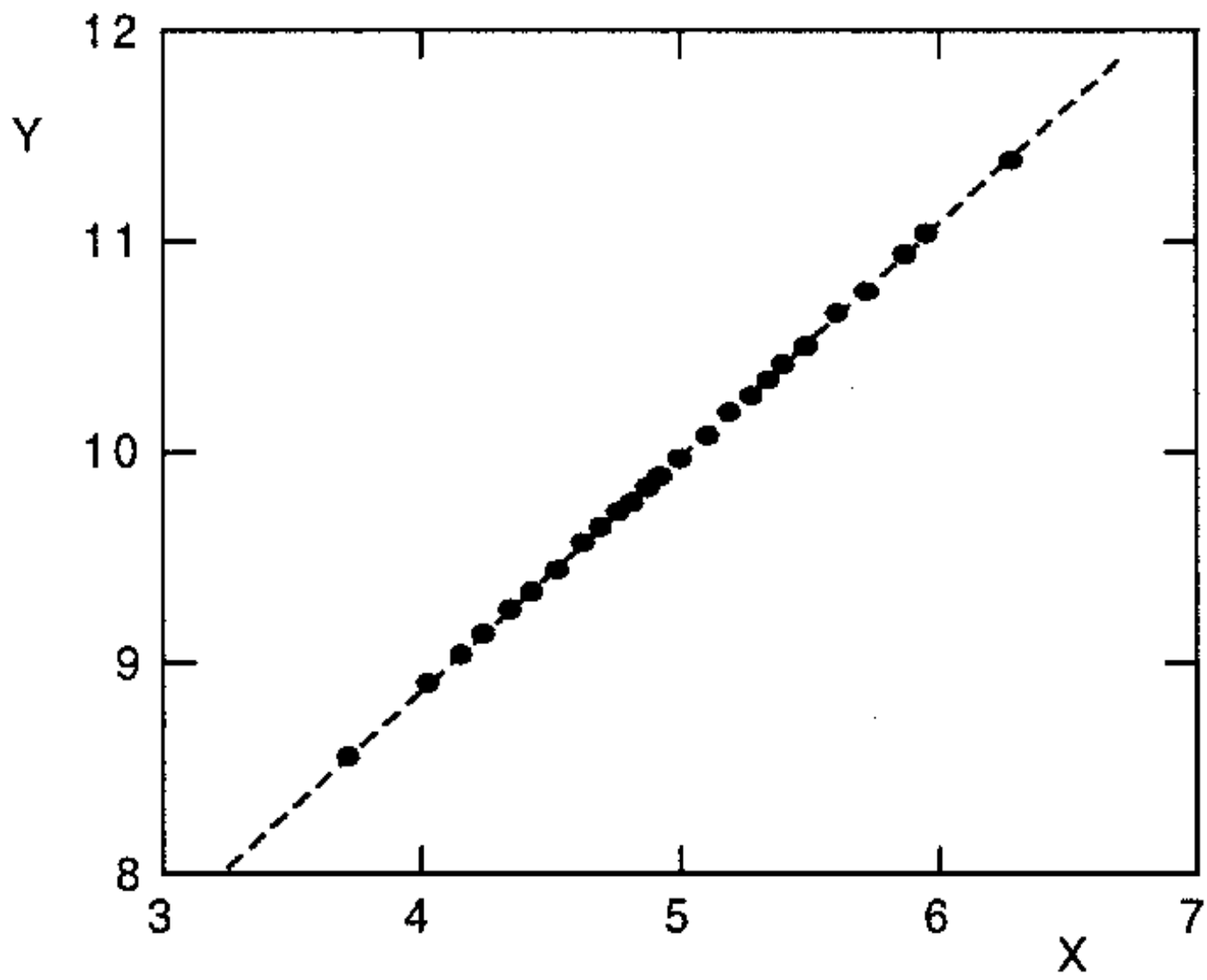


Pozitivní korelace



Pokud však budeme měřit data v příliš malém intervalu, nemusí se závislost vůbec prokázat!!

Lineární závislost



Cíl regresní analýzy

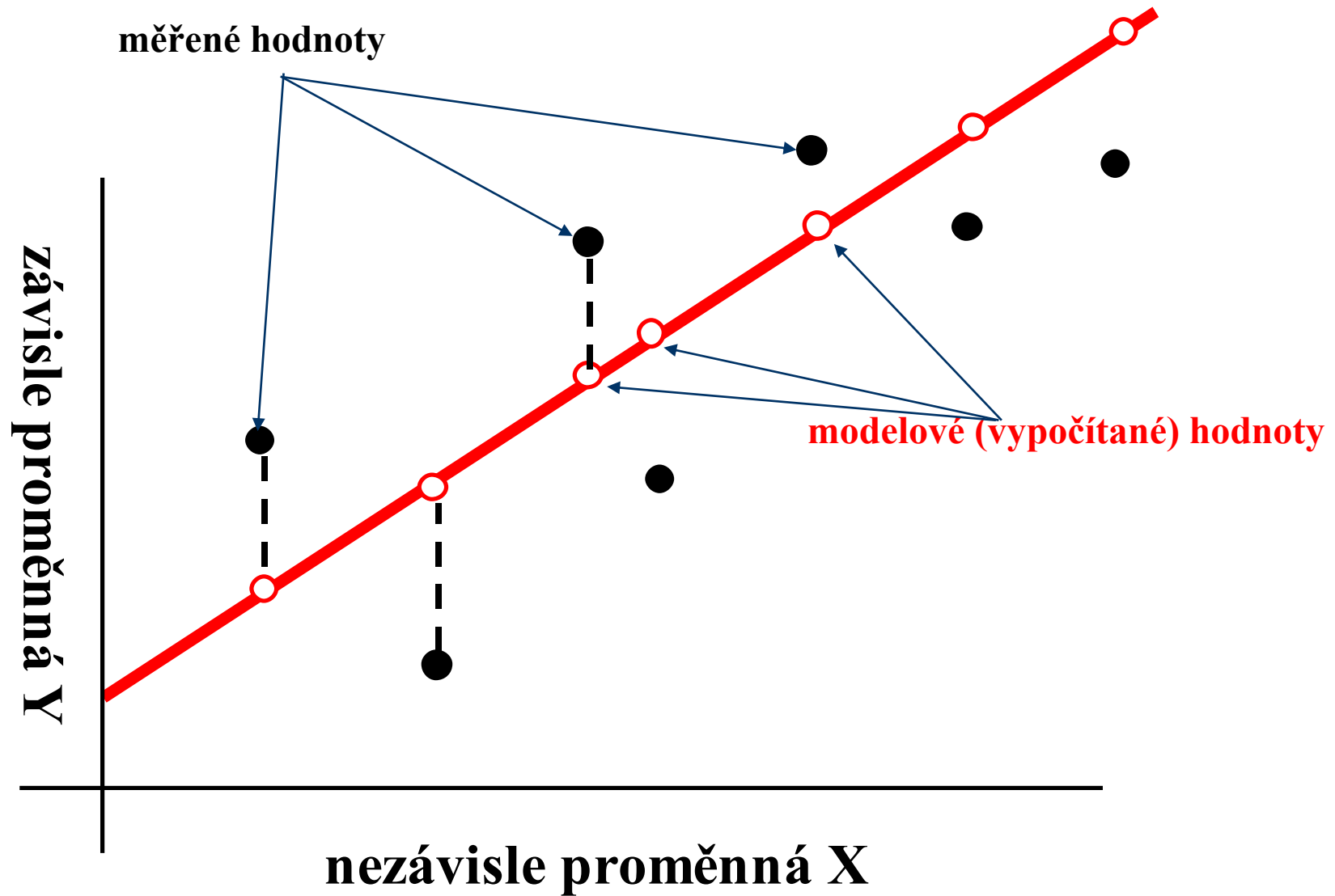
Cílem regresní analýzy je nalezení vhodného modelu studované závislosti tak, že se **snažíme nahradit**

každou měřenou (experimentální) hodnotu závisle proměnné y_{exp}

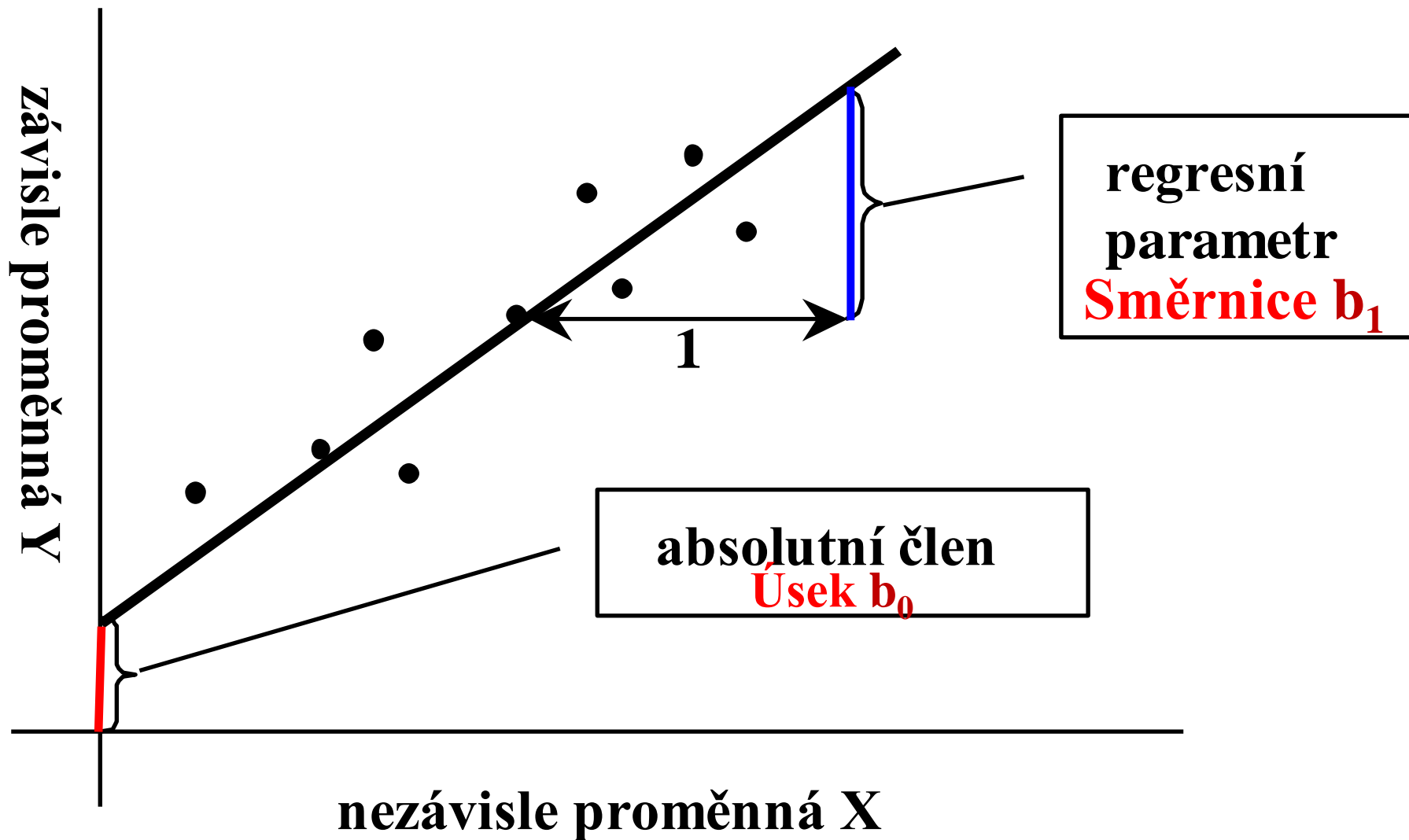
hodnotou vypočtenou (predikovanou) y_{vyp}

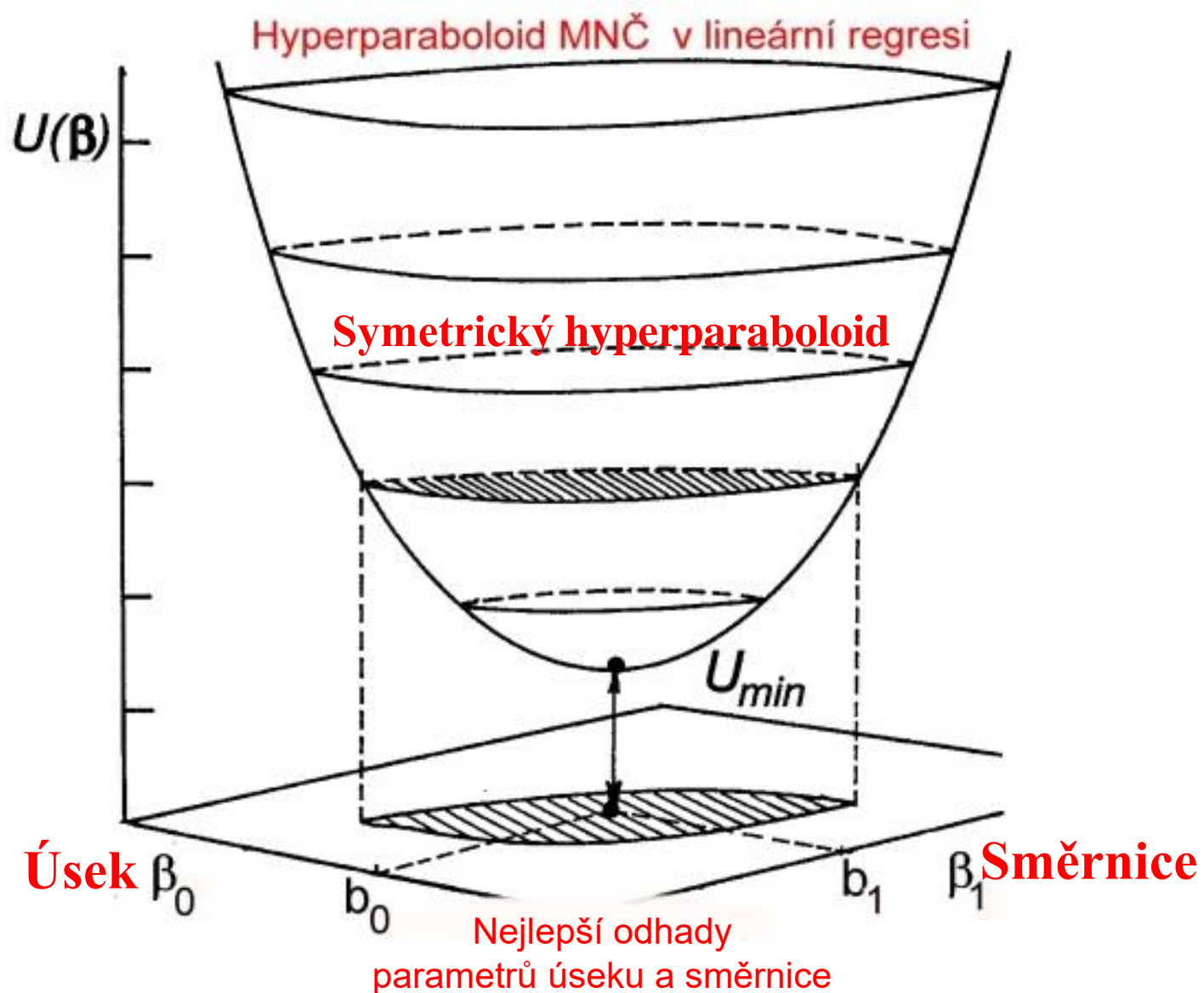
čili **hodnotou** ležící na spojitě funkci (modelu) **nezávisle proměnné x** .

Grafické vysvětlení cíle regresní analýzy



Grafické vysvětlení jednoduchého lineárního regresního modelu $y = b_0 + b_1x$





Účelová funkce U dosáhne minima pro nejlepší odhady parametrů úseku a směrnice

Popis závislostí čili regresních modelů

Příklad: Český hydrometeorologický ústav v Praze měřil na stanicích s různou nadmořskou výškou průměrnou roční teplotu půdy. Údaje jsou uvedeny v následující tabulce. Existuje mezi oběma proměnnými nějaká závislost?

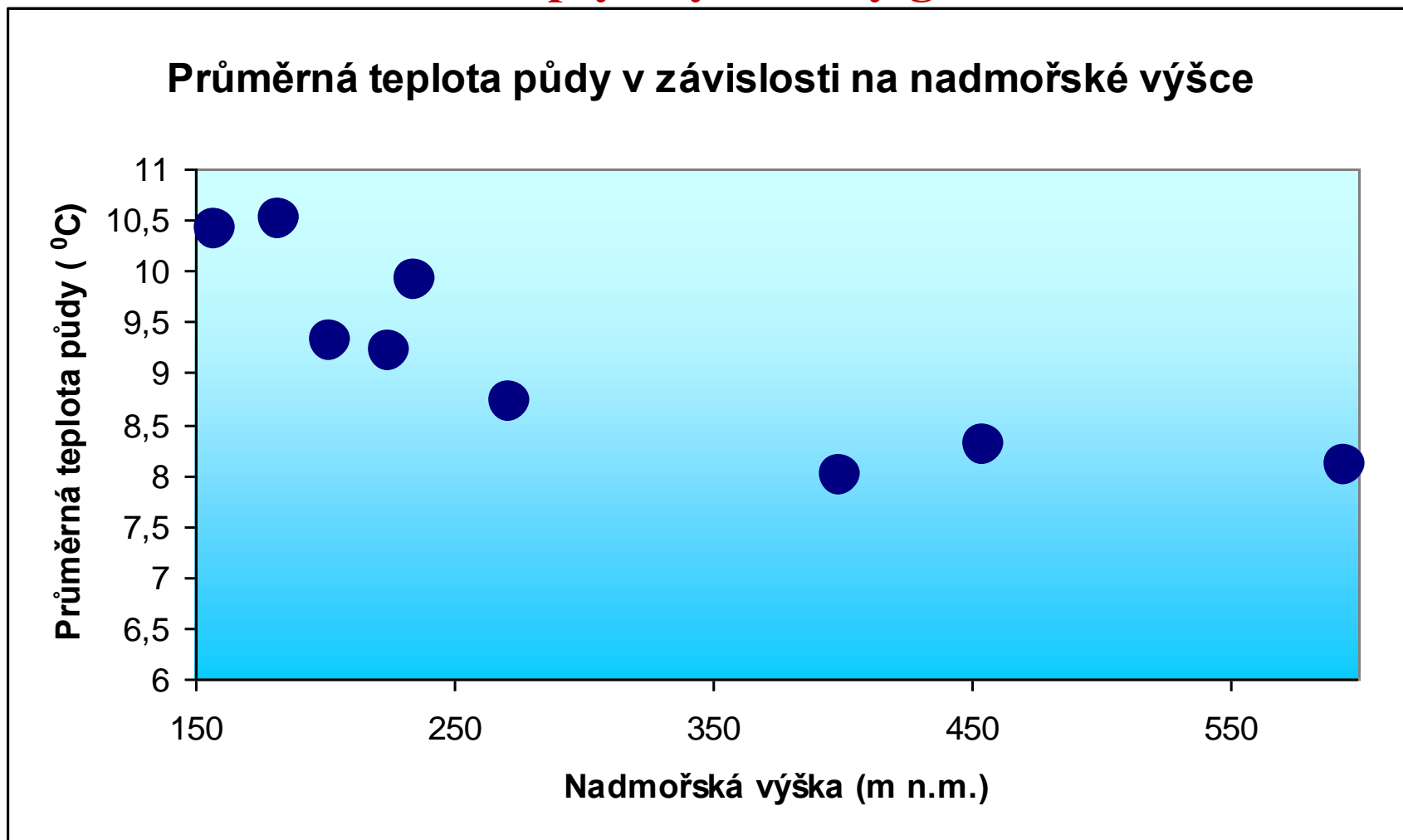
Data: Jednoduché metody k popisu závislosti vystihuje korelační koeficient

$r_{yx} = -0,835$ a také rozptylový bodový graf.

Nadmořská výška v m n.m.	158	183	203	225	235	272	400	455	595
Průměrná teplota půdy °C	10,4	10,5	9,3	9,2	9,9	8,7	8	8,3	8,1

Výklad: Jedná se o silnou závislost, ne však deterministickou, protože kromě výšky zde působí na teplotu i jiné faktory. S rostoucí výškou průměrná teplota půdy klesá.

Rozptylový bodový graf



Regresní model

Závislost vystihuje rovnice anebo graf určité křivky.

Např. pro nejjednodušší závislost je to graf přímky

$$y = \eta + \varepsilon = \beta_0 + \beta_1 x + \varepsilon$$

Deterministická složka obsahuje
vliv vysvětlující proměnné

Náhodná složka obsahuje
všechny ostatní proměnné
(nepopsané) vlivy

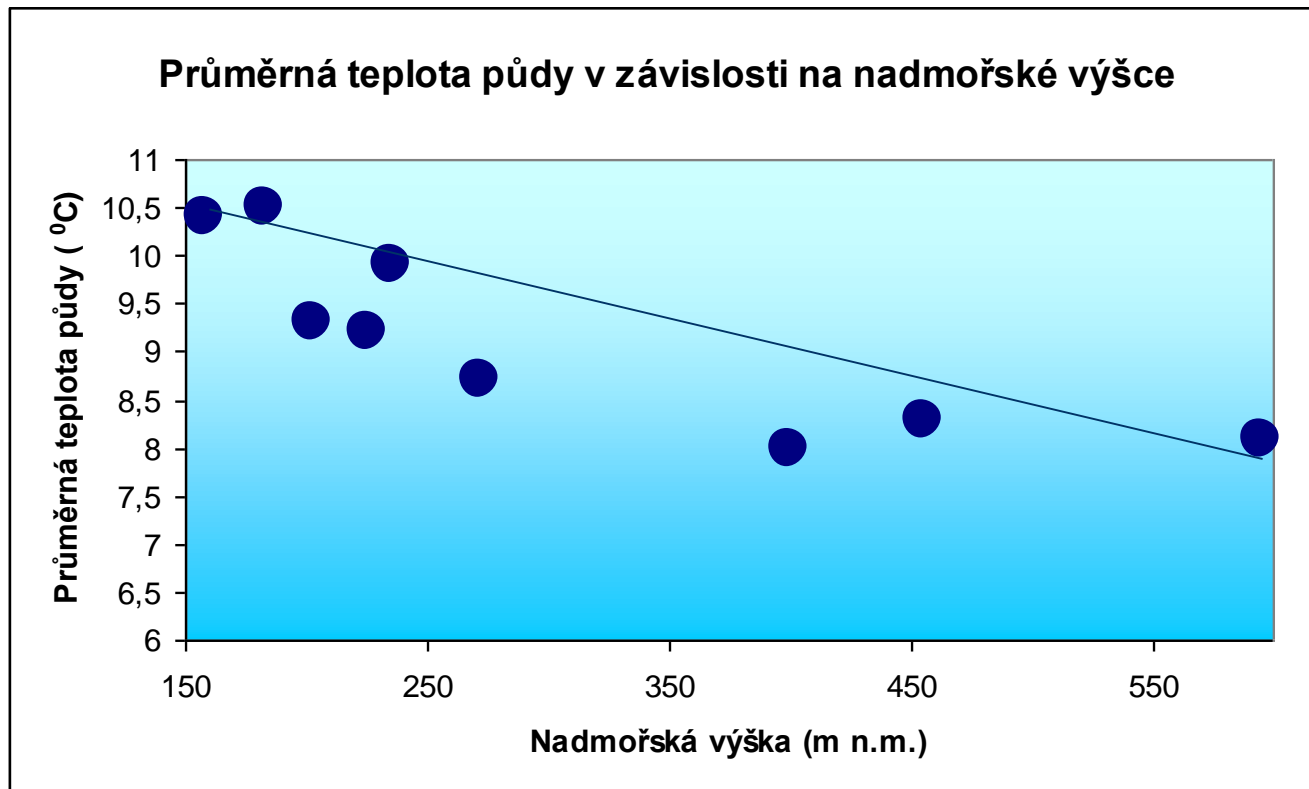
Deterministická složka η : Popisuje závislost mezi pozorovanými, naměřenými proměnnými.
Je vyjádřena konkrétní matematickou funkcí.

Náhodná složka ε : Popisuje závislost vysvětlované proměnné na neznámých nebo nepozorovaných proměnných a popisuje i vliv náhody. Vyjadřuje se pravděpodobnostní funkcí.

Pokud body proložíme přímkou, hovoříme o tzv. *regresní přímce*.

Pokud by všechny body ležely na přímce, šlo by o model pouze s deterministickou složkou η .

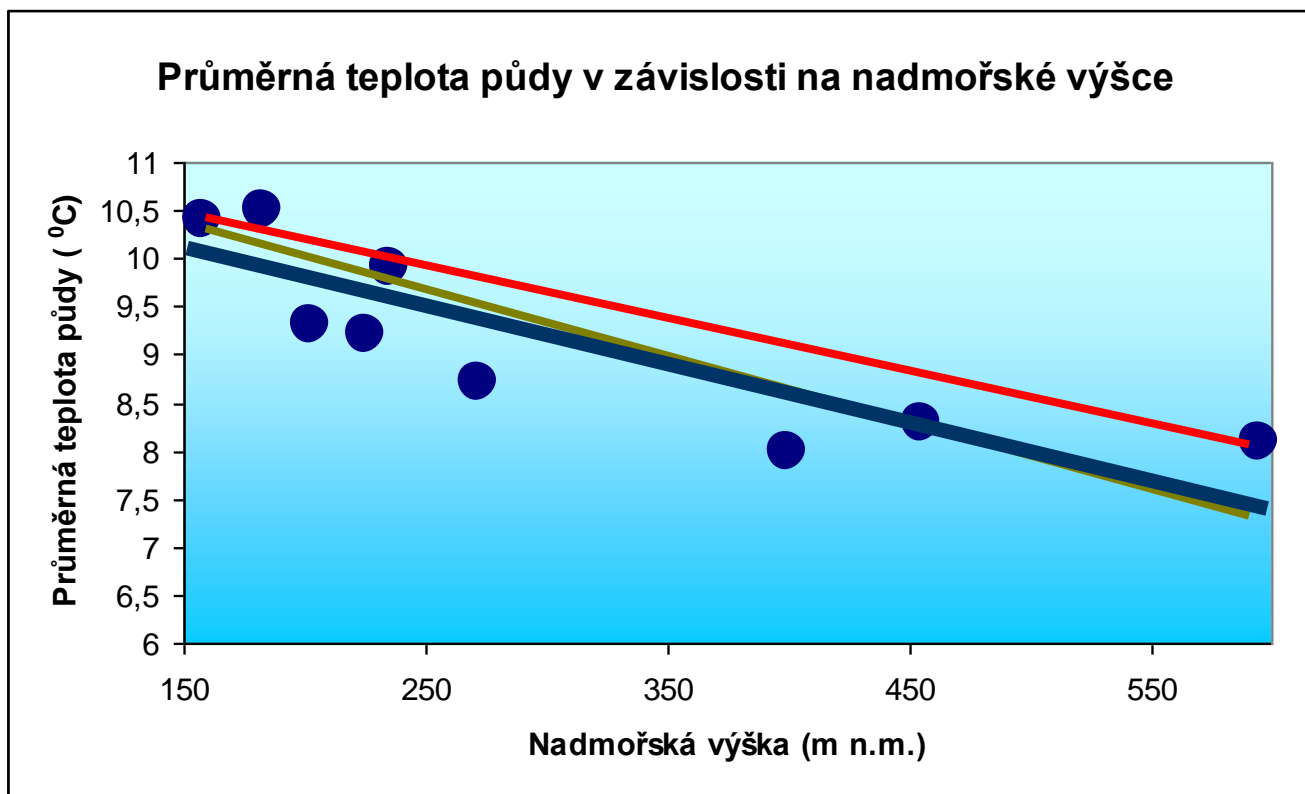
Body však leží i mimo – v modelu je deterministická složka η i náhodná složka ε .



Lze vložit přímku jen tak od oka - **zelená**.

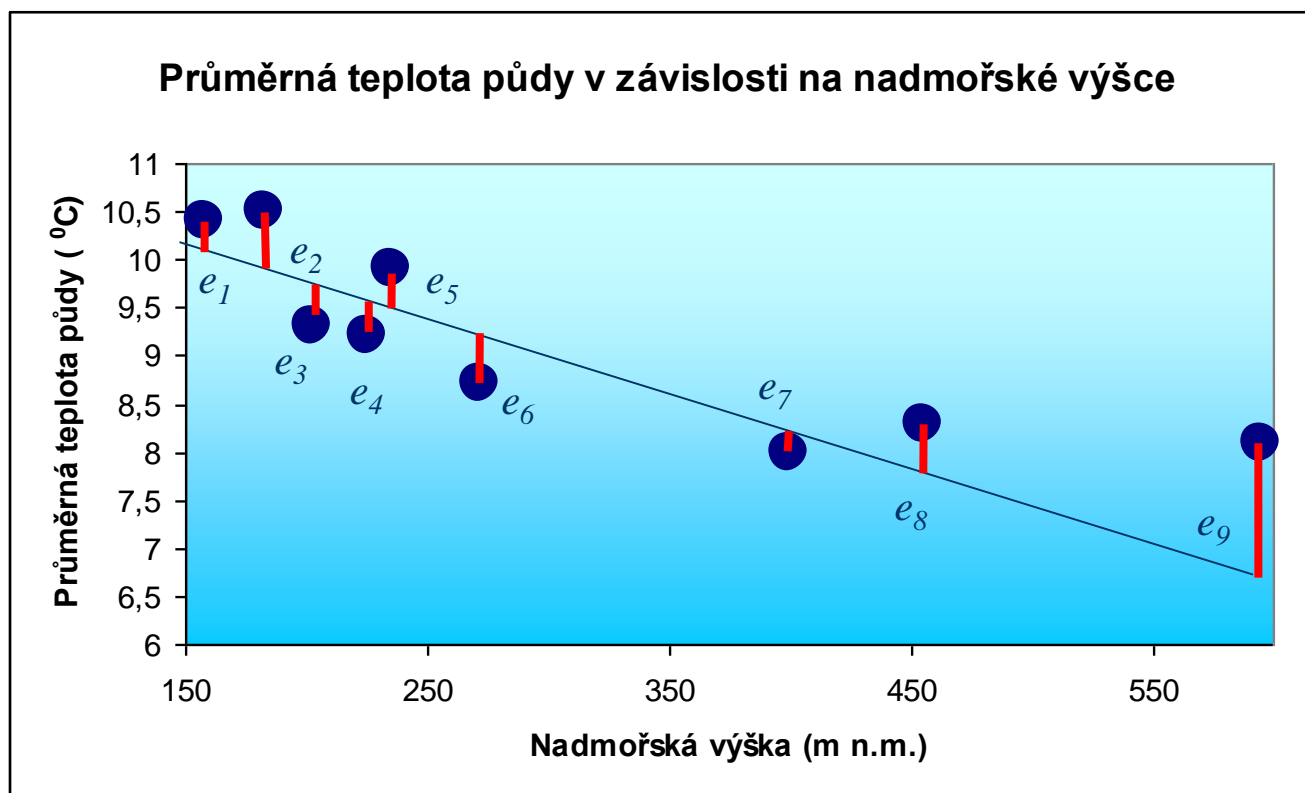
Lze spojit krajní body – **červená**.

Lze použít nástroje regresní analýzy a nalézt přímku, která prochází **nejblíže** všem bodům – **černá**)

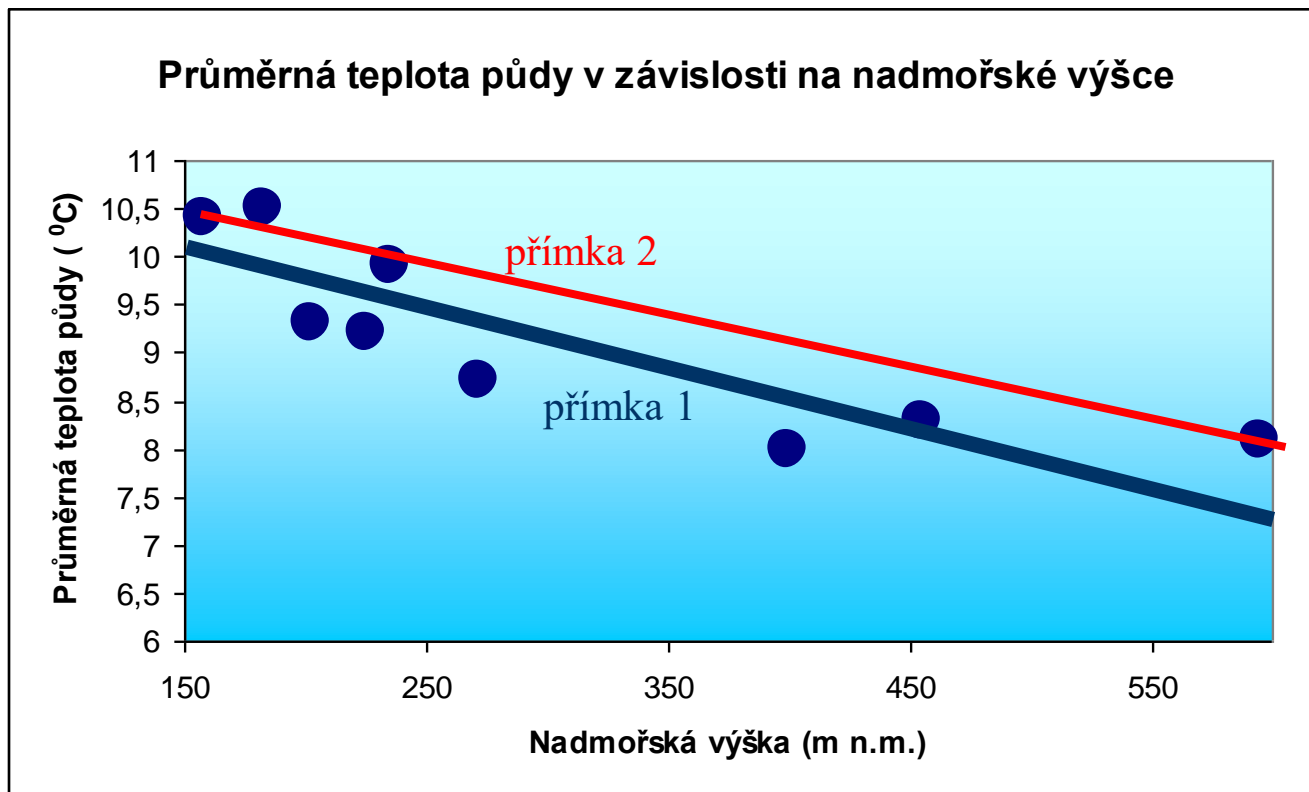


Přímka procházející nejbliže všem bodům je vždy *jen jedna!*
K jejímu nalezení slouží *metoda nejmenších čtverců* (MNC).

Vybere ze všech možných přímek takovou, pro kterou je součet druhých mocnin (čtverců) odchylek bodů od přímky e_i^2 minimální.



Přímka označená jako 1 je blíže k bodům, součet čtverců odchylek je menší než u přímky označené jako 2. Přímka 1 je proto vhodnější.



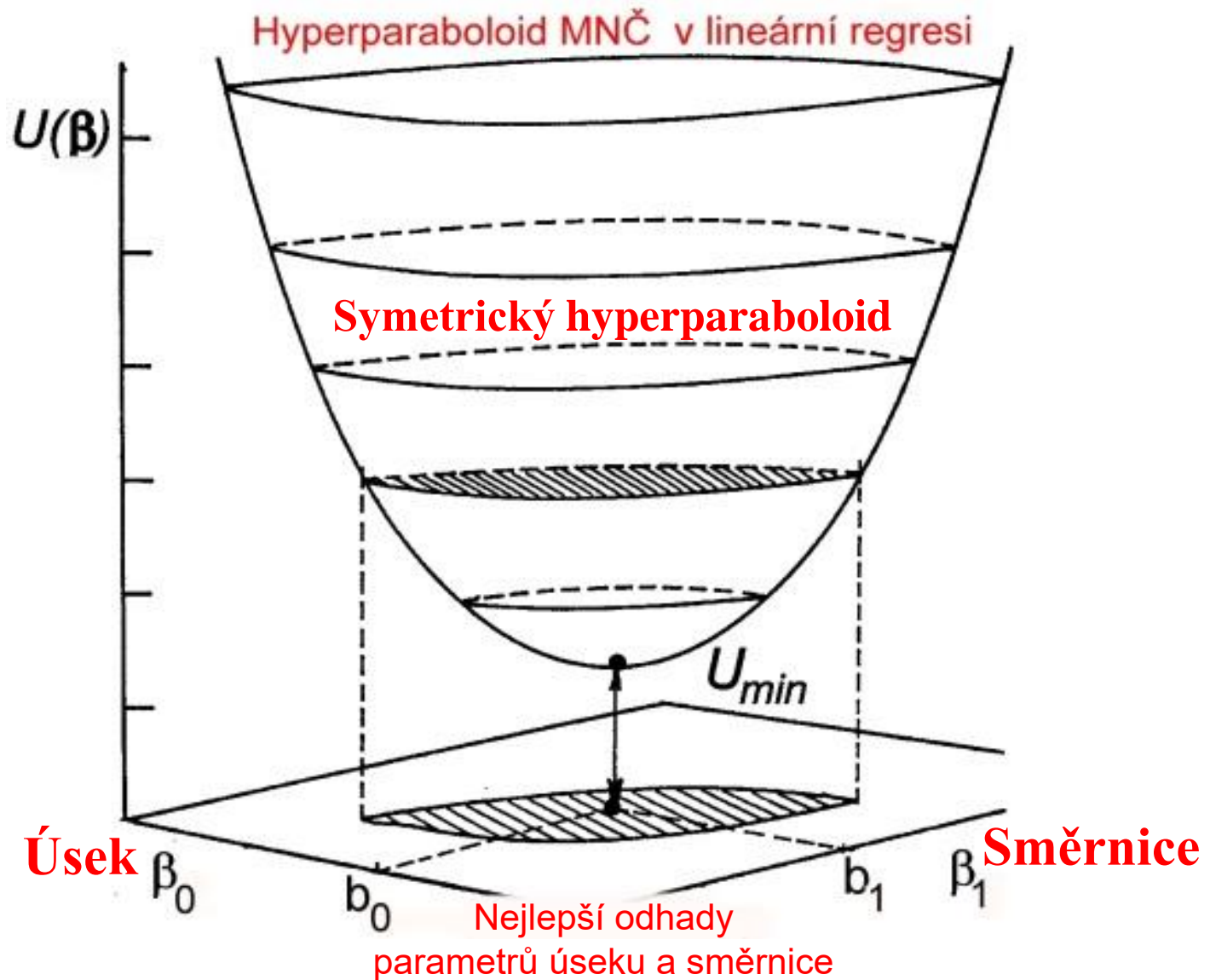
Metoda nejmenších čtverců

Nástroj k určení bodových odhadů parametrů *výběrové regresní přímky*: $\hat{y} = b_0 + b_1x$

Výběrová je protože je založena pouze na výběrových datech.

Parametry b_0 , b_1 jsou *výběrové* (empirické) *regresní parametry*.

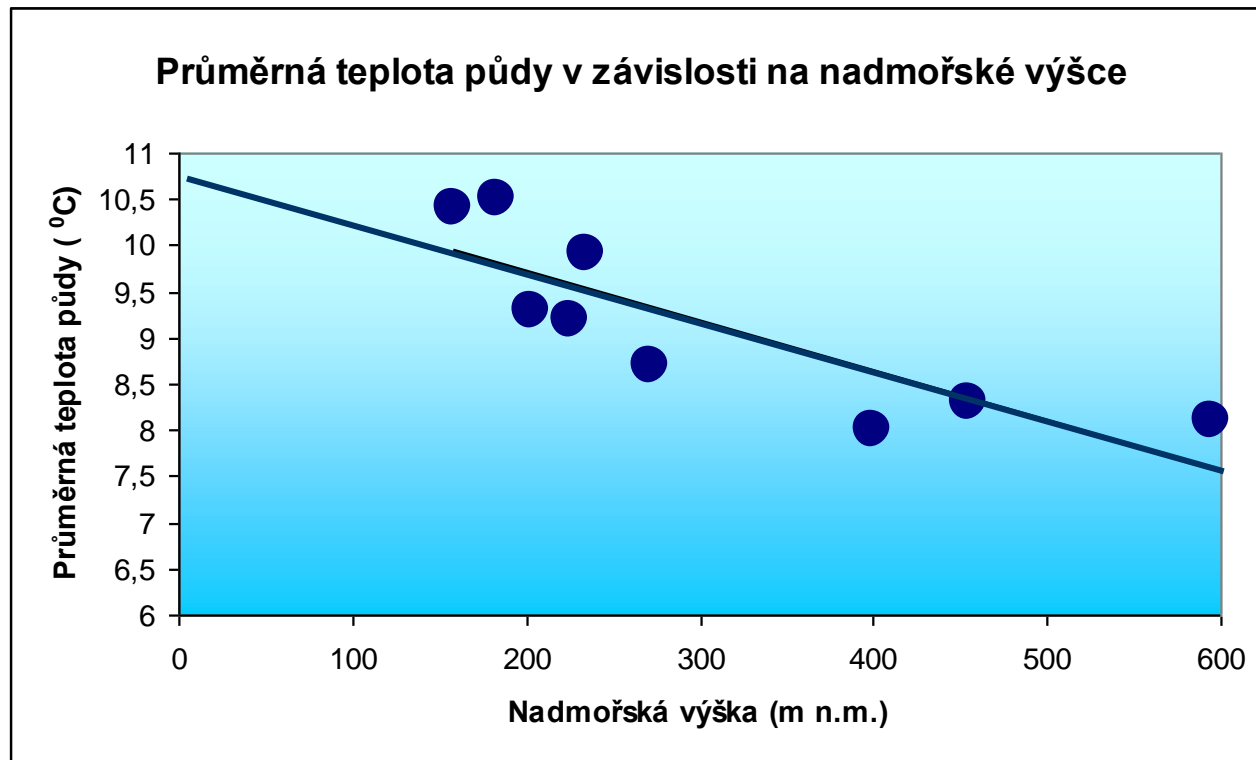
Oproti tomu *regresní přímka*: $\eta = \beta_0 + \beta_1x$, je založena na datech základního souboru, která ale nejsou k dispozici.



Účelová funkce U dosáhne minima pro nejlepší odhady parametrů úseku a směrnice

Pomocí metody nejmenších čtverců byla odhadnuta regresní přímka ve tvaru $\hat{y} = 10,795 - 0,00541 \cdot x$, kterou lze též zapsat:

průměrná teplota půdy = 10,795 – 0,00541 × nadmořská výška

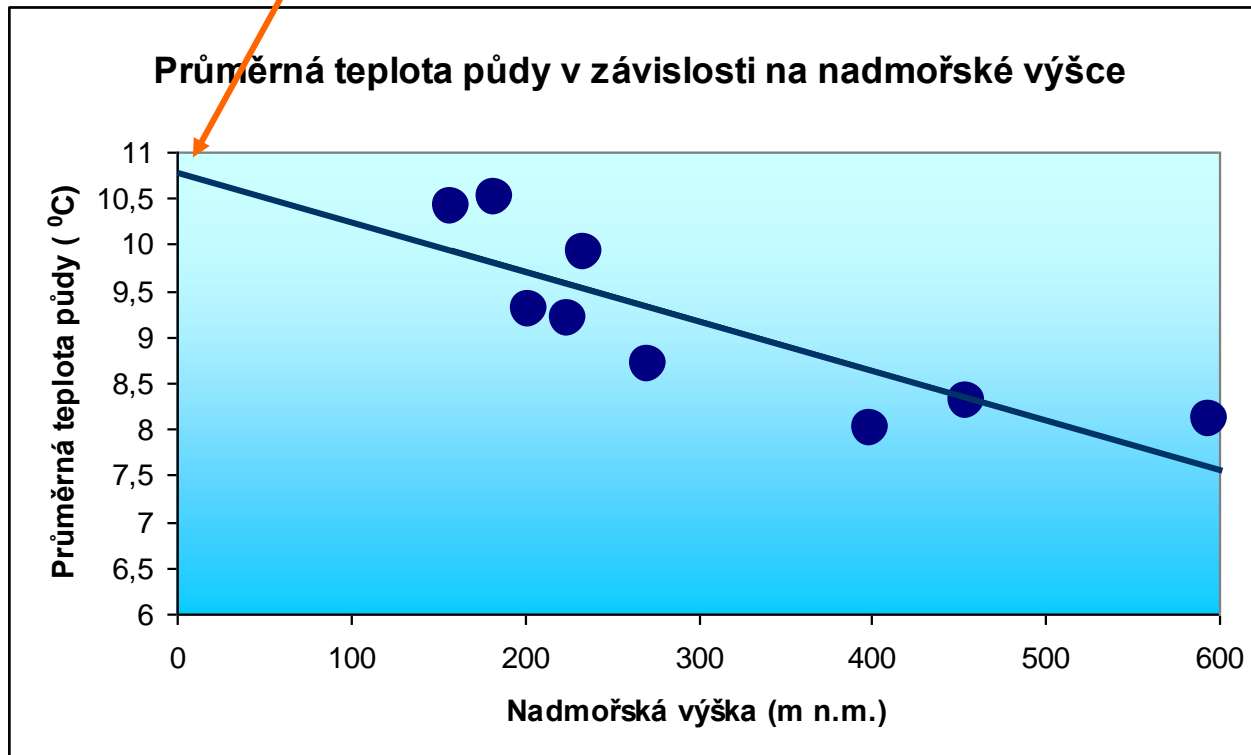


Pomocí metody nejmenších čtverců byla odhadnuta regresní přímka
ve tvaru $\hat{y} = 10,795 - 0,00541x$.

Vysvětlení fyzikálního významu parametru b_0 : $b_0 = 10,795$ je průsečík

Parametr
přímky s osou Y.

V nadmořské výšce 0
metrů n.m. by podle
modelu byla průměrná
teplota půdy 10,795 °C.

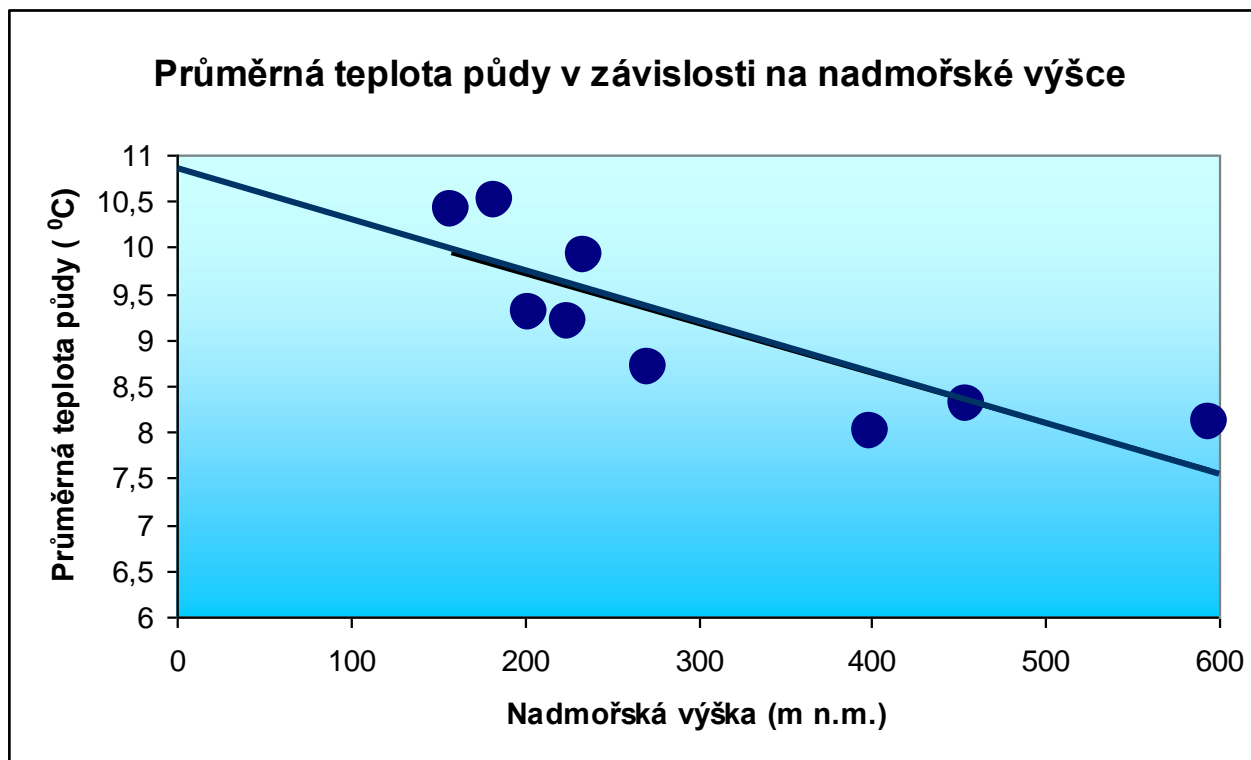


Pomocí metody nejmenších čtverců byla odhadnuta regresní přímka ve tvaru $\hat{y} = 10,795 - 0,00541x$.

Vysvětlení fyzikálního významu parametru b_1 : Parametr $b_1 = 0,00541$ je směrnicí přímky a udává její sklon.

Je záporný, protože přímka klesá.

S každým dalším metrem nadmořské výšky klesá průměrná teplota půdy v průměru o $0,00541^\circ\text{C}$.



Bodové a intervalové odhady neznámých parametrů úseku b_0 a směrnice b_1 v regresním modelu

$$y = \eta + \varepsilon = b_0 + b_1x + \varepsilon$$

	<i>Parametry</i>	<i>Směrodatná odchylka</i>	<i>t stat</i>	<i>Hodnota P</i>	<i>Dolní mez 95%</i>	<i>Horní mez 95%</i>
Hranice	10,79504	0,446866	24,15722	5,3E-08	9,738368	11,85171
Nadmořská výška	-0,00541	0,00134	-4,03872	0,004941	-0,00858	-0,00224

Regresní přímka $\hat{y} = 10,795 - 0,00541x$

Obecný lineární regresní model

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

y závisle proměnná

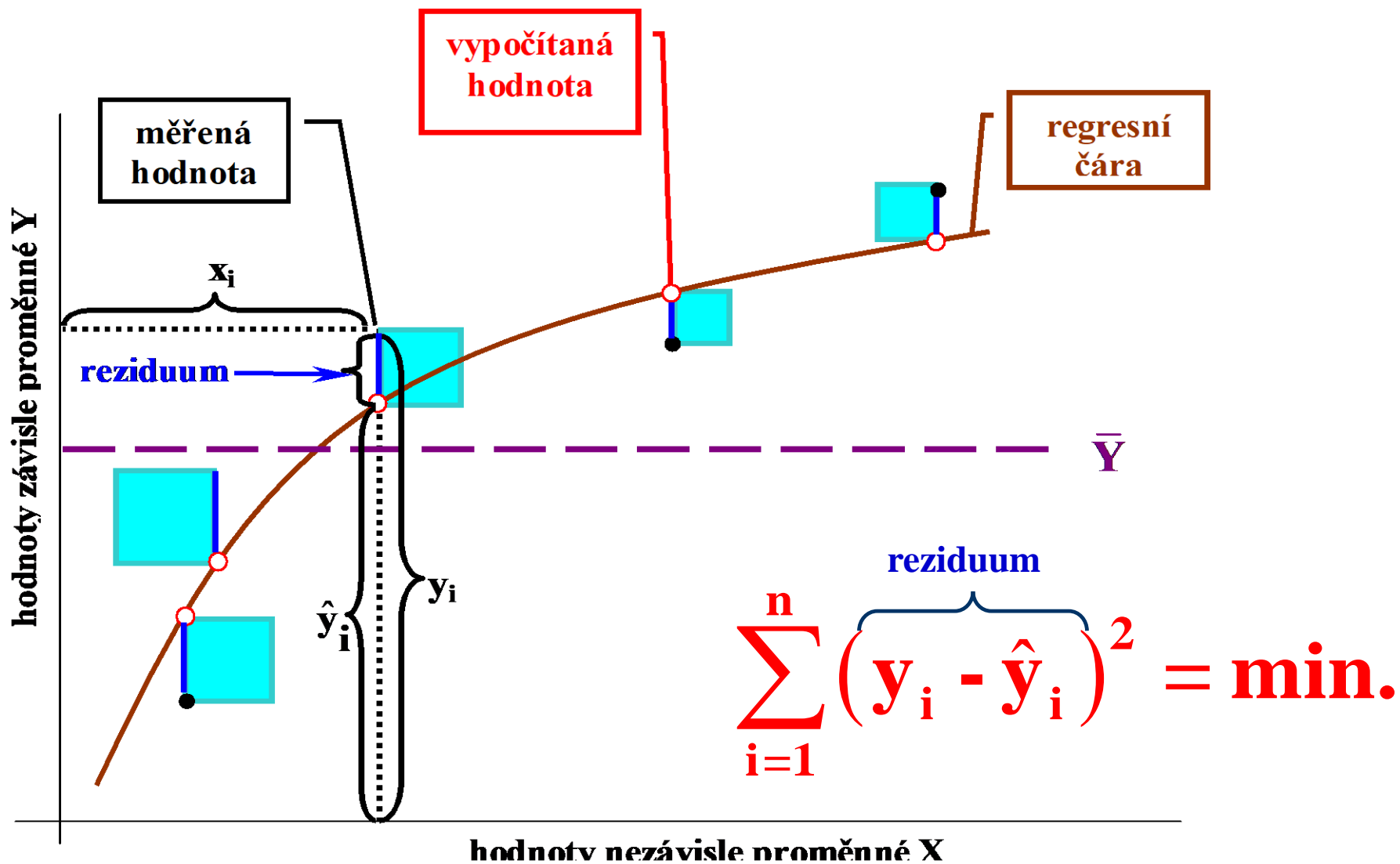
X nezávisle proměnná

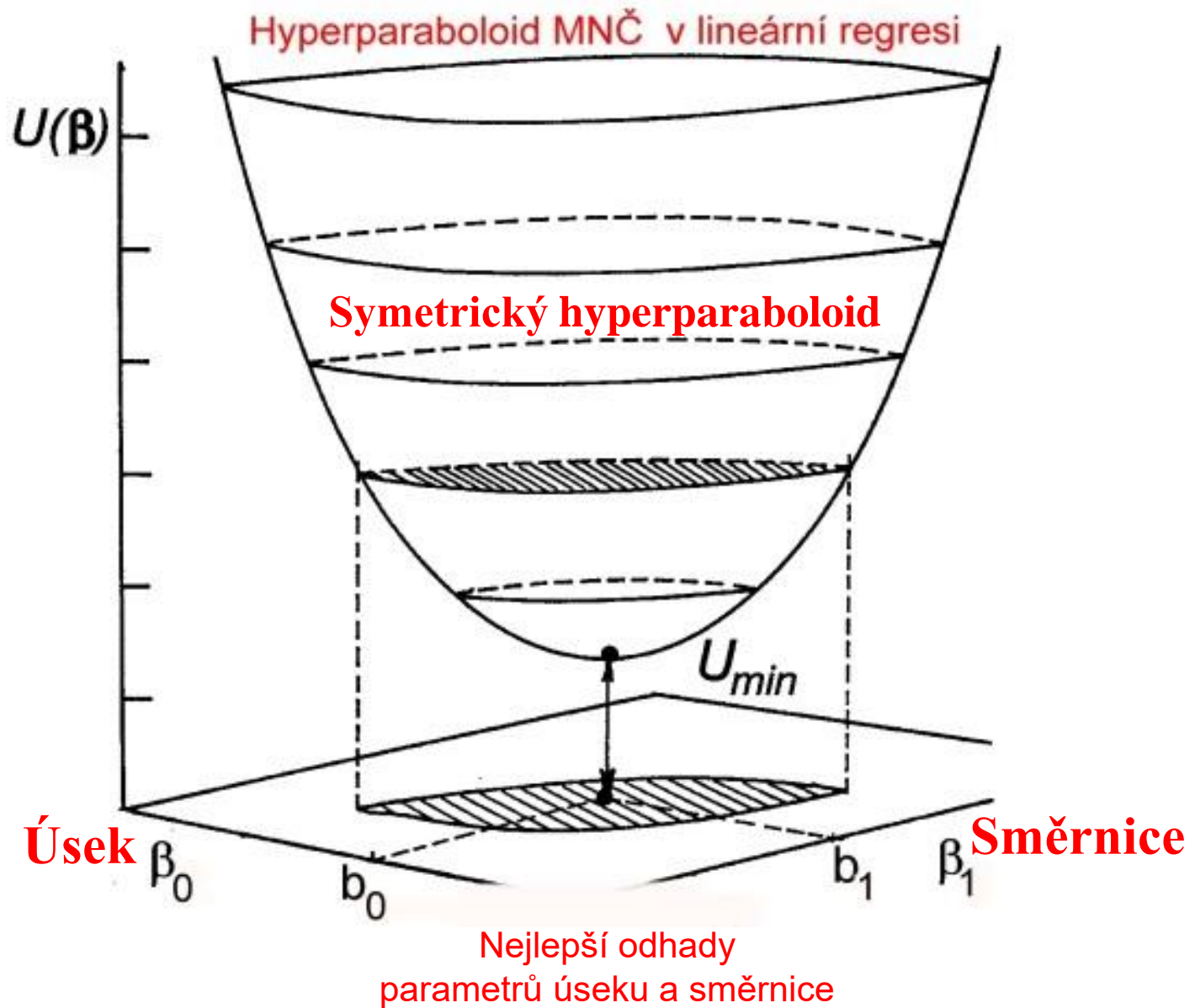
β regresní parametry

ε náhodná chyba

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

Vyčíslení odhadů parametrů lineárního regresního modelu metodou nejmenších čtverců (MNČ)





Účelová funkce U dosáhne minima pro nejlepší odhady parametrů úseku a směrnice

Předpoklady metody nejmenších čtverců

1. Parametry β mohou nabývat libovolných hodnot. Omezení jsou pouze fyzikálního smyslu.
2. Model je lineární v parametrech β ; platí přitom additivní model měření $y = \mathbf{X} \beta + \epsilon$.
3. Matice \mathbf{X} je nenáhodná, nastavitelných hodnot nezávisle proměnných. Má hodnost m :
 - (a) Žádné dva sloupce x_j a x_k nejsou kolineární (čili paralení).
 - (b) $\mathbf{X}^T \mathbf{X}$ je pak symetrická regulární matice.
 - (c) Rovina L je m rozměrná a vektory $\mathbf{X} \mathbf{b}$ jsou jednoznačné.

4. Náhodné chyby ϵ_i mají nulovou střední hodnotu $E(\epsilon_i) = 0$.

Je-li $E(\epsilon_i) = K$, je nutno zavést absolutní člen a pak bude
 $E(\epsilon_i) = 0$.

5. Náhodné chyby ϵ_i mají konstantní rozptyl, homoskedasticita,

$$E(\epsilon_i^2) = \sigma^2.$$

6. Náhodné chyby ϵ_i jsou vzájemně nekorelované, $\text{cov}(\epsilon_i, \epsilon_j) =$

$$E(\epsilon_i \cdot \epsilon_j) = 0.$$

7. Náhodné chyby mají normální rozdělení $\epsilon \approx N(0, \sigma^2)$.

Typy regresního modelu

Příklady **lineárních** regresních modelů:

$$y = \mathbf{a} + \mathbf{b}x$$

- přímka

$$y = \mathbf{a} + \mathbf{b}x + \mathbf{c}x^2$$

- parabola

$$y = \mathbf{a} + (\mathbf{b}/x)$$

- hyperbola

Lineární modely mohou být i modely, jejichž grafickým vyjádřením je křivka!!

Příklady **nelineárních** regresních modelů:

$$y = \mathbf{a} \cdot x^{\mathbf{b}}$$

$$y = \mathbf{a} \cdot e^{\mathbf{b}x}$$

$$y = \mathbf{a} \cdot e^{\frac{\mathbf{b}}{x}}$$

Výhody nelineární modelů: jsou schopny modelovat složité reálné děje, např. růst, včetně reálné predikce.

Nevýhody nelineárních modelů: daleko složitější výpočet

Obecný postup regresní analýzy

1. Navrhnout **vhodný tvar regresního modelu** čili postavit příslušnou rovnici či vzorec, který bude popisovat závislost y na x .
2. Určit **parametry modelu β** vyčíslením jejich konkrétních odhadů b .
3. Určit **statistickou významnost modelu**, tj. testovat zda navržený model významným způsobem přispěje ke zpřesnění odhadu závisle proměnné oproti použití pouhého průměru všech hodnot y .
4. Predikované hodnoty regresním modelem **vysvětlit** z hlediska zadání.

Koeficient determinace R^2 ($\times 100\%$):

je čtverec vícenásobného korelačního koeficientu. Vyjadřuje procento bodů, popsaných regresním modelem.

Testování: $H_0: \beta_c = 0$ nebo $H_0: R^2 = 0$

Testační kritérium F-testu

$$F_R = \frac{(\text{CSC} - \text{RSC})(n - m)}{\text{RSC}(m - 1)} = \frac{\hat{R}^2(n - m)}{(1 - \hat{R}^2)(m - 1)}$$

Je-li $F_R < F_{1-\alpha}(m - 1, n - m)$, je H_0 přijata.

Významnost korelačního koeficientu je shodná s testem významnosti všech regresních koeficientů vyjma absolutního členu.

JEDNOROZMĚRNÉ LINEÁRNÍ REGRESNÍ MODELY

Jak vysvětlit výstup programu a napsat semestrální práci je
v *Kompendiu*, str. 583

Návod k sestavení 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 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.

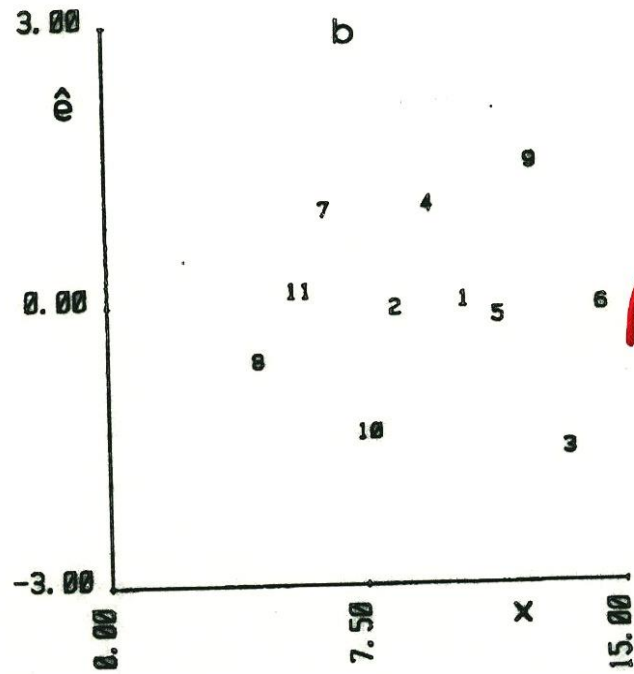
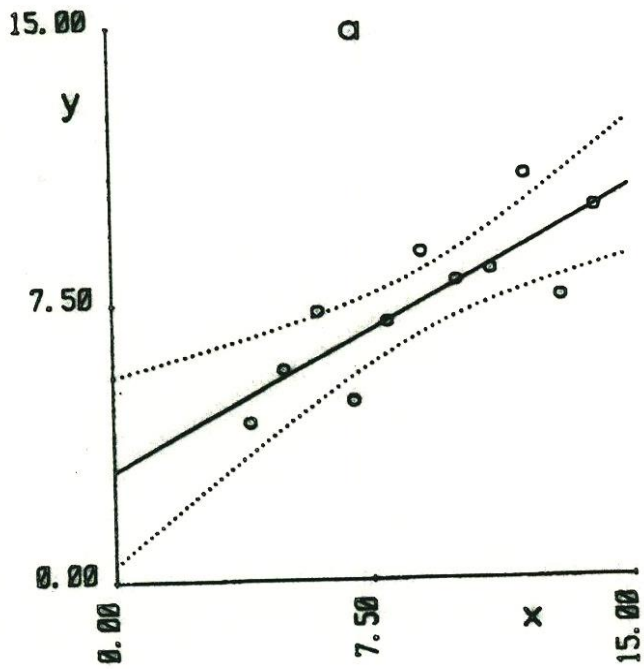
Příklad 6.8 *Omezení klasické analýzy lineárního modelu*

Anscomb⁵ uvádí testovací data pro čtyři simulované výběry. Testujte statistickou významnost obou parametrů β_1 a β_2 a proveďte grafickou analýzu reziduí.

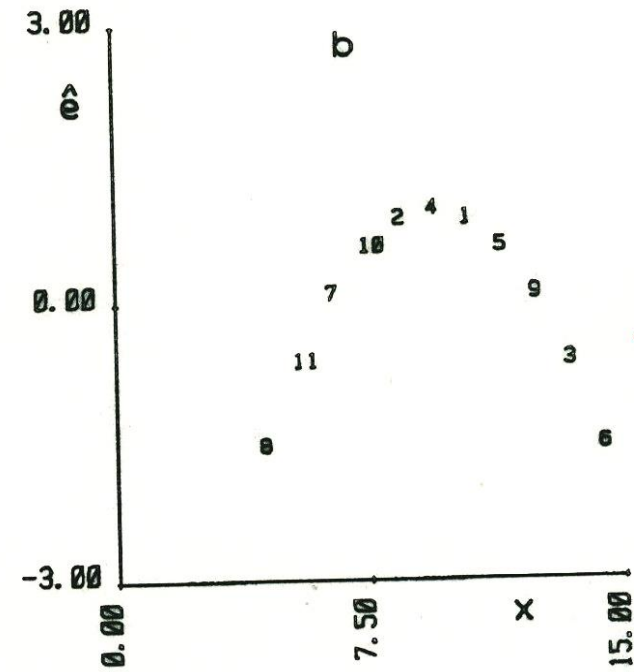
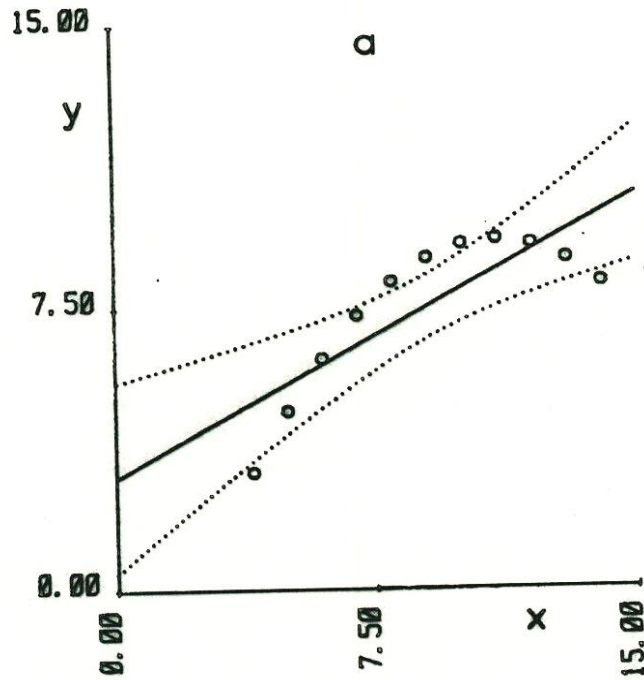
Data: čtyři simulované výběry vykazují stejné charakteristiky

$$b_1 = 0.5, b_2 = 3.0, D(b_1) = 0.0139 \text{ a } D(b_2) = 1.2656.$$

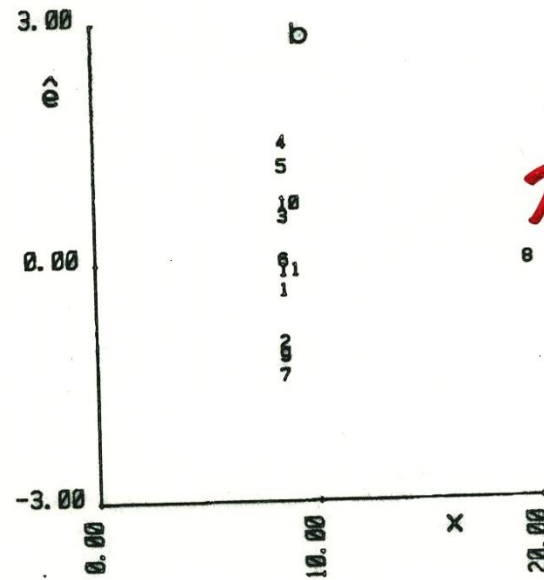
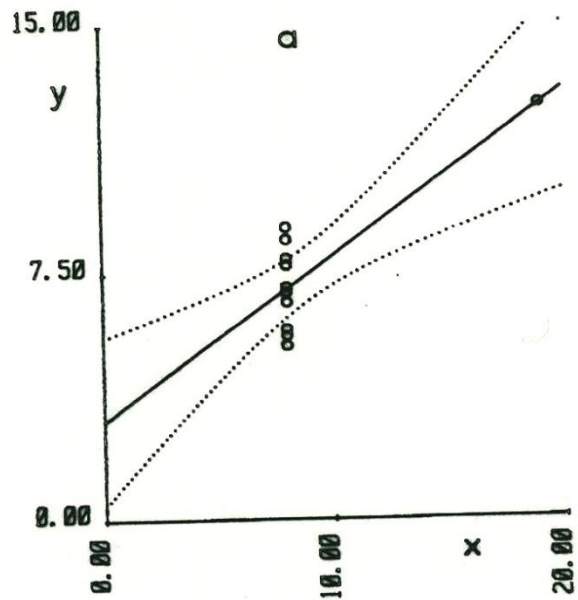
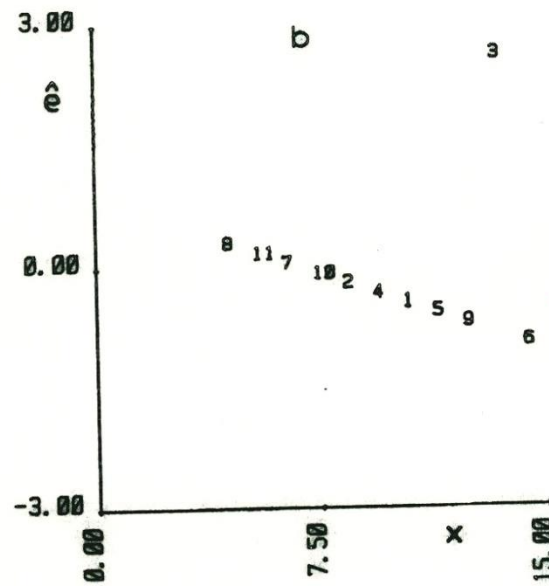
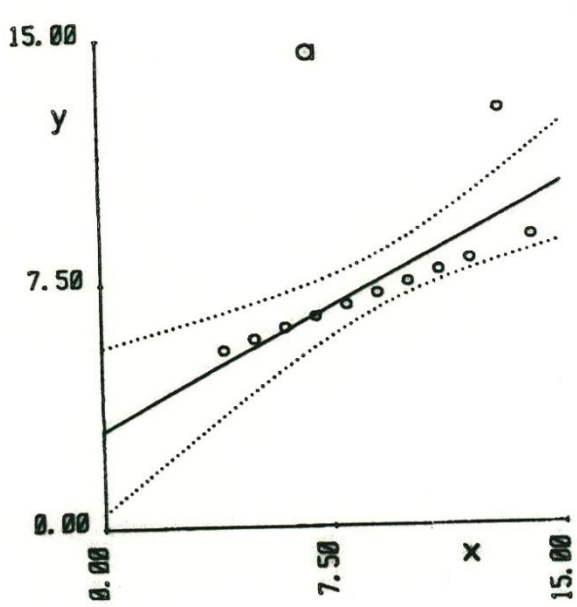
Výběr		A	B	C	D	
Proměnná	x	y	y	y	x	y
1	10	8.04	9.14	7.46	8	6.58
2	8	6.95	8.14	6.77	8	5.76
3	13	7.58	8.74	12.74	8	7.71
4	9	8.81	8.77	7.11	8	8.84
5	11	8.33	9.26	7.81	8	8.47
6	14	9.96	8.10	8.84	8	7.04
7	6	7.24	6.13	6.08	8	5.25
8	4	4.26	3.10	5.39	19	12.50
9	12	10.84	9.13	8.15	8	5.56
10	7	4.82	7.26	6.42	8	7.91
11	5	5.68	4.74	5.73	8	6.89



weak



trend



Závěr: Neshodu modelu s daty indikuje grafická analýza reziduí

Tento postup na dvou stranách si před analýzou dat v ADSTATu vytiskněte!

Návod k sestavení a napsání semestrální práce k této úloze najdete ve cvičebnici *M. Meloun, J. Miltký: Kompendium statistického zpracování dat, Karolinum Praha 1992, ISBN 978-80-246-2196-8*, ve vzorovém příkladu na str. 583 – 593 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.

DC

6.1 JEDNO-ROZMĚRNÉ LINEÁRNÍ REGRESNÍ MODELY - - Postup analýzy dat v ADSTAT 1.25

Příklad 6.8 Omezení klasické analýzy lineárního modelu dle Anscomba-výklad R a R^2

- 1. Modul:** V červeném menu zvolte **Lineární regrese** a pak klávesou **Enter** otevřete menu **Data**.
- 2. Zadání dat:** Stiskněte klávesu **F3** a smažte *demo.dat* a napište **P608a.txt** a ukončete klávesou **Esc** a **A**.
- 3. Metoda:** V nabídce zvolte **Regresní diagnostika** a **Enter**.
- 4. Zadání:** Ve **Podmínky** nechte vše předvolené a 7x **Enter**. Ve **Volby** nechte vše předvolené a 7x **Enter**.
- 5. Výpočet:** dejte **Enter**.
- 6. Grafy:** dejte **Enter** a prohlédněte si pouze **Graf regresní křivky** a **Residual-Index plot**.
- 7. Výsledky:** dejte **Enter** a z **Výstupu** vyberte níže požadované hodnoty a vyplňte tabulku.

Celý postup nyní opakujte pro data **P608b.txt**, **P608c.txt**, **P608d.txt** k vyplnění tabulky.

Statistika	P608a	P608b	P608c	P608d
Úsek, b_0, s_0				
Směrnice b_1, s_1				
Test významnosti úseku, t_0				
Test významnosti směrnice, t_1				
Test celkové regrese, F_R				
Korelační koeficient, R				
Koeficient determinace, D				
Směrodatná odchylka, $s(y)$				
Trend v reziduích				
Závěr: model je				

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

A D S T A T - Hlavní menu statistického systému
(c) TriloByte 1990

Jednorozměrná data
Víceřozměrná data
Kalibrace
Lineární regrese
Nelineární regrese
Uyhlazování
Pravděpodobnostní modely
Růstové křivky
Chemometrie
Analýza rozptylu
Konec práce

Pro Lineární regresi dejte **Enter**.

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

F1=nápověda F2=uložení F3=čtení F4=numerický mód F5=zoom ESC=konec
Počet řádků: 11 Počet sloupců: 2 Uodorovně P608a

Numerický editor

	1	2
1	1.000000 E+01	8.040000
2	8.000000	
3	1	
4	9	
5	1.100000 E+01	8.330000
6	1.400000 E+01	9.960000
7	6.000000	7.240000
8	4.000000	4.260000
9	1.200000 E+01	1.084000 E+01
10	7.000000	4.820000
11	5.000000	5.680000

Čtení ze souboru: P608a

značení: S=sloupec R=řádek N=nová poloha MEZERA=změna směru
blokové operace: F6=kopírování F7=zaplnění F8=transformace
ostatní operace: F9=změna velikosti matice F10=transpozice

Otevřít **Data**, pak stisk **F3**, pak smazat *demo.dat*, napsat **P608a** a ukončit **Esc a A**.

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 1 Sloupec 5 Insert NUM P608a

Řádek	1	Sloupec	5
P608ax	P608ay		
10.00000	8.04000		
8.00000	6.95000		
13.00000	7.58000		
9.00000	8.81000		
11.00000	8.33000		
14.00000	9.96000		
6.00000	7.24000		
4.00000	4.26000		
12.00000	10.84000		
7.00000	4.82000		
5.00000	5.68000		

Otevřít **Data** a pak **F5** a ukončit **Esc a A**.

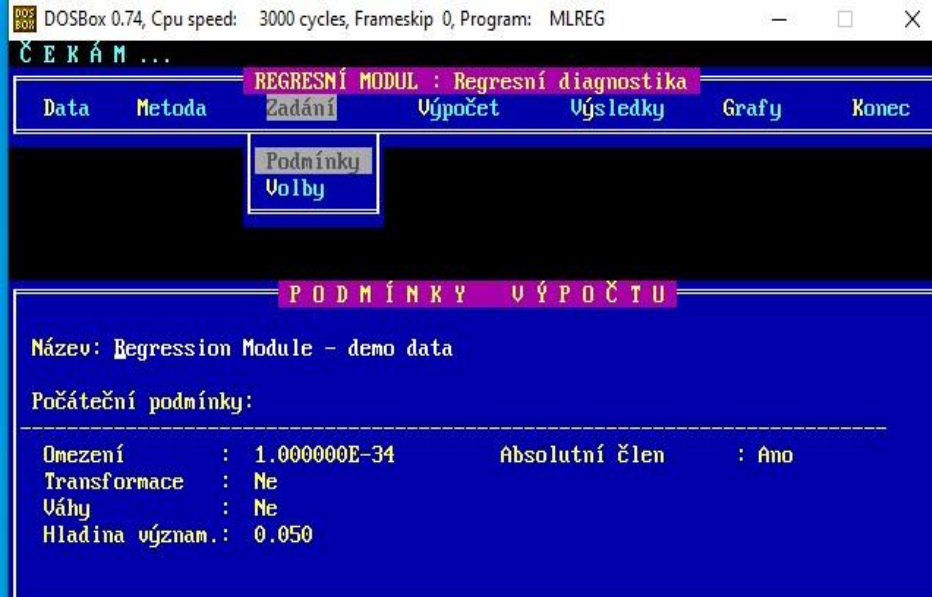
DOSBox 0.74, Cpu speed: 3000 cycles, Frameskip 0, Program: MLREG

REGRESNÍ MODUL : Lineární nejmenší čtverce

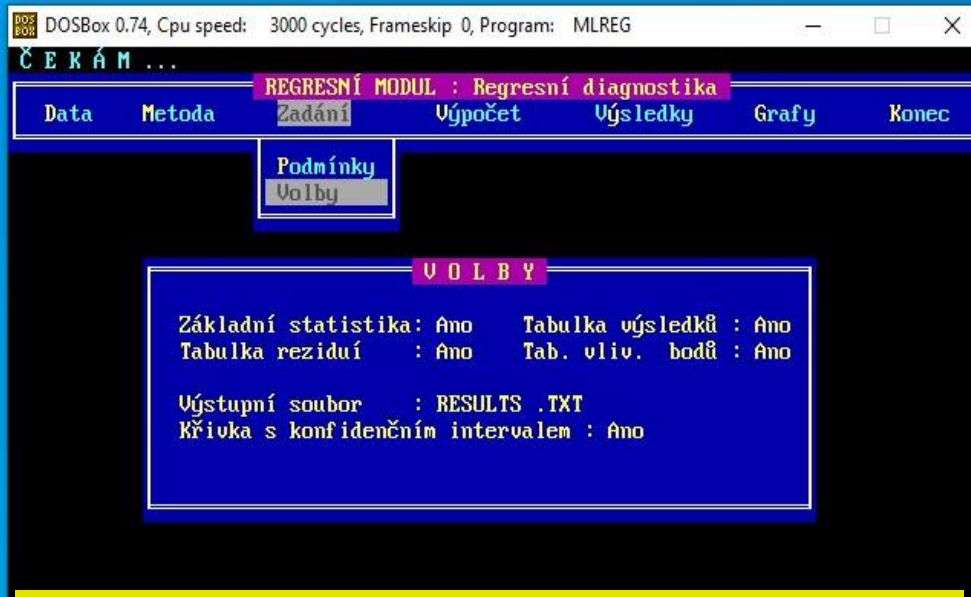
Data	Metoda	Zadání	Účpočet	Výsledky	Grafy	Konec
	Lineární regrese MNC					
	Regresní diagnostika					

Blok **Metoda** otevřít klávesou **Enter** a pak zvolit **Regresní diagnostika** a **Enter**.

Lineární regresní diagnostika



V Podmínky nechte vše předvolené defaultní a dejte 7x Enter.



Ve Volby nechte vše defaultní a dejte 7x Enter a pak ve Účpočet klikněte opět na Enter.

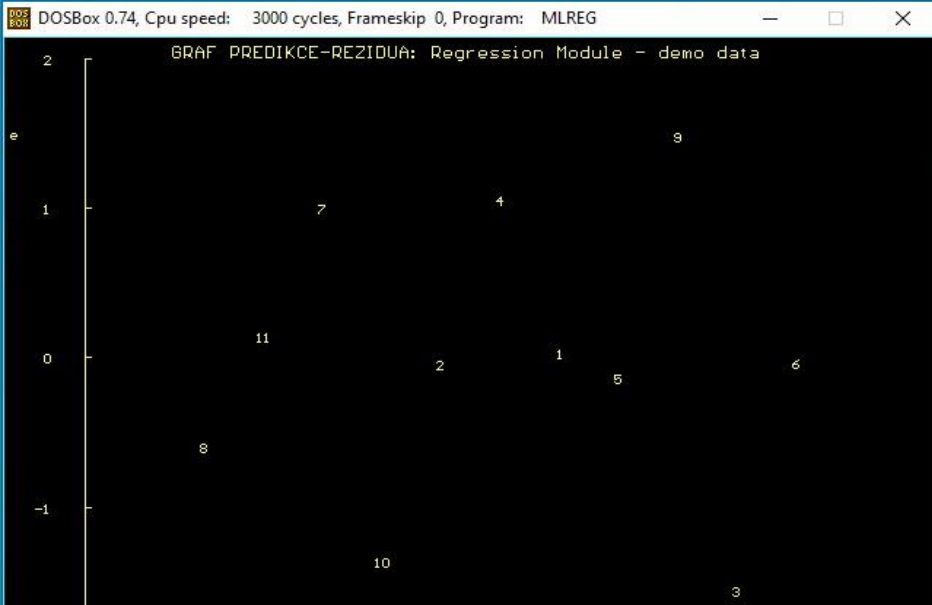


Běžte do bloku Grafy a pak zvolte Graf regresní křívky pomocí 2x kliknutí na Enter.



Kliknutí na mezerník dá síť a Esc.

Zobrazení modelu



Esc vrátí zpět a pak zvolte Residual-Index plot. Vložte síť mezerníkem. Pak Esc zpět.

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

U V Y S L E D K Y

(3) ODHADY PARAMETRŮ A TESTY VÝZNAMNOSTI:

Parametr	Odhad	Směrodatná odchylka	Test H0: $BI_{ij} = 0$ vs. HA: $BI_{ij} <> 0$	t-kriterium	hypoteza H0 je	Hlad. výz.
BI 01	3.0001E+00	1.1247E+00	2.6673E+00	Zamítnuta	0.026	
BI 11	5.0009E-01	1.1791E-01	4.2415E+00	Zamítnuta	0.002	

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R	: 8.1642E-01
Koeficient determinace, R ²	: 6.6654E-01
Predikovaný korelační koeficient, R _p ²	: 7.0808E-01
Střední kvadratická chyba predikce, MEP	: 1.8708E+00
Akaikeho informační kritérium, AIC	: 6.4647E+00

(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	8.0400E+00	8.0010E+00	3.9105E-01	3.9000E-02	4.8507E-01
2	6.9500E+00	7.0008E+00	3.9105E-01	-5.0818E-02	-7.3120E-01
3	7.5800E+00	9.5013E+00	6.0120E-01	-1.9213E+00	-2.5347E+01
4	8.8100E+00	7.5009E+00	3.7285E-01	1.3091E+00	1.4859E+01
5	8.3300E+00	8.5011E+00	4.4116E-01	-1.2109E-01	-2.0539E+00

Ve Výsledky táhněte dolů na bloky (3) a (4).

U V Y S L E D K Y

Bod	Meřená hodnota	Predikovaná hodnota	Směrodatná odchylka	Klasické reziduum	Relativní reziduum
i	yexp[i]	yvyp[i]	s(yvyp[i])	e[i]	er[i]
1	8.0400E+00	8.0010E+00	3.9105E-01	3.9000E-02	4.8507E-01
2	6.9500E+00	7.0008E+00	3.9105E-01	-5.0818E-02	-7.3120E-01
3	7.5800E+00	9.5013E+00	6.0120E-01	-1.9213E+00	-2.5347E+01
4	8.8100E+00	7.5009E+00	3.7285E-01	1.3091E+00	1.4859E+01
5	8.3300E+00	8.5011E+00	4.4116E-01	-1.2109E-01	-2.0539E+00
6	7.2800E+00	5.0005E+00	8.3734E-01	7.1013E-01	1.7302E+01
7	1.0840E+01	9.0012E+00	5.1394E-01	1.8388E+00	1.6963E+01
8	4.8200E+00	6.5007E+00	4.4116E-01	-1.6807E+00	-3.4870E+01
9	5.6800E+00	5.5005E+00	6.0120E-01	1.7945E-01	3.1594E+00

Rezidualní součet čtverců, RSC : 1.3763E+01
Průměr absolutních hodnot reziduí, Me : 8.3740E-01
Průměr relativních reziduí, Mer : 1.2126E+01
Odhad reziduálního rozptylu, s²(e) : 1.5292E+00
Odhad směrodatné odchylky reziduí, s(e) : 1.2366E+00
Odhad šikmosti reziduí, g1(e) : -1.2216E-01
Odhad špičatosti reziduí, g2(e) : 2.2438E+00

Napověda-F1 Řádek: 80 - 102 Celkem: 180 Délka: 9464

Táhněte na blok (5) Analýza klasických reziduí

U V Y S L E D K Y

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

Fisher-Snedocorův test významnosti regrese, F	: 1.7990E+01
Tabulkový kvantil, F(1-alpha, m-1, n-m)	: 5.1174E+00
Závěr: Navržený model je přijat jako významný.	
Spočtená hladina významnosti	: 0.002
Scottovo kritérium multikolinearity, M	: 9.8717E-17
Závěr: Navržený model je korektní.	
Cook-Weisbergův test heteroskedasticity, Sf	: 1.3551E+00
Tabulkový kvantil, Chi ² (1-alpha, 1)	: 3.8415E+00
Závěr: Rezidua vykazují homoskedasticitu.	
Spočtená hladina významnosti	: 0.244
Jarque-Berraův test normality reziduí, L(e)	: 2.8943E-01
Tabulkový kvantil, Chi ² (1-alpha, 2)	: 5.9915E+00
Závěr: Normalita je přijata.	
Spočtená hladina významnosti	: 0.865
Waldův test autokorelace, Wa	: 6.4778E+00
Tabulkový kvantil, Chi ² (1-alpha, 1)	: 3.8415E+00
Závěr: Rezidua jsou autokorelována.	

Táhněte na blok (6) Testování regresního tripletu.

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

U Ý S L E D K Y

(7) INDIKACE VLIVNÝCH BODŮ:
(* indikuje odlehlý nebo vlivný bod)

Bod	Standardizované reziduum eS[i]	Jackknife reziduum eJ[i]	Predikované reziduum eP[i]	Diagonální prvky H[i, i]
1	3.3244E-02	3.1345E-02	4.3333E-02	1.0000E-01
2	-4.3318E-02	-4.0845E-02	-5.6465E-02	1.0000E-01
3	-1.7779E+00	-2.0811E+00	-2.5160E+00	2.3636E-01
4	1.1103E+00	1.1268E+00	1.4400E+00	9.0909E-02
5	-1.4810E-01	-1.3980E-01	-1.9604E-01	1.2727E-01
6	-4.0509E-02	-3.8196E-02	-6.0667E-02	3.1818E-01
7	1.1019E+00	1.1170E+00	1.4981E+00	1.7273E-01
8	-7.2516E-01	-7.0458E-01	-1.0860E+00	3.1818E-01
9	1.6349E+00	1.8383E+00	2.2227E+00	1.7273E-01
10	-1.4549E+00	-1.5685E+00	-1.9258E+00	1.2727E-01
11	1.6607E-01	1.5681E-01	2.3500E-01	2.3636E-01

Bod	Zobecněné diag. prvky Hm[i, i]	Cookova vzdálenost D[i]	Atkinsonova vzdálenost A[i]	Vliv na predikci DF[i]
1	1.0011E-01	6.1398E-05	2.2164E-02	1.0448E-02
2	1.0019E-01	1.0425E-04	2.8882E-02	-1.3615E-02

Napověda-F1 Řádek: 135 - 157 Celkem: 180 Délka: 9464

Táhněte na blok (7) Indikace vlivných bodů, kde jsou diagnostiky v číselné formě. Ukončete úlohu klikem na Esc.

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

REGRESNÍ MODUL : Regresní diagnostika

Data Metoda Zadání Výpočet Výsledky Grafy **Konec**

Ne
Ano

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

A D S T A T - Hlavní menu statistického systému
(c) TriloByte 1990

- Jednorozměrná data
- Vícerozměrná data
- Kalibrace
- Lineární regrese**
- Nelineární regrese
- Vyhlazování
- Pravděpodobnostní modely
- Růstové křivky
- Chemometrie
- Analýza rozptylu
- Konec práce

Analýzu dat úlohy P608a a návrat do červeného menu ADSTATu k analýze další regresní úlohy provedete kliknutím na písmena K a potom A.

Pro Lineární regresi nové úlohy dejte Enter.

Konstrukce intervalů spolehlivosti

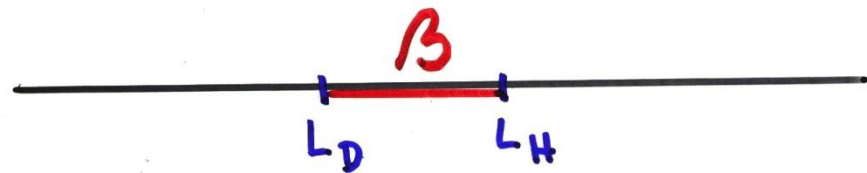
Bodové odhady parametrů β mají pro praxi menší význam.

Intervaly spolehlivosti (konfidenční oblasti), ve kterých leží teoretická hodnota β se zvolenou pravděpodobností $(1 - \alpha)$.

1. Interval spolehlivosti parametrů β

Pro hladinu významnosti $\alpha = 0.10, 0.05$ nebo 0.01 odpovídají 90%ní, 95%ní nebo 99%ní intervaly spolehlivosti.

$\alpha = 0.10$ odpovídá
90%ní jistota



$\alpha = 0.05$ odpovídá
95%ní jistota



$\alpha = 0.01$ odpovídá
99%ní jistota



Intervalové odhady regresních parametrů

Pro různý počet pozorování se mohou odhadnuté regresní parametry b_0 a b_1 lišit.

Vedle bodových odhadů regresních parametrů lze vyčíslit i jejich **intervalové odhady**:

$$b_i - t_{1-\alpha/2}(n-m) \cdot s(b_i) < \beta_i < b_i + t_{1-\alpha/2}(n-m) \cdot s(b_i)$$

kde b_i je bodový odhad regresního parametru,

$t_{1-\alpha/2}(n-p)$ je kvantil Studentova t rozdělení,

m je počet parametrů modelu,

$s(b_i)$ je směrodatná chyba odhadu parametru.

Příklad: Český hydrometeorologický ústav v Praze měřil ...

95% interval spolehlivosti pro parametr b_0 .

	<i>Parametry</i>	<i>Směrodat. odchylka</i>	<i>t exp</i>	<i>Hodnota P</i>	<i>Dolní 95%</i>	<i>Horní 95%</i>
Hranice	10,79504	0,446866	24,15722	5,3E-08	9,738368	11,85171
Nadm. výška	-0,00541	0,00134	-4,03872	0,004941	-0,00858	-0,00224

$$b_i - t_{1-\alpha/2}(n-m) \cdot s(b_i) < \beta_i < b_i + t_{1-\alpha/2}(n-m) \cdot s(b_i)$$

Interval spolehlivosti lze vyčíslit ručně podle tohoto vzorce,

nebo jej přímo přečíst z výstupu.

Příklad: Český hydrometeorologický ústav v Praze měřil ...

95% interval spolehlivosti pro oba parametry.

	<i>Koeficienty</i>	<i>Chyba stř. hodnoty</i>	<i>t stat</i>	<i>Hodnota P</i>	<i>Dolní 95%</i>	<i>Horní 95%</i>
Hranice	10,79504	0,446866	24,15722	5,3E-08	9,738368	11,85171
Nadm. výška	-0,00541	0,00134	-4,03872	0,004941	-0,00858	-0,00224

Výklad úseku: V nadmořské výšce 0 metrů n.m. by se s pravděpodobností 95% měla průměrná teplota půdy nacházet v rozmezí 9,738 °C až 11,852 °C.

Výklad směrnice: S každým dalším metrem nadmořské výšky klesá s pravděpodobností 95% průměrná teplota půdy v rozmezí od 0,00858 °C do 0,00224 °C.

IS y -hodnot – Working-Hottelingův pás spolehlivosti

udává rozpětí, ve kterém se budou nacházet hodnoty závisle proměnné se zvolenou pravděpodobností $1 - \alpha$

$$Y_{i(\min, \max)} = y'_i \pm t_{\frac{\alpha}{2}; n-m} \cdot \sigma$$

TEST VÝZNAMNOSTI REGRESNÍHO MODELU – co testujeme?

$$Y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \dots + b_mx_m$$

Testuje se významnost odhadů jednotlivých parametrů:

Test: Jestliže je testovaný odhad parametru statisticky nevýznamný, pak jeho příslušná proměnná x_j **nepřispívá** ke zpřesnění odhadu závisle proměnné y a tato proměnná x_j je v modelu zbytečná.

Testuje se model jako celek:

Test: zda příslušná kombinace všech nezávisle proměnných statisticky významně zpřesní odhad závisle proměnné y oproti použití pouhého průměru všech hodnot y .

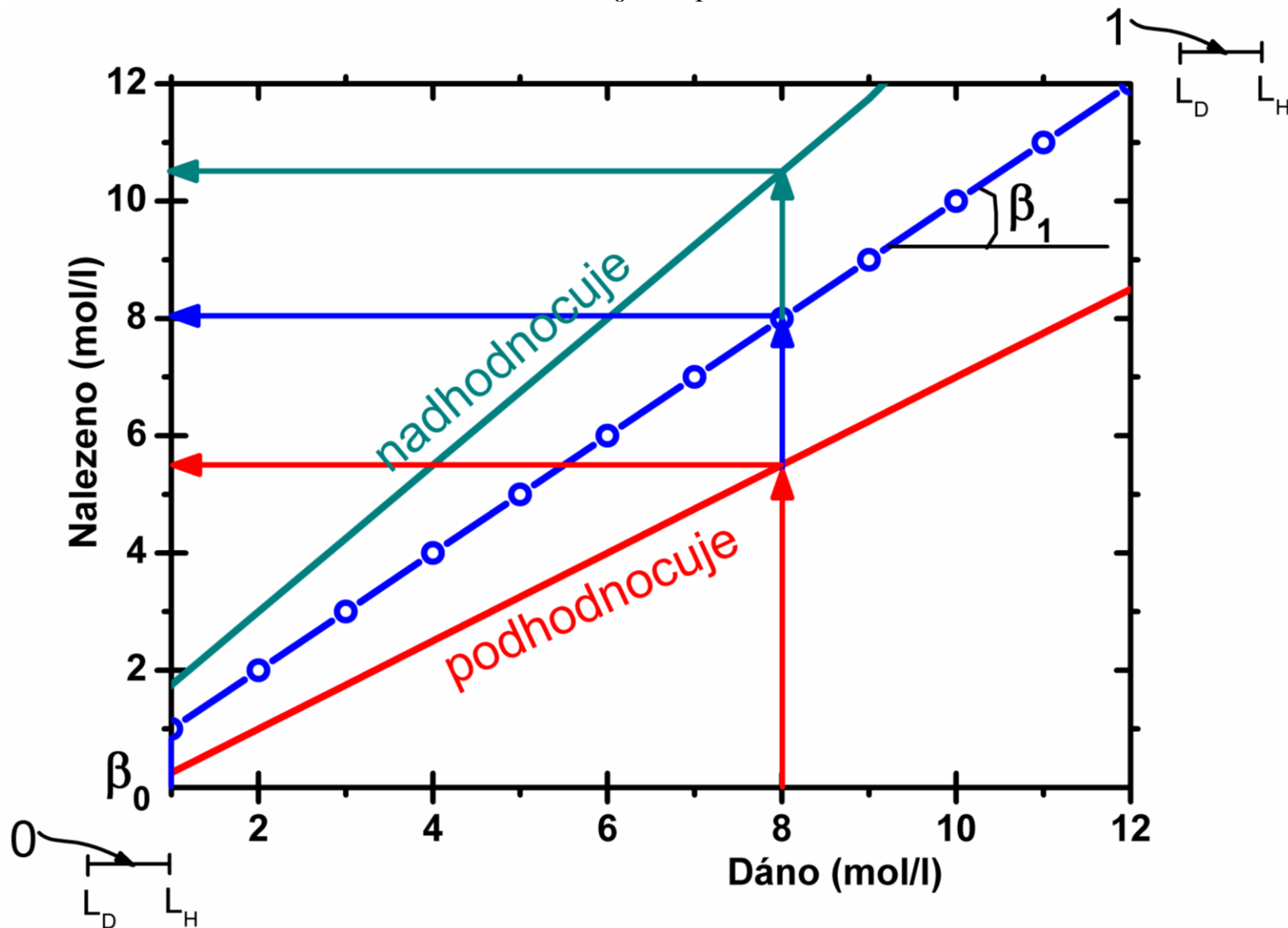
Regresní analýza v programu Validace v programu ADSTAT 1.25

Jak vysvětlit výstup programu a napsat semestrální práci je
v *Kompendiu*, str. 609

Návod k sestavení 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. 609 – 618 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.

Validace nové analytické metody

$$y(\text{nalezeno}) = \beta_0 + \beta_1 \cdot x(\text{dáno})$$



Příklad 6.7 Validizace nové analytické metody

Proveďte validizaci nové analytické metody porovnáním jejich výsledků y vůči standardům x .

- Určete odhady b_1 a b_2 ,
- zkonstruujte 95%ní interval spolehlivosti úseku a směrnice,
- 95%ní elipsoid spolehlivosti,
- 95%ní interval spolehlivosti predikce v těžišti.

Data: $n = 24$, $m = 2$

Obsah látky určený standardní metodou (x) a novou metodou (y) v g:

x	40.2	43.8	47.6	50.7	56.8	81.3	83.3	97.1	102.5	118.7
y	48.9	39.1	42.6	56.9	70.3	71.5	97.6	99.9	105.2	102.3
x	129.4	184.8	287.5	295.4	420.3	421.3	427.9	566.1	608.5	640.7
y	106.8	162.9	234.0	303.4	388.8	391.1	369.3	611.6	580.2	643.3
	x		692.8		705.2		714.4		881.4	
	y		596.6		612.6		633.5		669.8	

Řešení:

- (a) Odhady b_1 a b_2 :
- odhad úseku $b_2 = 14.73 (\pm 12.61)$,
 - odhad směrnice $b_1 = 0.868 (\pm 0.0302)$,
 - koeficient determinace $\hat{R}^2 = 0.974$,
 - odhad směrodatné odchylky reziduí $\hat{\sigma} = 39.54$.

(b) Interval spolehlivosti úseku: $H_0: \beta_2 = 0$ vs. $H_A: \beta_2 \neq 0$

$$b_2 - t_{1-\alpha/2}(22) \sqrt{D(b_2)} \leq \beta_2 \leq b_2 + t_{1-\alpha/2}(22) \sqrt{D(b_2)}$$

a dosazením

$$14.73 - 2.08 \times 12.61 \leq \beta_2 \leq 14.73 + 2.08 \times 12.61$$

vyjde

$$-11.499 \leq \beta_2 \leq 40.959$$

Závěr testování: interval spolehlivosti úseku zahrnuje nulu, takže H_0 je přijata a úsek β_2 lze považovat za nulový.

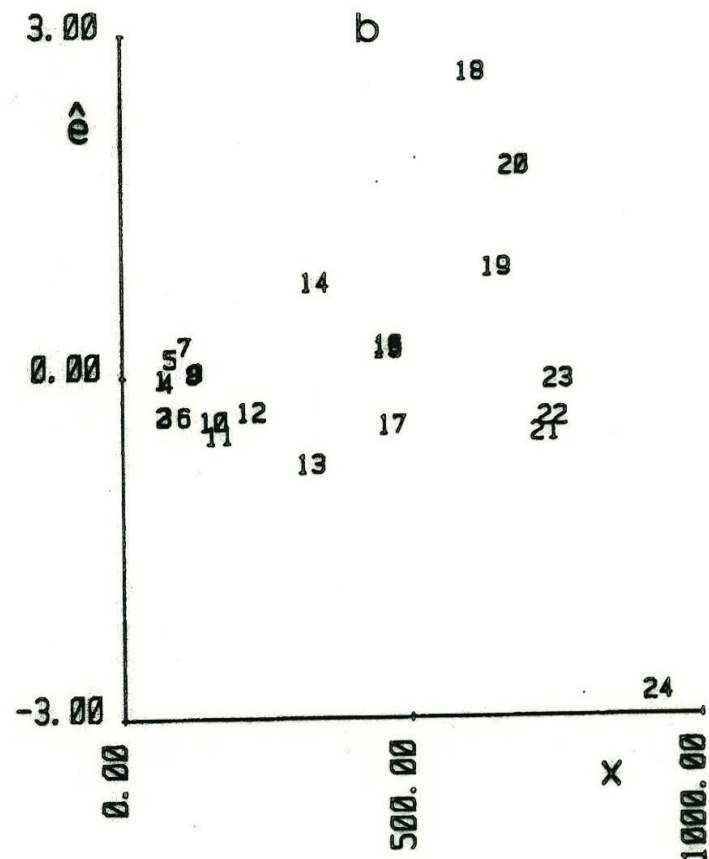
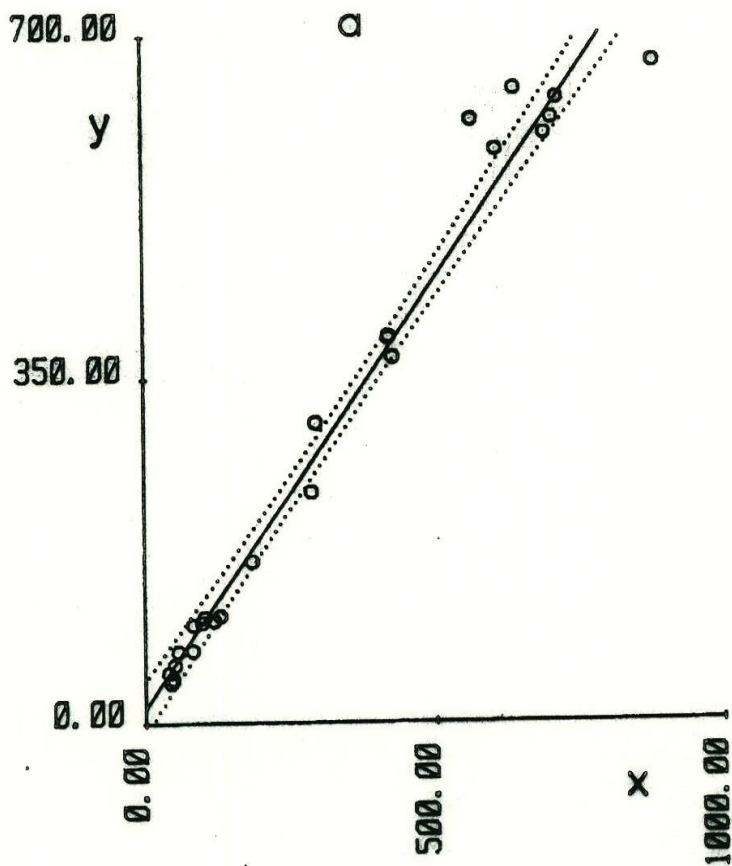
Interval spolehlivosti směrnice: $H_0: \beta_1 = 1$ vs. $H_A: \beta_1 \neq 1$

$$0.868 - 2.08 \times 0.0302 \leq \beta_1 \leq 0.868 + 2.08 \times 0.0302$$

a vyčíslením

$$0.805 \leq \beta_1 \leq 0.93$$

Závěr testování: interval spolehlivosti neobsahuje jedničku, takže H_0 je zamítnuta a směrnici β_1 nelze považovat za jednotkovou.



Závěr: intervaly spolehlivosti indikují, že úsek regresní přímky lze považovat za nulový $\beta_2 = 0$, zatímco směrnice β_1 je významně odlišná od jedničky.

Nová analytická metoda vede k odlišným výsledkům od standardní.

Tento postup na dvou stranách si před analýzou dat v ADSTATu vytiskněte!

Návod k sestavení a napsání semestrální práce k této úloze najdete ve cvičebnici *M. Meloun, J. Miličák: Kompendium statistického zpracování dat, Karolinum Praha 1992, ISBN 978-80-246-2196-8*, ve vzorovém příkladu na str. 609 – 618 a obecný výklad outputu v lineární regresi na str. 582 - 583. V ADSTATu používejte na přenos výstupů a obrázků na papír **PrintScreen** a upravte pak negativ na černobílý pozitiv.

DC

6.2 VALIDACE – - Postup analýzy dat v ADSTAT 1.25

Příklad 6.7 Validace nové analytické metody (Je z učebnice nikoliv z Kompendia)

- 1. Modul:** V červeném menu zvolte **Lineární regrese** a pak klávesou **Enter** otevřete menu **Data**.
- 2. Zadání dat:** Nyní stiskněte klávesu **F3** a smažte *demo.dat* a napište **P607.txt** a ukončete **Esc** a **A**.
- 3. Metoda:** V nabídce zvolte **Regresní diagnostika** a **Enter**.
- 4. Zadání:** Ve **Podmínky** nechte vše předvolené a 7x **Enter**. Ve **Volby** nechte vše předvolené a 7x **Enter**.
- 5. Výpočet:** V menu **Výpočet** dejte **Enter**.
- 6. Grafy:** V menu na **Grafy** dejte **Enter** a prohlédněte si nejprve **Graf regresní křivky** a **Residual-Index plot**, u kterého stiskněte mezerník pro síť os. Všimněte si podezřelých odlehlých bodů 18, 20, 24. Rozhodnou grafy odlehlých bodů o vyloučení odlehlých bodů. Dejte **Esc** a zvolte **Grafy vlivných bodů**, které detekují 18, 20, 24 k odstranění.
- 7. Výsledky:** dejte **Enter** a z **Výstupu** vyberte a запиšte si k pozdějšímu porovnání: **(3) Odhady parametrů** a **(4) Statistické charakteristiky regrese** a **(5) Analýza klasických reziduí** a **(6) Testování regresního** **tržbu**.

Celý postup nyní opakujte pro data bez odlehlých hodnot

2. Zadání dat: Vraťte se nyní do bloku **Data**, pak **Enter** a **F4** a **F5**. Najed'te na řádek 18 a ťukejte **Delete** až po obou číselech zůstane prázdné místo. Totéž opakujte v řádku 20 a 24. Pak stiskněte **Esc** a **A**.

3. Metoda: V nabídce zvolte **Regresní diagnostika** a **Enter**.

4. Zadání: V **Podmínky** nechte vše předvolené a dejte 7x **Enter**. Ve **Volby** nechte vše předvolené a 7x **Enter**.

5. Výpočet: V menu **Výpočet** dejte **Enter**.

6. Grafy: V menu na **Grafy** dejte **Enter** a prohlédněte si nejprve **Graf regresní křivky** a **Residual-Index plot**, u kterého stiskněte mezerník sít' os.

7. Výsledky: dejte **Enter** a z **Výstupu** vyberte a zapište si k porovnání s předešlým výstupem: **(3) Odhady parametrů** a **(4) Statistické charakteristiky regrese** a **(5) Analýza klasických reziduí** a **(6) Testování regresního tripletu**.

Podrobný rozbor výstavby regresního modelu s výkladem celého výstupu programu ADSTAT je ve cvičebnici Kompendium a začíná na str. 609 a dále.

A D S T A T - Hlavní menu statistického systému
(c) TriloByte 1990

DC

Jednorozměrná data
 Vícerozměrná data
 Kalibrace
Lineární regrese
 Nelineární regrese
 Uhlazování
 Pravděpodobnostní modely
 Růstové křivky
 Chemometrie
 Analýza rozptylu
 Konec práce

Pro Lineární regrese dejte **Enter**.

Numerický editor

	1	2
1	4.020000 E+01	4.890000 E+01
2		
3		
4	Čtení ze souboru: P607	
5		
6	5.680000 E+01	7.030000 E+01
7	8.130000 E+01	7.150000 E+01
8	8.330000 E+01	9.760000 E+01
9	9.710000 E+01	9.990000 E+01
10	1.025000 E+02	1.052000 E+02
11	1.187000 E+02	1.023000 E+02
12	1.294000 E+02	1.068000 E+02
13	1.848000 E+02	1.629000 E+02
14	2.875000 E+02	2.340000 E+02
15	2.954000 E+02	3.034000 E+02
16	4.203000 E+02	3.888000 E+02
17	4.213000 E+02	3.911000 E+02

Otevřít **Data**, pak stisk **F3**, smazat *demo.dat*, napsat **P607** a ukončit **Esc** a **A**.

Rádek	1	Sloupec	5	Insert	NUM	P607
P607x	P607y					
40	2.20000	48	9.0000			
43	8.0000	39	1.0000			
47	6.0000	42	6.0000			
50	7.0000	56	9.0000			
56	8.0000	70	3.0000			
81	3.0000	71	5.0000			
83	3.0000	97	6.0000			
97	7.10000	99	9.0000			
102	5.0000	105	2.0000			
118	7.0000	102	3.0000			
129	4.0000	106	8.0000			
184	8.0000	162	9.0000			
287	5.0000	234	0.0000			
295	4.0000	303	4.0000			
420						
421						
427						
566	1.0000	611	6.0000			
608	5.0000	580	2.0000			
640	7.0000	643	3.0000			
692	8.0000	596	6.0000			
705	2.0000	612	6.0000			

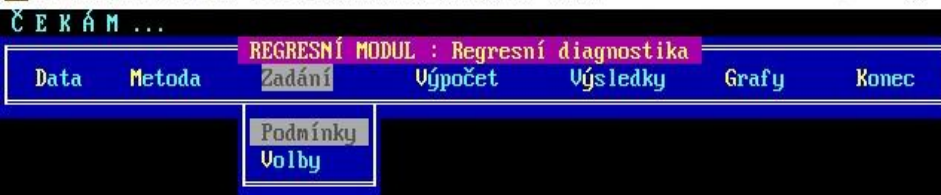
Otevřít **Data** a **F5** a ukončit **Esc** a **A**.

REGRESNÍ MODUL : Lineární nejmenší čtverce

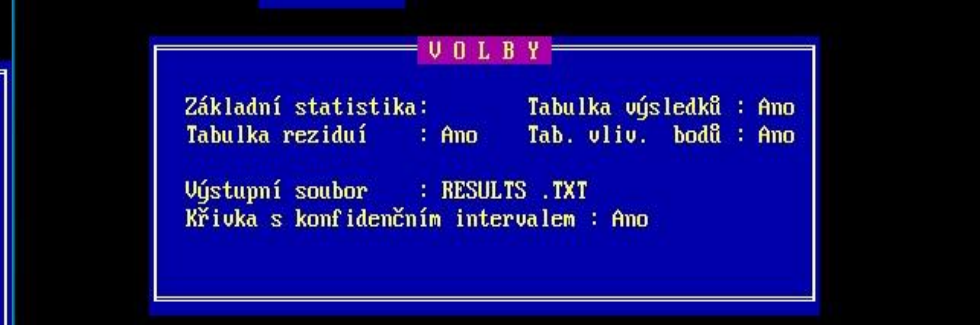
Data	Metoda	Zadání	Úpočet	Úsledek	Grafy	Konec
	Lineární regrese MNC					
	Regresní diagnostika					

Blok **Metoda** otevřít klávesou **Enter** a zvolit **Regresní diagnostika** a **Enter**.

Lineární regresní diagnostika



V Podmínky nechte vše předvolené defaultní a dejte 7x Enter.



Ve Volby nechte vše default a dejte 7x Enter a pak ve Výpočet klikněte opět Enter.



Běžte do Grafy a zvolte Graf regresní křivky pomocí 2x kliknutí na Enter.



Kliknutí na mezerník dá síť a Esc.



Esc vrátí zpět a zvolte Residual-Index plot. Vložte síť mezerou. Pak Esc zpět.

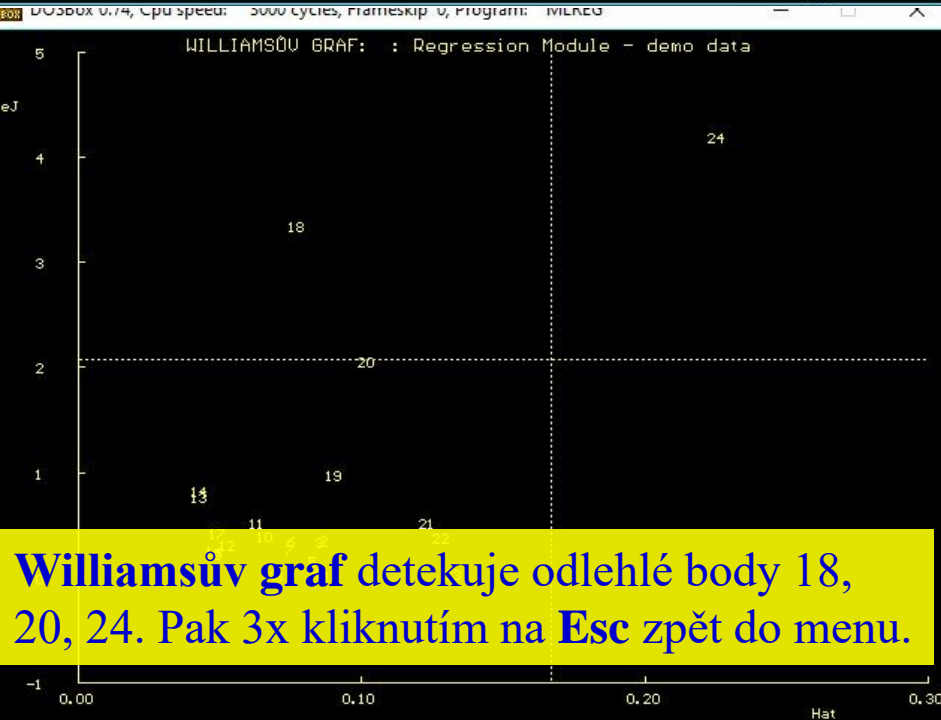
DOSBox 0.74, Cpu speed: 3000 cycles, Frameskip 0, Program: MLREG

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 XI]
- Rezid. komponenty vs XI]
- Grafy vlivných bodů
- Graf predikovaných reziduí
- Pregibonův graf
- Williamsův graf**
- McCulloh-Meeterův graf
- L-R graf

Z 9 grafů otevřete klikem na Enter Grafy vlivných bodů, pak si otevřete všech pět kliknutím na Enter, nakonec ponechte otevřený nejlepší Williamsův graf.



Williamsův graf detekuje odlehlé body 18, 20, 24. Pak 3x kliknutím na Esc zpět do menu.

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

U 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.
B[0]	1.4729E+01	1.2610E+01	1.1680E+00		Akceptována	0.255
B[1]	8.6840E-01	3.0205E-02	2.8750E+01		Zamítnuta	0.000

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R	: 9.8695E-01
Koeficient determinace, R ²	: 9.7407E-01
Predikovaný korelační koeficient, Rp ²	: 9.8228E-01
Střední kvadratická chyba predikce, MEP	: 1.9417E+03
Akaikeho informační kritérium, AIC	: 1.7843E+02

(5) ANALÝZA KLASICKÝCH REZIDUÍ:

Bod	Meřená hodnota	Predikovaná hodnota	Směrodatná odchylka	Klasické reziduum	Relativní reziduum
i	yexp[i]	yvp[i]	s(yvp[i])	e[i]	er[i]
4	5.6900E+01	5.8757E+01	1.1475E+01	-1.8571E+00	-3.2638E+00

Ve Výsledky táhněte dolů na bloky (3) a (4).

4 5.6900E+01 5.8757E+01 1.1475E+01 -1.8571E+00 -3.2638E+00

Nanověda-F1 Řádek: 72 - 94 Celkem: 240 Délka: 13721

U V Ý S L E D K Y

10	1.0230E+02	1.1781E+02	1.0119E+01	-1.5508E+01	-1.5160E+01
11	1.0680E+02	1.2710E+02	9.9277E+00	-2.0300E+01	-1.9008E+01
12	1.6290E+02	1.7521E+02	9.0563E+00	-1.2310E+01	-7.5567E+00
13	2.3400E+02	2.6439E+02	8.1342E+00	-3.0395E+01	-1.2989E+01
14	3.0340E+02	2.7126E+02	8.1083E+00	3.2145E+01	1.0595E+01
15	3.8880E+02	3.7972E+02	8.6140E+00	9.0813E+00	2.3357E+00
16	3.9110E+02	3.8059E+02	8.6246E+00	1.0513E+01	2.6880E+00
17	3.6930E+02	3.8632E+02	8.6969E+00	-1.7019E+01	-4.6083E+00

Nyní na blok (5) Analýza klasických reziduí

21	5.9660E+02	6.1636E+02	1.3837E+01	-1.9758E+01	-3.3118E+00
22	6.1260E+02	6.2713E+02	1.4143E+01	-1.4527E+01	-2.3713E+00
23	6.3350E+02	6.3512E+02	1.4372E+01	-1.6159E+00	-2.5507E-01
24	6.6980E+02	7.8014E+02	1.8760E+01	-1.1034E+02	-1.6473E+01

Rezidualní součet čtverců, RSC : 3.4403E+04
 Průměr absolutních hodnot reziduí, Me : 2.3777E+01
 Průměr relatiivních reziduí, Mer : 1.0198E+01
 Odhad rezidualního rozptylu, s²(e) : 1.5638E+03
 Odhad směrodatné odchylky reziduí, s(e) : 3.9545E+01
 Odhad šikmosti reziduí, g1(e) : 1.6990E-01
 Odhad špičatosti reziduí, g2(e) : 6.1368E+00

U V Ý S L E D K Y

(6) TESTOVÁNÍ REGRESNÍHO TRIPLETU (DATA + MODEL + METODA):
 Fisher-Snedocorův test významnosti regrese, F : 8.2658E+02
 Tabulkový kvantil, F(1-alpha,m-1,n-m) : 4.3009E+00
 Závěr: Navržený model je přijat jako významný.
 Spočtená hladina významnosti : 0.000
 Scottovo kritérium multikolinearity, M : -5.5015E-16
 Závěr: Navržený model je korektní.
 Cook-Weisbergův test heteroskedasticity, Sf : 1.1945E+02
 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) : 9.9551E+00
 Tabulkový kvantil, Chi²(1-alpha,2) : 5.9915E+00
 Závěr: Normalita není přijata.
 Spočtená hladina významnosti : 0.007

Táhněte na blok (6) Testování regresního tripletu .

U V Ý S L E D K Y

(7) INDIKACE VLIVNÝCH BODŮ:
 (* indikuje odlehlý nebo vlivný bod)

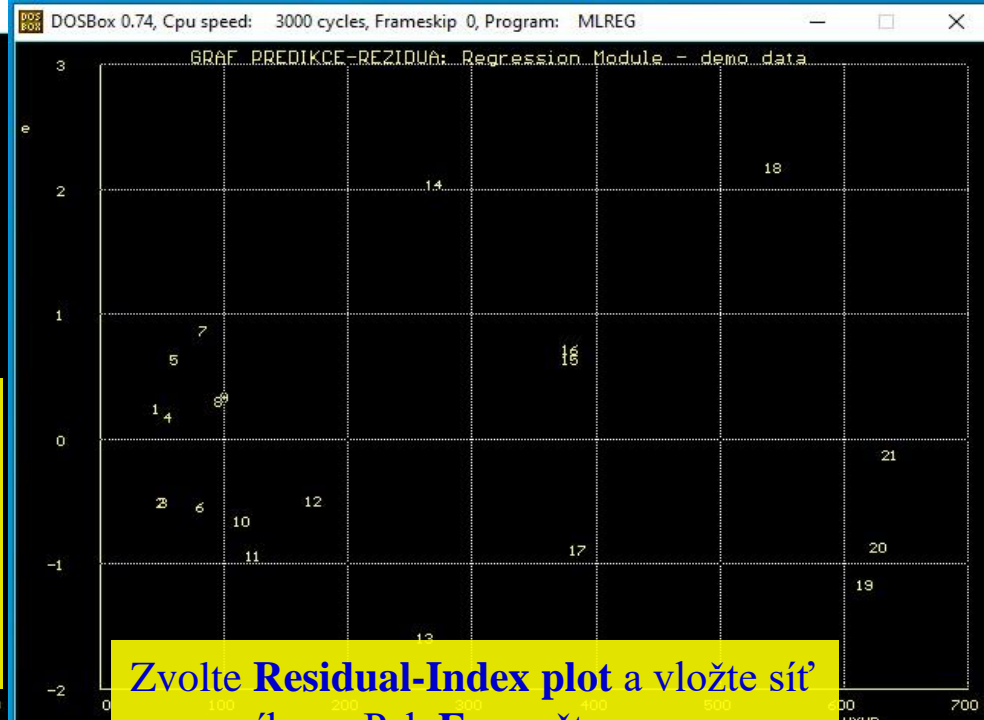
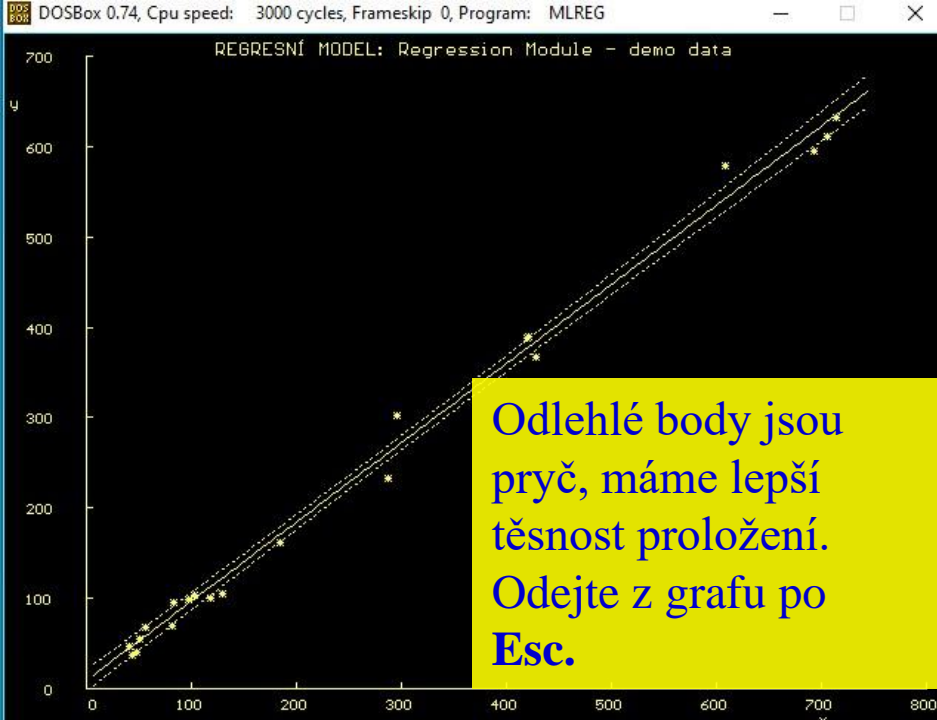
Bod	Standardizované reziduum	Jackknife reziduum	Predikované reziduum	Diagonální prvky
i	eS[i,i]	eJ[i,i]	eP[i,i]	H[i,i]
1	-1.9560E-02	-1.9110E-02	-8.0977E-01	8.7582E-02
2	-3.6153E-01	-3.5427E-01	-1.4958E+01	8.6412E-02
3	-3.5600E-01	-3.4882E-01	-1.4719E+01	8.5192E-02
4	-4.9073E-02	-4.7947E-02	-2.0278E+00	8.4210E-02
5	1.6487E-01	1.6118E-01	6.8059E+00	8.2309E-02
6	-3.6366E-01	-3.5637E-01	-1.4953E+01	7.5114E-02
7	2.7688E-01	2.7098E-01	1.1382E+01	7.4558E-02
8	2.2274E-02	2.1762E-02	9.1379E-01	7.0846E-02
9	3.8264E-02	3.7386E-02	1.5686E+00	6.9453E-02
10	-4.0568E-01	-3.9784E-01	-1.6595E+01	6.5481E-02

Táhněte na blok (7) Indikace vlivných bodů, kde jsou diagnostiky v číselné formě. Ukončete úlohu Esc.

F1=napověda F2=uložení F3=čtení F4=numerický mód F5=zoom ESC=konec
 Řádek 24 Sloupec 1 Insert NUM P607

4.380000	E+01	3.910000	E+01
4.760000	E+01	4.260000	E+01
5.070000	E+01	5.690000	E+01
5.680000	E+01	7.030000	E+01
8.130000	E+01	7.150000	E+01
8.330000	E+01	9.760000	E+01
9.710000	E+01	9.990000	E+01
1.025000	E+02	1.052000	E+02
1.187000	E+02	1.023000	E+02
1.294000	E+02	1.068000	E+02
1.848000	E+02	1.629000	E+02
2.875000	E+02	2.340000	E+02
2.954000	E+02	3.034000	E+02
4.203000	E+02	3.888000	E+02
4.213000	E+02	3.911000	E+02
4.279000	E+02	3.693000	E+02
5.085000	E+02	5.802000	E+02
6.928000	E+02	5.966000	E+02
7.052000	E+02	6.126000	E+02
7.144000	E+02	6.335001	E+02

Opakujte nyní celý výpočet po odstranění odlehlých bodů 18, 20, 24 v bloku Data, která otevřete klikem na Enter a pak F4 a F5.



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

VÝSLEDKY

(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	9.4132E+00	5.6728E+00	1.6594E+00		Akceptována	0.113
BI 11	8.7634E-01	1.5894E-02	5.5137E+01		Zamítnuta	0.000

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R : 9.9689E-01
 Koeficient determinace, R² : 9.9379E-01
 Predikovaný korelační koeficient, Rp² : 9.9614E-01
 Střední kvadratická chyba predikce, MEP : 3.3360E+02
 Akaikeho informační kritérium, AIC : 1.2149E+02

(5) ANALÝZA KLASICKÝCH REZIDUÍ:

Bod	Meřená hodnota	Predikovaná hodnota	Směrodatná odchylka	Klasické reziduum	Relativní reziduum
3	4.2600E+01	5.1127E+01	5.1311E+00	-8.5269E+00	-2.0016E+01
4	5.6900E+01	5.3844E+01	5.0978E+00	3.0564E+00	5.3716E+00

Ve Výsledky táhněte dolů na bloky (3) a (4).

Namoučte-F1 Řádek: 22 - 94 Celkem: 228 Délka: 12794

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

VÝSLEDKY

7	9.7600E+01	8.2412E+01	4.7635E+00	1.5188E+01	1.5561E+01
8	9.9900E+01	9.4506E+01	4.6322E+00	5.3943E+00	5.3997E+00
9	1.0520E+02	9.9238E+01	4.5827E+00	5.9621E+00	5.6674E+00
10	1.0230E+02	1.1343E+02	4.4407E+00	-1.1135E+01	-1.0884E+01
11	1.0680E+02	1.2281E+02	4.3528E+00	-1.6011E+01	-1.4992E+01
12	1.6290E+02	1.7136E+02	3.9834E+00	-8.4606E+00	-5.1938E+00
13	2.3400E+02	2.6426E+02	3.7762E+00	2.7761E+01	4.1682E+01
14	3.9110E+02	3.7861E+02	4.4900E+00	1.2485E+01	3.1923E+00
15	3.6930E+02	3.8440E+02	4.5481E+00	-1.5099E+01	-4.0884E+00
16	5.8020E+02	5.4267E+02	6.6027E+00	3.7535E+01	6.4693E+00
17	5.9660E+02	6.1654E+02	7.7415E+00	-1.9941E+01	-3.3424E+00
18	6.1260E+02	6.2741E+02	7.9143E+00	-1.4807E+01	-2.4171E+00
19	6.3350E+02	6.3547E+02	8.0432E+00	-1.9696E+00	-3.1090E-01

Táhněte na blok (5) Analýza klasických reziduí

Rezidualní součet čtverců, RSC : 5.6480E+03
 Průměr absolutních hodnot reziduí, Me : 1.3444E+01
 Průměr relativních reziduí, Mer : 8.9802E+00
 Odhad reziduálního rozptylu, s²(e) : 2.9726E+02
 Odhad směrodatné odchylky reziduí, s(e) : 1.7241E+01
 Odhad šikmosti reziduí, g1(e) : 6.6492E-01
 Odhad špičatosti reziduí, g2(e) : 2.9878E+00

Namoučte-F1 Řádek: 97 - 119 Celkem: 228 Délka: 12794

```
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 : 3.0401E+03
Tabulkový kvantil, F(1-alpha,m-1,n-m) : 4.3807E+00
Závěr: Navržený model je přijat jako významný.
Spočtená hladina významnosti : 0.000

Scottovo kritérium multikolinearity, M : 2.9919E-16
Závěr: Navržený model je korektní.

Cook-Weisbergův test heteroskedasticity, Sf : 8.9361E+01
Tabulkový kvantil, Chi^2(1-alpha,1) : 3.8415E+00
Závěr: Rezidua vykazují heteroskedasticitu.
Spočtená hladina významnosti : 0.000

Ljung-Boxova test normality rezidui, Lj(p) : 1.5476E+00
Tabulkový kvantil, Chi^2(1-alpha,1) : 5.9915E+00
Závěr: Navržený model je přijat.
Spočtená hladina významnosti : 0.461

Waldův test autokorelace, Wa : 1.0230E+00
Tabulkový kvantil, Chi^2(1-alpha,1) : 3.8415E+00
Závěr: Rezidua nejsou autokorelována.
Spočtená hladina významnosti : 0.312

Napověda-F1 Řádek: 122 - 144 Celkem: 228 Délka: 12794
```

Táhněte na blok (6) Testování regresního tripletu.

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

REGRESNÍ MODUL : Regresní diagnostika

Data Metoda Zadání Výpočet Výsledky Grafy **Konec**

Ne
Ano

Pokračovat

Analýzu dat úlohy P607 a návrat do červeného menu ADSTATu k analýze další regresní úlohy provedete kliknutím na písmena **K** a **A**.

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

A D S T A T - Hlavní menu statistického systému
(c) TriloByte 1990

- Jednorozměrná data
- Vícerozměrná data
- Kalibrace
- Lineární regrese
- Nelineární regrese
- Vyhlažování
- Pravděpodobnostní modely
- Růstové křivky
- Chemometrie
- Analýza rozptylu
- Konec práce**

Ukončete analýzu kliknutím na **K** a **A**.

Úlohy k procvičování lineární regresní analýzy

Jak vysvětlit výstup programu a napsat semestrální práci je
v *Kompendiu*, str. 609

Návod k sestavení 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. 609 – 618 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.

6.2.1 Úlohy na validaci nové analytické metody

Úloha V6.01 Validace stanovení molybdenu rentg.-fluoresc. metodou

Zadání: U stanovení obsahu molybdenu porovnejte výsledky z rentg.-fluorescenční metody y s deklarováním obsahem standardů ocelí x .

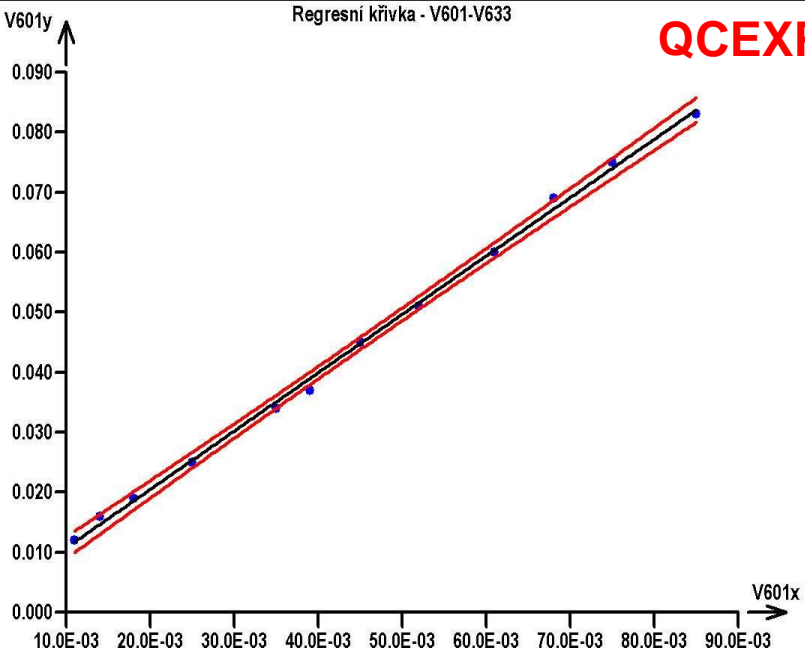
Úkoly:

- (1) Určete velikost systematické chyby metody (= velikost úseku β_0).
- (2) Správnost metody (= směrnice měla být 1).
- (3) Pokuste se vyjádřit i přesnost metody.
- (4) Jsou v datech vlivné a vybočující body?
- (5) Tabulkové indikace vlivných bodů a pět nejdůležitějších grafů identifikace vlivných bodů.

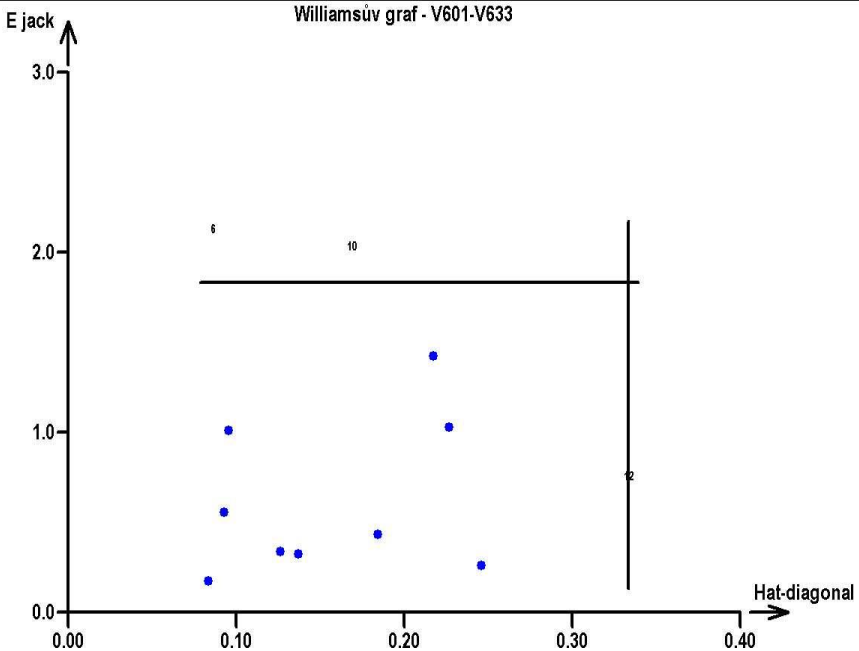
Data: Obsah molybdenu, dáno x [%], stanoveny y [%]:

Dáno x	Stanoveny y
0.011	0.012
...	...
0.085	0.083

Regresní křivka - V601-V633



Williamsův graf - V601-V633



Odhady parametrů

Proměnná	Odhad	Směr.Odch.	Závěr	Pravděpodobnost
Abs	0.001034	0.000686	Nevýznamný	0.163
V601x	0.972702	0.013748	Významný	7.77E-015

Spodní mez

-0.00049559
0.9420701358

Horní mez

0.0025644
1.003335592

Statistické charakteristiky regrese

Vícenásobný korelační koeficient R :	0.99900
Koeficient determinace R ² :	0.99800
Predikovaný korelační koeficient R _p :	0.99434
Střední kvadratická chyba predikce MEP :	1.50063E-006
Akaikeho informační kritérium :	-161.13

Fisher-Snedecorův test významnosti modelu

Hodnota kritéria F :	5005.80
Kvantil F (1-alfa, m-1, n-m) :	4.96460
Pravděpodobnost :	7.75E-015
Závěr :	Model je významný

DC

- Jednorozměrná data
- Vícerozměrná data
- Kalibrace
- Lineární regrese**
- Nelineární regrese
- Vyhlazování
- Pravděpodobnostní modely
- Růstové křivky
- Chemometrie
- Analýza rozptylu
- Konec práce

Pro Lineární regrese dejte Enter.

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

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

Počet řádků: 12 Počet sloupců: 2 Uodorově U601

Numerický editor

1	2
1.100000 E-02	1.200000 E-02
3.500000 E-02	3.400000 E-02
3.900000 E-02	3.700000 E-02
4.500000 E-02	4.500000 E-02
5.200000 E-02	5.100000 E-02
6.100000 E-02	6.000000 E-02
6.800000 E-02	6.900000 E-02
7.500000 E-02	7.500000 E-02
8.500000 E-02	8.300000 E-02

Čtení ze souboru: U601

S=sloupec R=řádek N=nová poloha MEZERA=změna směru
blokové operace: F6=kopírování F7=zaplnění F8=transformace
ostatní operace: F9=změna velikosti matice F10=transpozice

Oteřít Data, pak stisk F3, smazat *demo.dat*, napsat *V601* a ukončit *Esc* a *A*.

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 1 Sloupec 5 Insert NUM U601

Textový editor

U601x	U601y
0.01100	0.01200
0.01400	0.01600
0.01800	0.01900
0.02500	0.02500
0.03500	0.03400
0.03900	0.03700
0.04500	0.04500
0.05200	0.05100
0.06100	0.06000
0.06800	0.06900
0.07500	0.07500
0.08500	0.08300

Vyhledávání chyb: F6=první F7=předchozí F8=následující F9=poslední

Otevřít Data a F5 a ukončit *Esc* a *A*.

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

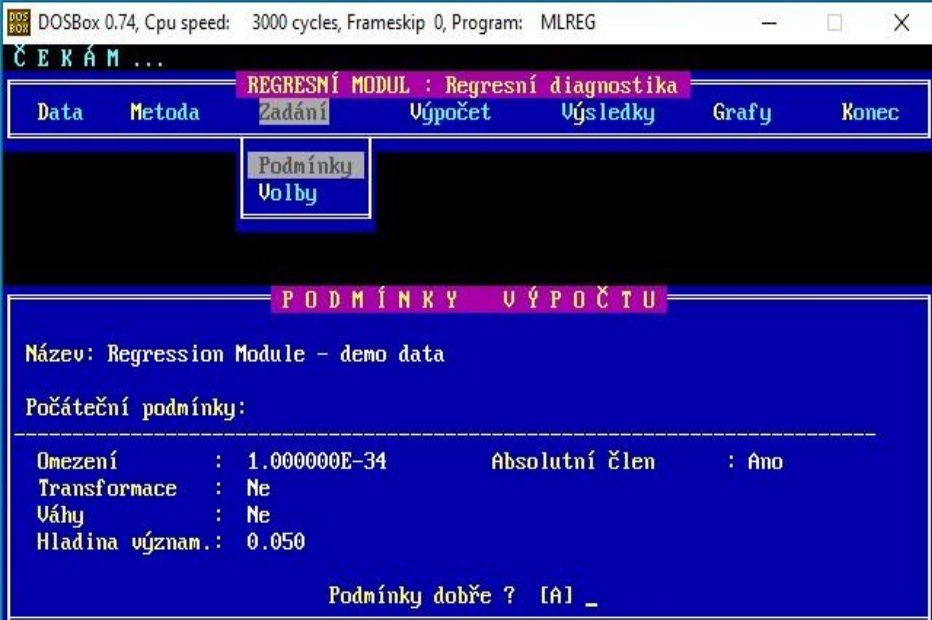
REGRESNÍ MODUL : Lineární nejmenší čtverce

Data **Metoda** Zadání Výpočet Výsledky Grafy Konec

Lineární regrese MMC
Regresní diagnostika

Lineární regresní diagnostika

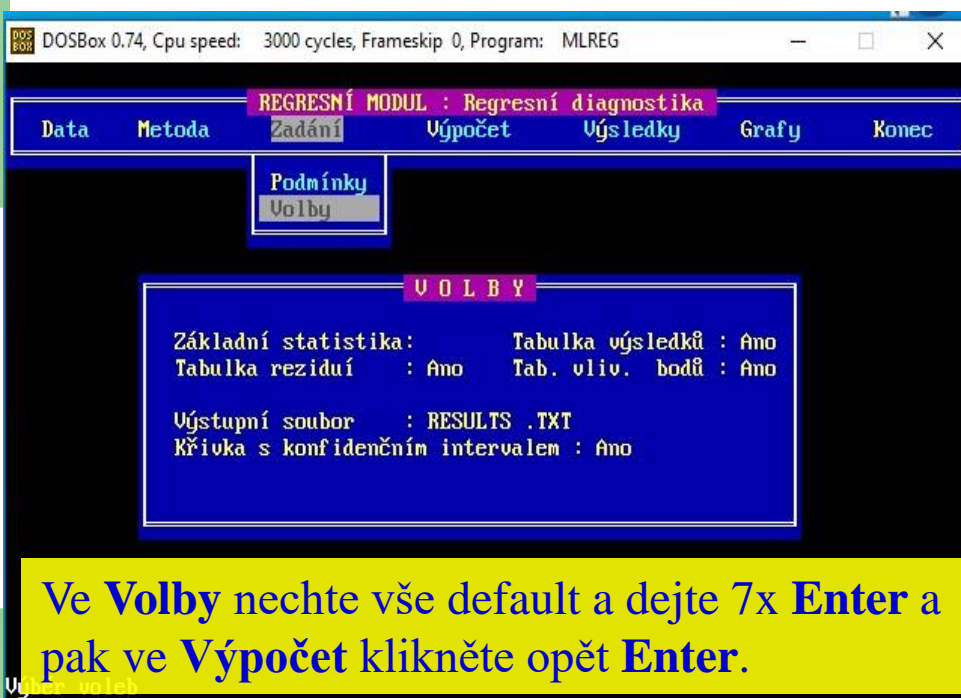
Blok **Metoda** otevřít klávesou **Enter** a zvolit **Regresní diagnostika** a **Enter**.



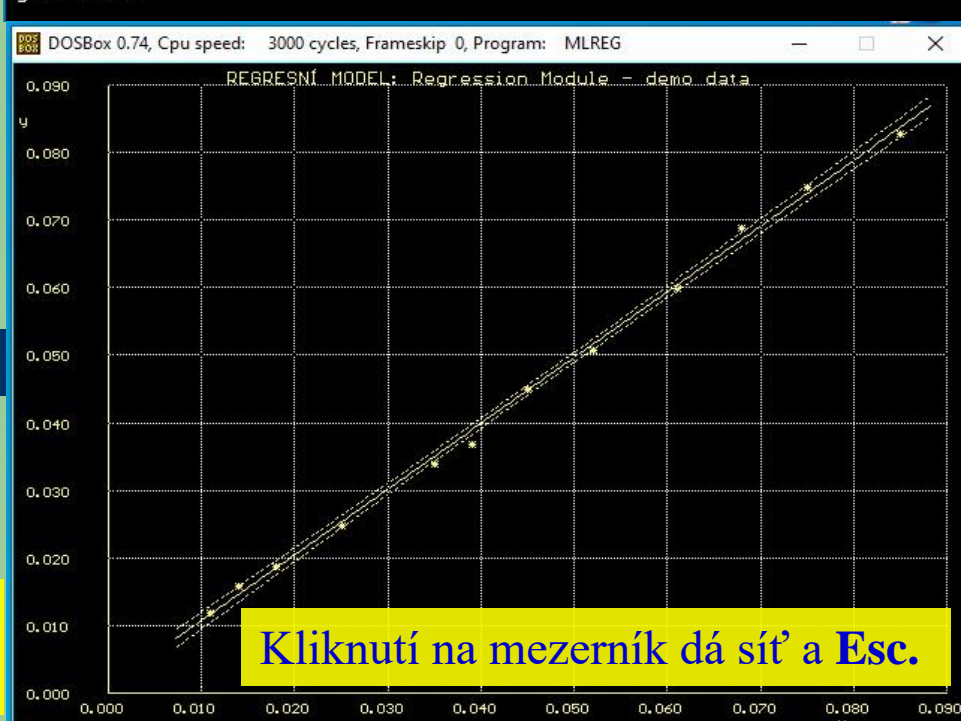
V Podmínky nechte vše předvolené (default) a dejte 7x Enter.



Běžte do Grafy a zvolte Graf regresní křivky pomocí 2x kliknutí na Enter.



Ve Volby nechte vše default a dejte 7x Enter a pak ve Výpočet klikněte opět Enter.



Kliknutí na mezerník dá síť a Esc.



Esc vrátí zpět a zvolte Residual-Index plot. Vložte síť mezeríkem. Pak Esc zpět.

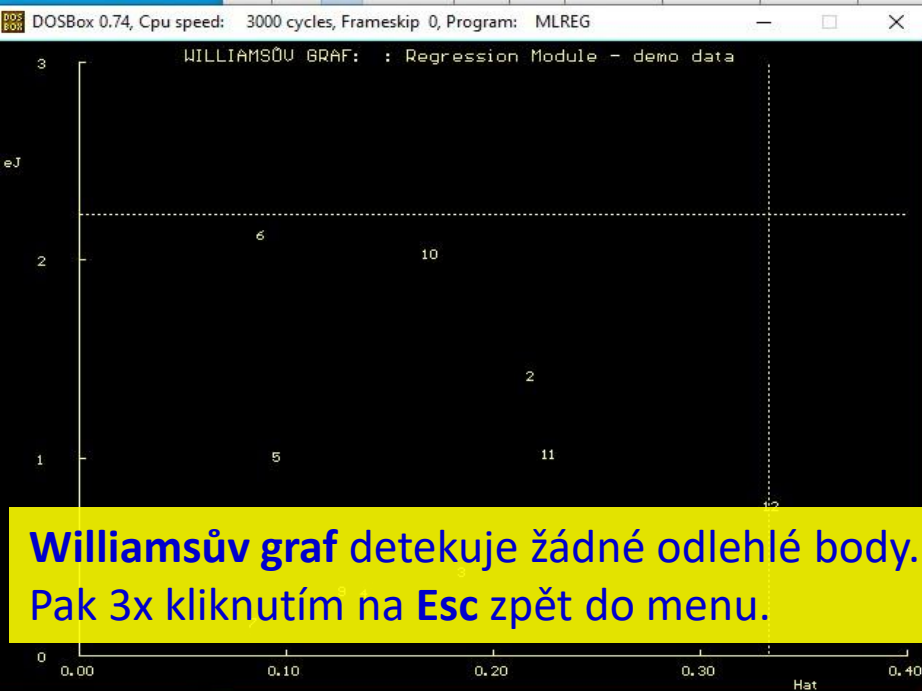
DOSBox 0.74, Cpu speed: 3000 cycles, Frameskip 0, Program: MLREG

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 XI 1
- Rezid. komponenty vs XI 1
- Grafy vlivných bodů
- Graf predikovaných reziduí
- Pregibonův graf
- Williamsův graf
- McCulloh-Meeterův graf
- L-R graf

Z 9 grafů otevřete klikem na Enter blok **Grafy vlivných bodů**, pak si otevřete všech pět grafů kliknutím na **Enter**, nakonec ponechte otevřený nejlepší, a to **Williamsův graf**.



Williamsův graf detekuje žádné odlehlé body. Pak 3x kliknutím na Esc zpět do menu.

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

U V Ý S L E D K Y

(3) ODHADY PARAMETRŮ A TESTY VÝZNAMNOSTI:

Parametr	Odhad	Směrodatná odchylnka	Test H0: B[i,j] = 0 vs. HA: B[i,j] <> 0	t-kriterium	hypoteza H0 je	Hlad. výz.
BI 01	1.0344E-03	6.8667E-04	1.5064E+00		Akceptována	0.163
BI 11	9.7270E-01	1.3748E-02	7.0752E+01		Zamítnuta	0.000

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

- Vícenásobný korelační koeficient, R : 9.9900E-01
- Koeficient determinace, R² : 9.9801E-01
- Predikovaný korelační koeficient, Rp² : 9.9858E-01
- Střední kvadratická chyba predikce, MEP : 1.5006E-06
- Akaikeho informační kritérium, AIC : -1.6113E+02

(5) ANALÝZA KLASICKÝCH REZIDUÍ:

Bod	Meřená hodnota	Predikovaná hodnota	Směrodatná odchylnka	Klasické reziduum	Relativní reziduum
i	uexp[i]	uyup[i]	s(uyup[i])	e[i]	er[i]
4	2.5000E-02	2.5352E-02	4.1693E-04	-3.5198E-04	-1.4079E+00

Napověda-F1 Řádek: 64 - 86 Celkem: 184 Délka: 9793

Ve **Výsledky** táhněte dolů na bloky (3) a (4).

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

U Ý S L E D K Y

Bod	Meřená hodnota yexp[i]	Predikovaná hodnota yup[i]	Směrodatná odchylka s(yup[i])	Klasické reziduum e[i]	Relativní reziduum er[i]
1	1.2000E-02	1.1734E-02	5.5806E-04	2.6586E-04	2.2155E+00
2	1.6000E-02	1.4652E-02	5.2507E-04	1.3478E-03	8.4235E+00
3	1.9000E-02	1.8543E-02	4.8308E-04	4.5694E-04	2.4050E+00
4	2.5000E-02	2.5352E-02	4.1693E-04	-3.5198E-04	-1.4079E+00
5	3.4000E-02	3.5079E-02	3.4771E-04	-1.0790E-03	-3.1735E+00
6	3.7000E-02	3.8970E-02	3.3214E-04	-1.9698E-03	-5.3238E+00

Táhněte na blok (5) Analýza klasických reziduí

10	8.9000E-02	8.7170E-02	7.8510E-04	1.8210E-03	2.0405E+00
11	7.5000E-02	7.3987E-02	5.3594E-04	1.0129E-03	1.3505E+00
12	8.3000E-02	8.3714E-02	6.5063E-04	-7.1415E-04	-8.6042E-01

Rezidualní součet čtverců, RSC : 1.2671E-05
 Průměr absolutních hodnot reziduí, Me : 8.4987E-04
 Průměr relativních reziduí, Mer : 2.5044E+00
 Odhad reziduálního rozptylu, s²(e) : 1.2671E-06
 Odhad směrodatné odchylky reziduí, s(e) : 1.1257E-03
 Odhad šikmosti reziduí, g1(e) : 4.9856E-03
 Odhad špičatosti reziduí, g2(e) : 2.4115E+00

Napověda-F1 Řádek: 80 - 102 Celkem: 184 Délka: 9793

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-Snedcorův test významnosti regrese, F : 5.0058E+03
 Tabulkový kvantil, F(1-alpha, m-1, n-m) : 4.9646E+00
 Závěr: Navržený model je přijat jako významný.
 Spočtená hladina významnosti : 0.000

Scottovo kritérium multikolinearity, M : 5.5415E-15
 Závěr: Navržený model je korektní.

Cook-Weisbergův test heteroskedasticity, Sf : 2.2794E+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) : 1.7323E-01
 Tabulkový kvantil, Chi²(1-alpha, 2) : 5.9915E+00
 Závěr: Normalita je přijata.
 Spočtená hladina významnosti : 0.917

Waldův test autokorelace, Wa : 1.0793E+00
 Tabulkový kvantil, Chi²(1-alpha, 1) : 3.8415E+00
 Závěr: Rezidua nejsou autokorelována.
 Spočtená hladina významnosti : 0.299

Napověda-F1 Řádek: 105 - 127 Celkem: 184 Délka: 9793

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

U Ý S L E D K Y

(7) INDIKACE VLIVNÝCH BODŮ:
 (* indikuje odlehlý nebo vlivný bod)

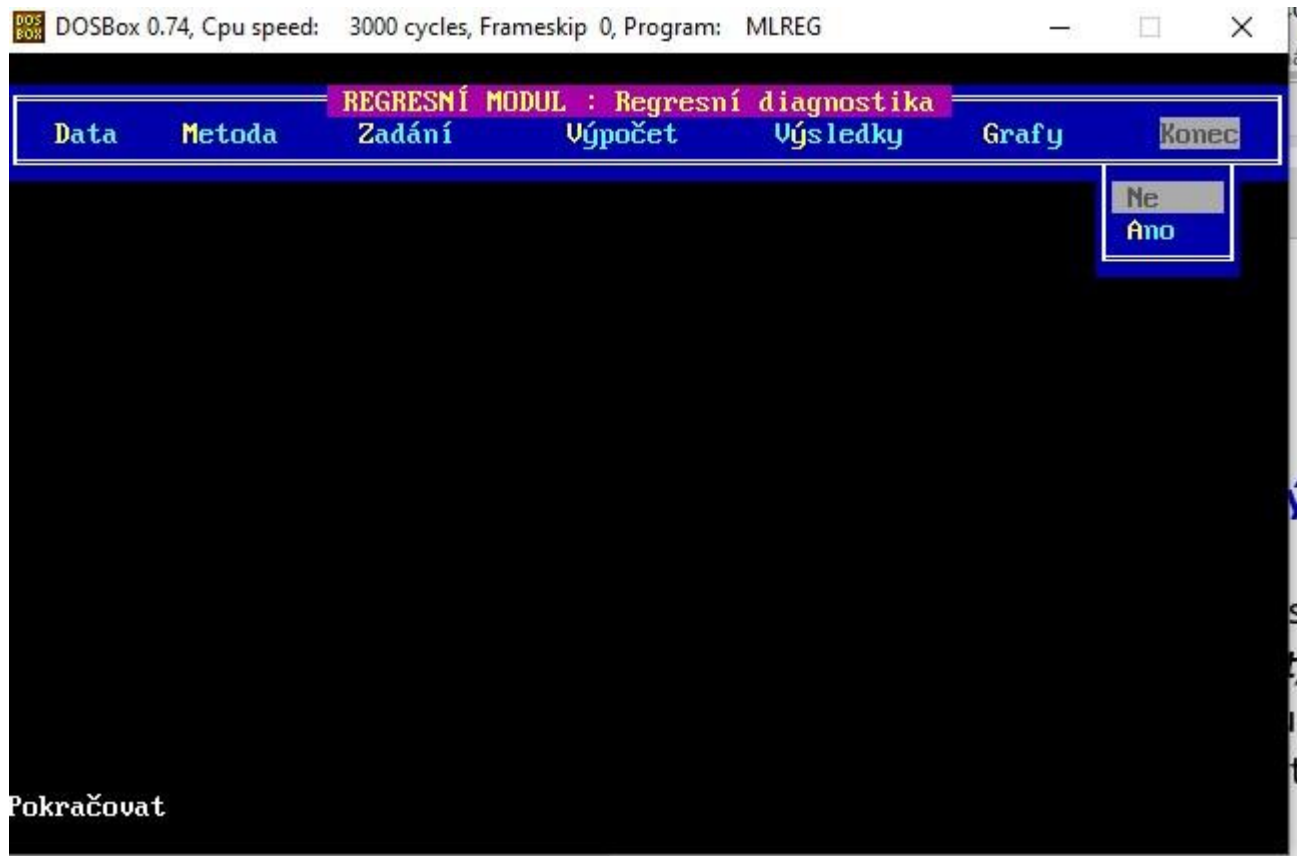
Bod	Standardizované reziduum eS[i]	Jackknife reziduum eJ[i]	Predikované reziduum eP[i]	Diagonální prvky H[i, i]
1	2.7195E-01	2.5896E-01	3.5249E-04	2.4577E-01
2	1.3536E+00	1.4208E+00	1.7225E-03	2.1758E-01
3	4.4942E-01	4.3073E-01	5.6009E-04	1.8417E-01
4	-3.3663E-01	-3.2118E-01	-4.0794E-04	1.3718E-01
5	-1.0078E+00	-1.0087E+00	-1.1928E-03	9.5416E-02
6	-1.8315E+00	-2.1313E+00	-2.1577E-03	8.7062E-02
7	1.7999E-01	1.7103E-01	2.1163E-04	8.3482E-02
8	-5.7359E-01	-5.5333E-01	-6.7792E-04	9.2880E-02
9	-3.5100E-01	-3.3505E-01	-4.2273E-04	1.2644E-01
10	1.7756E+00	2.0357E+00	2.1930E-03	1.6925E-01
11	1.0232E+00	1.0259E+00	1.3098E-03	2.2668E-01
12	-7.7744E-01	-7.6090E-01	-1.0724E-03	3.3408E-01*

Bod	Zobecněné diag. prvky Hm[i, i]	Cookova vzdálenost D[i]	Atkinsonova vzdálenost A[i]	Vliv na predikci DF[i]
1	2.5135E-01	1.2050E-02	3.3054E-01	1.4782E-01

Napověda-F1 Řádek: 136 - 158 Celkem: 184 Délka: 9793

Táhněte na blok (6) Testování regresního tripletu .

Táhněte na blok (7) Indikace vlivných bodů, kde jsou diagnostiky v číselné formě. Ukončete úlohu klikem na Esc.



Analýzu dat úlohy **V601** a návrat do červeného menu ADSTATu k analýze další regresní úlohy provedete kliknutím na písmena **K** a potom **A**.

Závěry a vysvětlení výstupu úlohy V601:

Úkoly

Odpovědi, závěry a vysvětlení

(1) Určete velikost systematické chyby metody (= velikost úseku β_0).

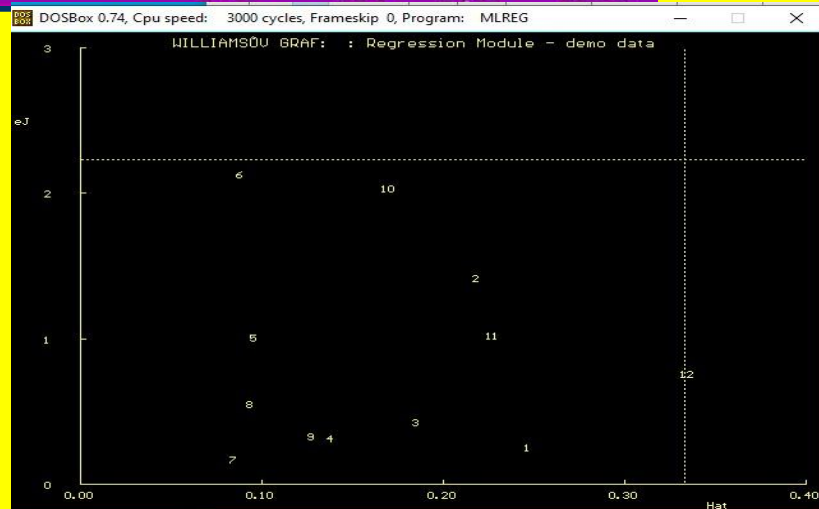
Parametr	Odhad	Směrodatná odchylka	Test $H_0: B[j] = 0$ vs. $H_A: B[j] \neq 0$	t-kriterium	hypoteza H_0 je	Hlad. úz
B[0]	1.0344E-03	6.8667E-04	1.5064E+00		Akceptována	0.163
B[1]	9.7270E-01	1.3748E-02	7.0752E+01		Zamítnuta	0.000

(2) Pokuste se vyjádřit také přesnost metody.

Reziduální součet čtverců, RSC	: 1.2671E-05
Průměr absolutních hodnot reziduí, Me	: 8.4987E-04
Průměr relativních reziduí, Mer	: 2.5044E+00
Odhad reziduálního rozptylu, $s^2(e)$: 1.2671E-06
Odhad směrodatné odchylky reziduí, $s(e)$: 1.1257E-03
Odhad šikmosti reziduí, $g1(e)$: 4.9856E-03
Odhad špičatosti reziduí, $g2(e)$: 2.4115E+00

(3) Jsou v datech vlivné a vybočující body graficky?

Williamsův graf nenašel žádné odlehlé body.



(4) Tabulkové indikace vlivných bodů.

Nejsou žádné odlehlé body.

Úloha V6.02 *Bichromátometrická metoda stanovení železitých iontů*

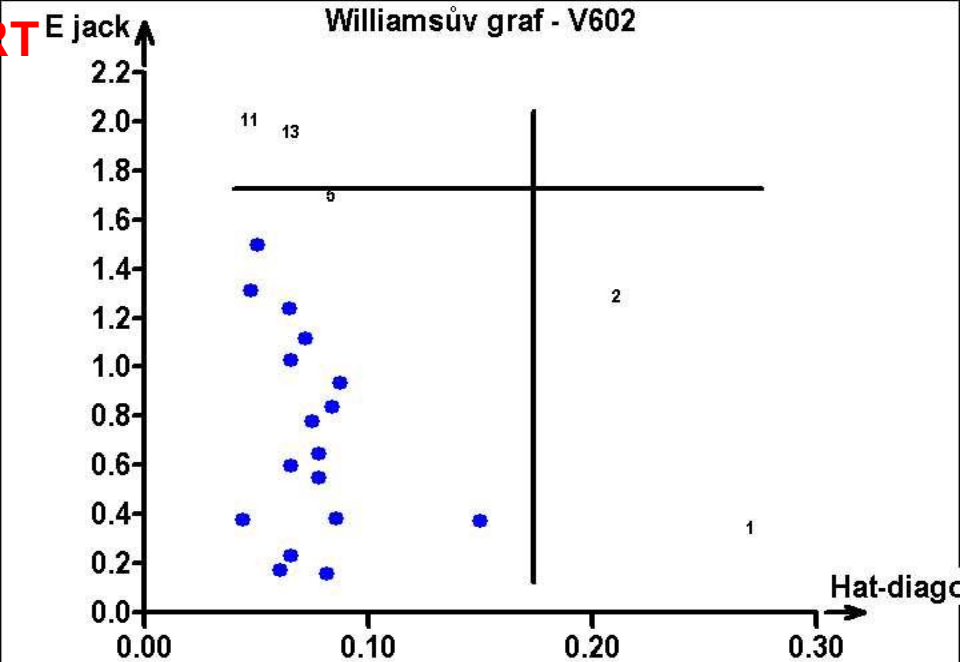
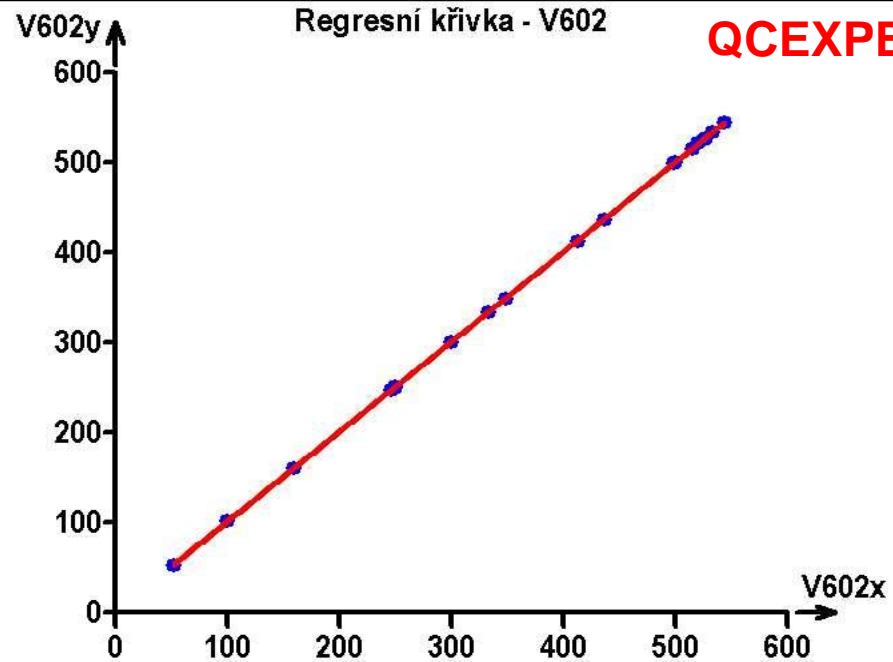
Zadání: Kraft a Dosch⁶⁰ navrhli titrační stanovení železa ve vodách. Železité ionty Fe^{3+} v Fe_2O_3 se redukuje titanitou solí v přebytku a vzniklé ionty Fe^{2+} se pak stanoví bichromátometricky.

Úkoly:

- (1) Vede titrační stanovení ke správným výsledkům?
- (2) Proved'te Studentův t -test významnosti úseku b_0 (má být $\beta_0 = 0$).
- (3) Proved'te Studentův t -test jednotkové směrnice b_1 (má být $\beta_1 = 1$).
- (4) Jsou v datech nějaké odlehlé body?

Data: Obsah Fe_2O_3 [mg], dáno x , nalezeno y :

Dáno x	Stanoveno y
52.0	52.50
...	...
543.61	543.78



Odhady parametrů

Proměnná	Odhad	Směr.Odch.	Závěr	Pravděpodobnost
Abs	0.70845	0.23872	Významný	0.007343
V602x	0.99834	0.00056	Významný	0

Spodní mez

0.21200
0.99716

Horní mez

1.20490
0.99951

Statistické charakteristiky regrese

Vícenásobný korelační koeficient R :	0.99999
Koeficient determinace R ² :	0.99999
Predikovaný korelační koeficient R _p :	0.99998
Střední kvadratická chyba predikce MEP :	0.17699
Akaikeho informační kritérium :	-39.545

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

Uložení F3=čtení F4=numerický mód F5=zoom ESC=konec
Počet sloupců: 2 Uodorovně U601

Numerický editor

1 2
1 1
2 2
3 3.500000 E-02 3.400000 E-02
4 3.900000 E-02 3.700000 E-02
5 4.500000 E-02 4.500000 E-02
6 5.200000 E-02 5.100000 E-02
7 6.100000 E-02 6.000000 E-02
8 6.800000 E-02 6.900000 E-02
9 7.500000 E-02 7.500000 E-02
10 8.500000 E-02 8.300000 E-02

Čtení ze souboru: U602

Otevřít Data, pak stisk F3, smazat demo.dat, napsat V602 a ukončit Esc a A.

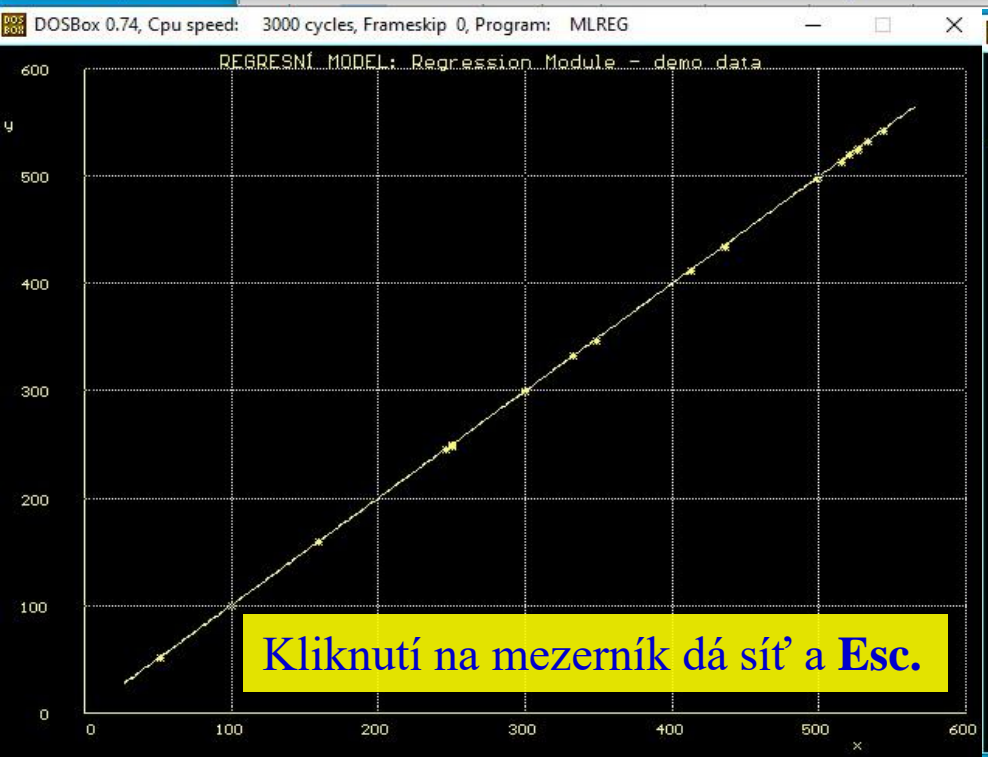
ostatní operace: F3=zmena velikosti matice F10=transpozice

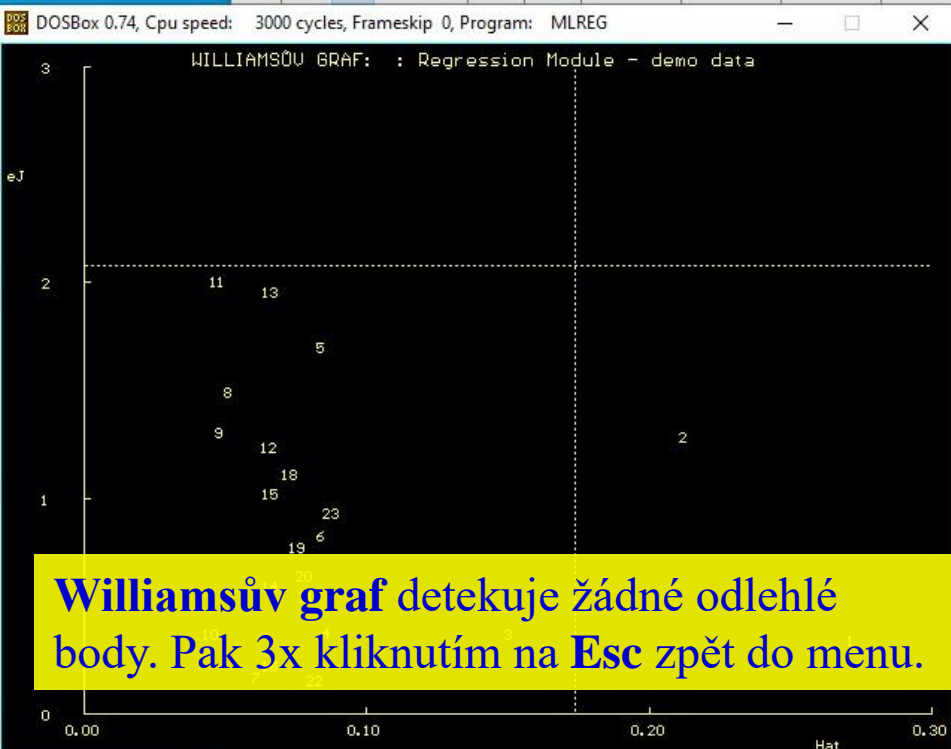
DOSBox 0.74, Cpu speed: 3000 cycles, Frameskip 0, Program: MLREG

F1=nápoověda F2=uložení F3=čtení F4=numerický mód F5=zoom ESC=konec
Řádek 1 Sloupec 5 Insert NUM U602

U602x U602y
52.00000 52.50000
100.00000 101.00000
160.00000 160.30000
246.55000 247.00000
250.00000 249.66000
250.00000 250.62000
300.00000 300.28000
333.07000 333.80000
348.25000 347.87000
412.75000 412.62000
436.36000 435.60000
499.07000 498.47000
500.00000 500.60000
500.00000 499.64000
500.00000 500.28000
500.00000 499.97000
500.00000 500.00000
526.26000 525.84000
526.82000 526.87000
533.17000 532.93000

Otevřít Data a F5 a ukončit Esc a A.





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: $BI_{jl} = 0$ vs. $H_A: BI_{jl} \neq 0$	t-kriterium	hypoteza H0 je	Hlad. výz.
BI 01	7.0845E-01	2.3872E-01	2.9677E+00	Zamítnuta	0.007	
BI 11	9.9834E-01	5.6702E-04	1.7607E+03	Zamítnuta	0.000	

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R	: 1.0000E+00
Koeficient determinace, R ²	: 9.9999E-01
Predikovaný korelační koeficient, R _p ²	: 1.0000E+00
Střední kvadratická chyba predikce, MEP	: 1.7699E-01
Akaikeho informační kritérium, AIC	: -3.9546E+01

(5) ANALÝZA KLASICKÝCH REZIDUÍ:

Bod	Meřená hodnota	Predikovaná hodnota	Směrodatná odchylka	Klasické reziduum	Relativní reziduum
i	y _{expli}	y _{vypli}	s(y _{vypli})	e _{li}	er _{li}
1	5.2500E+01	5.2622E+01	2.1141E-01	-1.2210E-01	-2.3257E-01
2	1.0100E+02	1.0054E+02	1.8679E-01	4.5761E-01	4.5308E-01

Ve Výsledky táhněte dolů na bloky (3) a (4).

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

V Ý S L E D K Y

10	4.1262E+02	4.1277E+02	8.5368E-02	-1.5304E-01	-3.7090E-02
11	4.3560E+02	4.3634E+02	8.8075E-01	-7.4381E-01	-1.7076E-01
12	4.9847E+02	4.9895E+02	1.0366E-01	-4.7970E-01	-9.6235E-02
13	5.0060E+02	4.9988E+02	1.0396E-01	7.2186E-01	1.4420E-01
14	4.9964E+02	4.9988E+02	1.0396E-01	-2.3814E-01	-4.7661E-02
15	5.0028E+02	4.9988E+02	1.0396E-01	4.0185E-01	8.0325E-02
16	4.9997E+02	4.9988E+02	1.0396E-01	8.1854E-02	1.6374E-02
17	5.2000E+02	5.2049E+02	1.1116E-01	3.0610E-01	5.8776E-02
20	5.2584E+02	5.2609E+02	1.1325E-01	-2.5453E-01	-4.8404E-02
21	5.2687E+02	5.2665E+02	1.1346E-01	2.1638E-01	4.1068E-02
22	5.3293E+02	5.3299E+02	1.1589E-01	-6.3058E-02	-1.1832E-02
23	5.4378E+02	5.4342E+02	1.2001E-01	3.6431E-01	6.6996E-02

Nyní na blok (5) Analýza klasických reziduí

Rezidualní součet čtverců, RSC	: 3.4632E+00
Průměr absolutních hodnot reziduí, Me	: 3.2816E-01
Průměr relativních reziduí, Mer	: 1.0809E-01
Odhad reziduálního rozptylu, s ² (e)	: 1.6492E-01
Odhad směrodatné odchylky reziduí, s(e)	: 4.0610E-01
Odhad šikmosti reziduí, g1(e)	: -1.2991E-01
Odhad špičatosti reziduí, g2(e)	: 2.1528E+00

Napověda-F1 Řádek: 100 - 122 Celkem: 236 Délka: 13444

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

V Ý S L E D K Y

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

Fisher-Snedecorův test významnosti regrese, F	: 3.1000E+06
Tabulkový kvantil, F(1-alpha, m-1, n-m)	: 4.3248E+00
Závěr: Navržený model je přijat jako významný.	
Spočtená hladina významnosti	: 0.000
Scottovo kritérium multikolinearity, M	: -1.3690E-12
Závěr: Navržený model je korektní.	
Cook-Weisbergův test heteroskedasticity, Sf	: 9.5731E+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)	: 7.5260E-01
Tabulkový kvantil, Chi ² (1-alpha, 2)	: 5.9915E+00
Závěr: Normalita je přijata.	
Spočtená hladina významnosti	: 0.686
Waldův test a ₁ = 0	: 2.4712E+00
Tabulkový kvantil, Chi ² (1-alpha, 1)	: 3.8415E+00
Závěr: Rezidua nevykazují korektnost.	
Spočtená hladina významnosti	: 0.116

Táhněte na blok (6) Testování regresního tripletu.

Napověda-F1 Řádek: 121 - 146 Celkem: 236 Délka: 13444

Závěry a vysvětlení výstupu úlohy V602:

Úkoly

Odpovědi, závěry a vysvětlení

(1) Vede titrační stanovení ke správným výsledkům?

Ne, přímka neprochází počátkem .

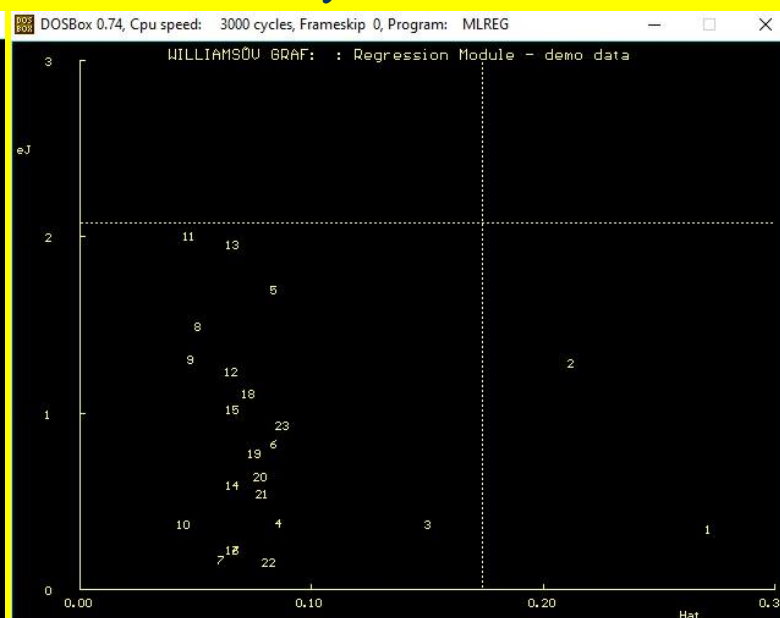
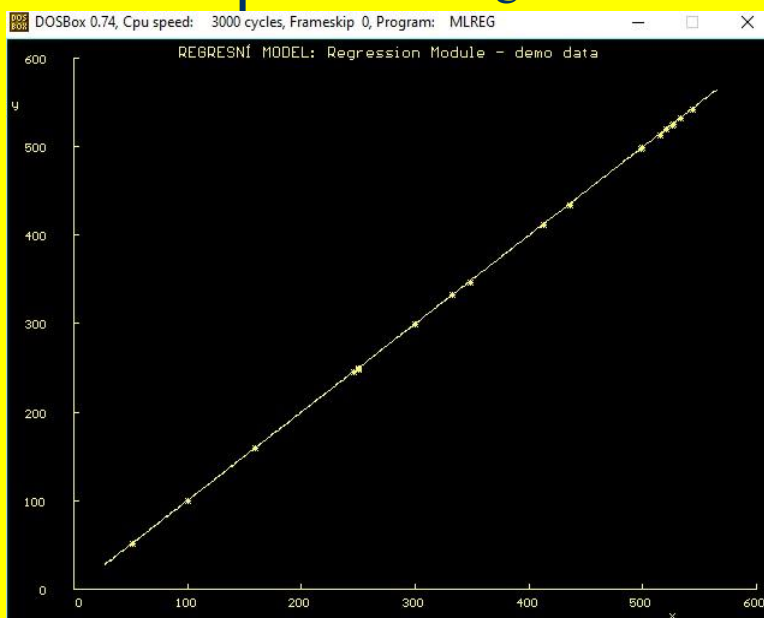
(2) Proved'te Studentův t -test významnosti úseku b_0 (má být $\beta_0 = 0$). Proved'te Studentův t -test jednotkové směrnice b_1 (má být $\beta_1 = 1$).

(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.
B[0]	7.0845E-01	2.3872E-01		2.9677E+00	Zamítnuta	0.007
B[1]	9.9834E-01	5.6702E-04		1.7607E+03	Zamítnuta	0.000

(3) Jsou v datech nějaké odlehlé body?.

Williamsův graf nenašel žádné odlehlé body.



Úloha V6.04 Stanovení kyseliny ftalové tenkovrstvou chromatografií

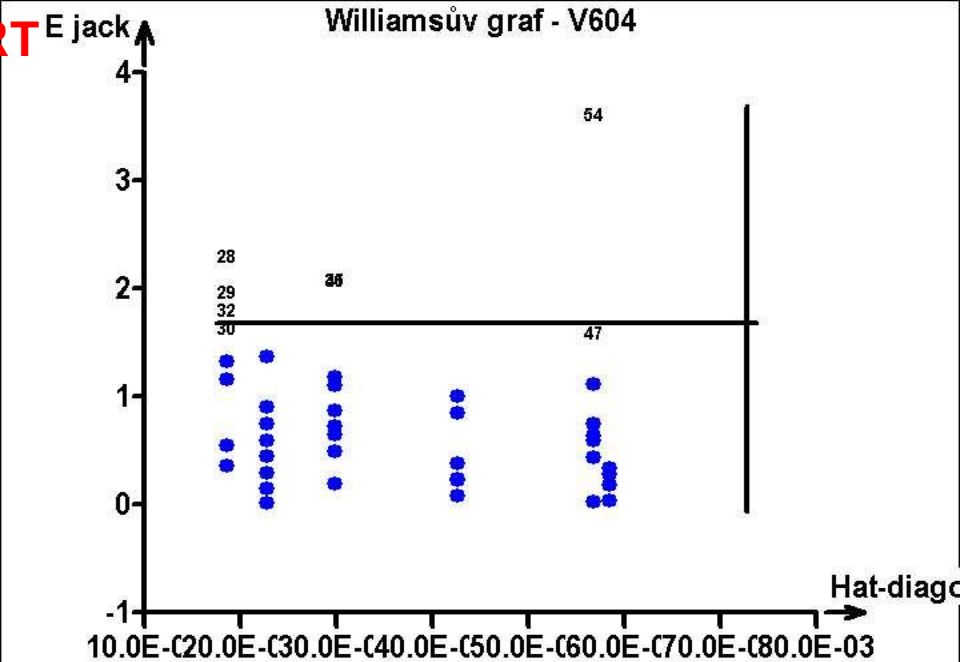
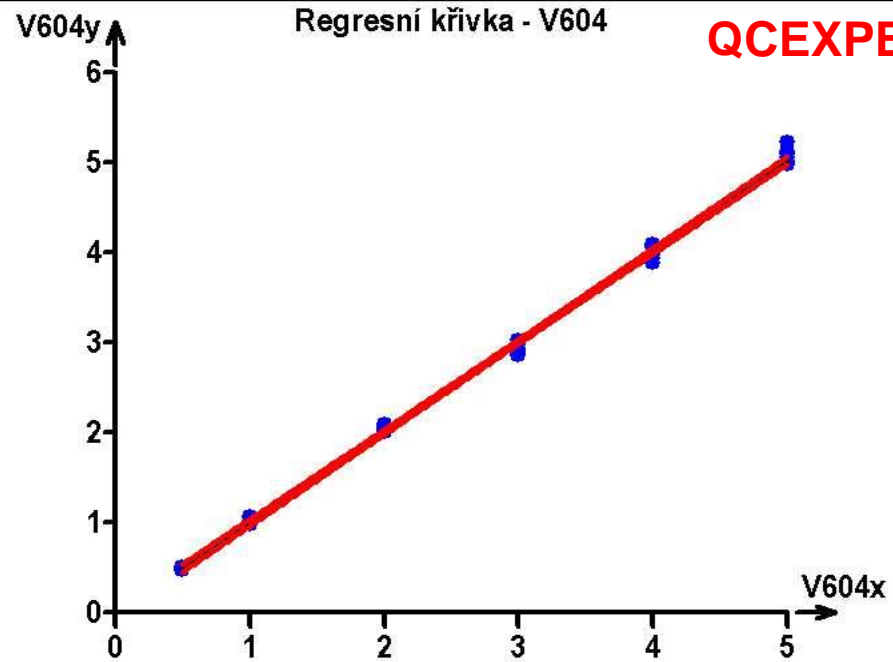
Zadání: Obsah kyseliny ftalové byl stanoven tenkovrstvou chromatografií a chromatogram byl vyhodnocován remisním fotometrem.

Úkoly:

- (1) Stanovte oba parametry lineárního regresního modelu a vyšetřete, zda je úsek nulový a směrnice jednotková.
- (2) Vyšetřete, zda jsou v datech vybočující hodnoty?
- (3) Je stanovení je správné?
- (4) Jaký je nutno zvolit postup při porušení předpokladů MNČ?

Data: Obsah kyseliny ftalové [μg], dáno x , nalezeno y (opakovaně).

Dáno x	Stanoveno y
0.50	0.48
...	...
5.23	65.02



Odhady parametrů

Proměnná	Odhad	Směr.Odch.	Závěr	Pravděpodobnost
Abs	-0.0110	0.0187	Nevýznamný	0.5565
V604x	1.00588	0.0059	Významný	0

Spodní mez
-0.048499
0.9940574

Horní mez
0.026400
1.017716

Statistické charakteristiky regrese

Vícenásobný korelační koeficient R : 0.9990902542
 Koeficient determinace R^2 : 0.9981813361
 Predikovaný korelační koeficient Rp : 0.9961043053
 Střední kvadratická chyba predikce MEP : 0.004614241291
 Akaikeho informační kritérium : -295.6513242

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

ození F3=čtení F4=numerický mód F5=zoom ESC=konec
 oupec 5 Insert NUM V604

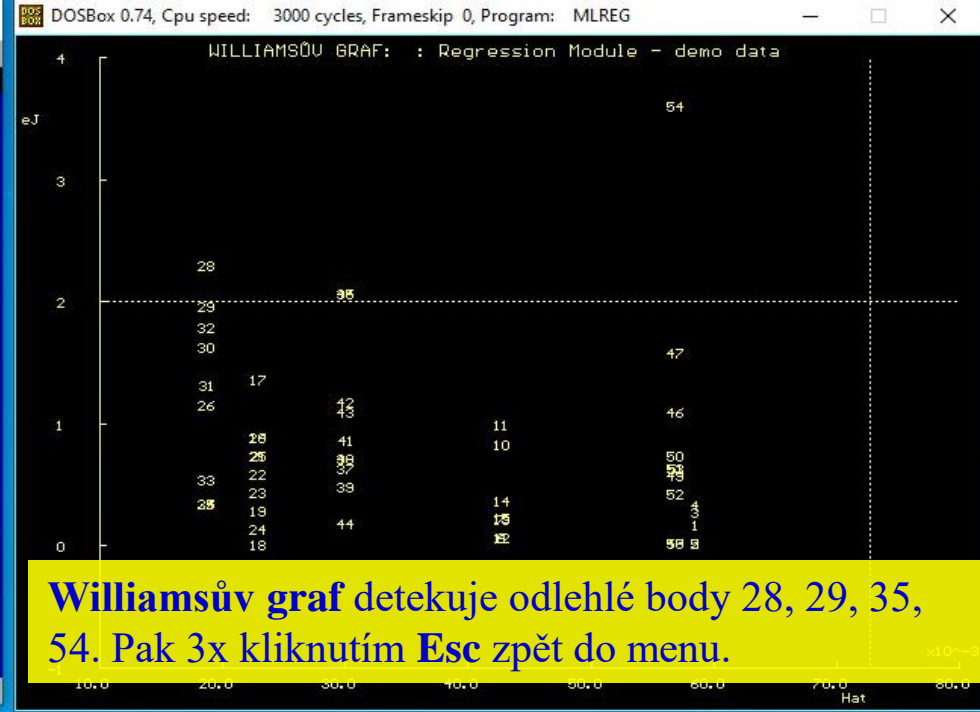
Textový editor

Čtení ze souboru: V604

0.50000	0.49000
0.50000	0.99000
0.50000	0.98000
1.00000	1.00000
1.00000	1.00000
1.00000	1.05000
1.00000	1.06000
1.00000	0.99000
1.00000	0.98000
1.00000	0.97000

Otevřít Data, pak stisk F3, smazat demo.dat, napsat V604, F4 a F5 a Esc a A a spustit výpočet.

Uhledávání chyb: F6=první F7=předchozí F8=následující F9=poslední



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 odchylna	Test H0: B _i j = 0 vs. H _A : B _i j <> 0	t-kriterium	hypoteza H0 je	Hlad. výz.
B _i 0	-1.1050E-02	1.8671E-02	-5.9179E-01		Ákceptována	0.557
B _i 1	1.0059E+00	5.8977E-03	1.7056E+02		Zamítnuta	0.000

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R : 9.9909E-01
 Koeficient determinace, R² : 9.9818E-01
 Predikovaný korelační koeficient, R_p² : 9.9902E-01
 Střední kvadratická chyba predikce, MEP : 4.6142E-03
 Akaikeho informační kritérium, AIC : -2.9565E+02

(5) ANALÝZA KLASICKÝCH REZIDUÍ:

Bod	Meřená hodnota	Predikovaná hodnota	Směrodatná odchylna	Klasické reziduum	Relativní reziduum
i	yexp[i]	yvyp[i]	s(yvyp[i])	e[i]	er[i]
1	4.8000E-01	4.9189E-01	1.6152E-02	-1.1894E-02	-2.4779E+00
2	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
3	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
4	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
5	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
6	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
7	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
8	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
9	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
10	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
11	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
12	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
13	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
14	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
15	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
16	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
17	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
18	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
19	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
20	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
21	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
22	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
23	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
24	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
25	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
26	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
27	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
28	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
29	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
30	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
31	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
32	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
33	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
34	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
35	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
36	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
37	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
38	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
39	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
40	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
41	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
42	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
43	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
44	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
45	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
46	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
47	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
48	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
49	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
50	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
51	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
52	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
53	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
54	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01
55	1.0000E-01	1.0188E-01	1.6152E-02	1.8885E-03	3.8613E-01

Napověda-F1 Řádek: 96 - 118 Celkem: 388 Délka: 24070

Sledujte Výsledek v bloku (3) a (4).

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

V Ý S L E D K Y

41	4.0700E+00	4.0125E+00	1.1558E-02	5.7503E-02	1.4128E+00
42	4.0900E+00	4.0125E+00	1.1558E-02	7.7503E-02	1.8949E+00
43	3.9400E+00	4.0125E+00	1.1558E-02	-7.2497E-02	-1.8400E+00
44	4.0000E+00	4.0125E+00	1.1558E-02	-1.2497E-02	-3.1243E-01
45	3.8800E+00	4.0125E+00	1.1558E-02	-1.3250E-01	-3.4149E+00
46	5.0900E+00	5.0184E+00	1.5930E-02	7.1616E-02	1.4070E+00
47	5.1200E+00	5.0184E+00	1.5930E-02	1.0162E-01	1.9847E+00
48	5.0200E+00	5.0184E+00	1.5930E-02	1.6159E-03	3.2189E-02
49	4.9800E+00	5.0184E+00	1.5930E-02	-3.8384E-02	-7.7076E-01
50	4.9700E+00	5.0184E+00	1.5930E-02	-4.8384E-02	-9.7353E-01
51	5.0000E+00	5.0184E+00	1.5930E-02	4.1616E-02	1.2245E-01
52	5.0000E+00	5.0184E+00	1.5930E-02	4.1616E-02	1.2245E-01
53	5.0600E+00	5.0184E+00	1.5930E-02	4.1616E-02	1.2245E-01
54	5.2300E+00	5.0184E+00	1.5930E-02	2.1162E-01	4.0462E+00
55	5.0200E+00	5.0184E+00	1.5930E-02	1.6159E-03	3.2189E-02

Proložení v bloku (5) Analýza klasických reziduí

Rezidualní součet čtverců, RSC	: 2.3672E-01
Průměr absolutních hodnot reziduí, Me	: 4.8687E-02
Průměr relativních reziduí, Mer	: 1.9264E-03
Odhad reziduálního rozptylu, s ² (e)	: 4.4664E+00
Odhad směrodatné odchylny reziduí, s(e)	: 6.6831E-02
Odhad šikmosti reziduí, g1(e)	: -4.1145E-02
Odhad špičatosti reziduí, g2(e)	: 4.0400E+00

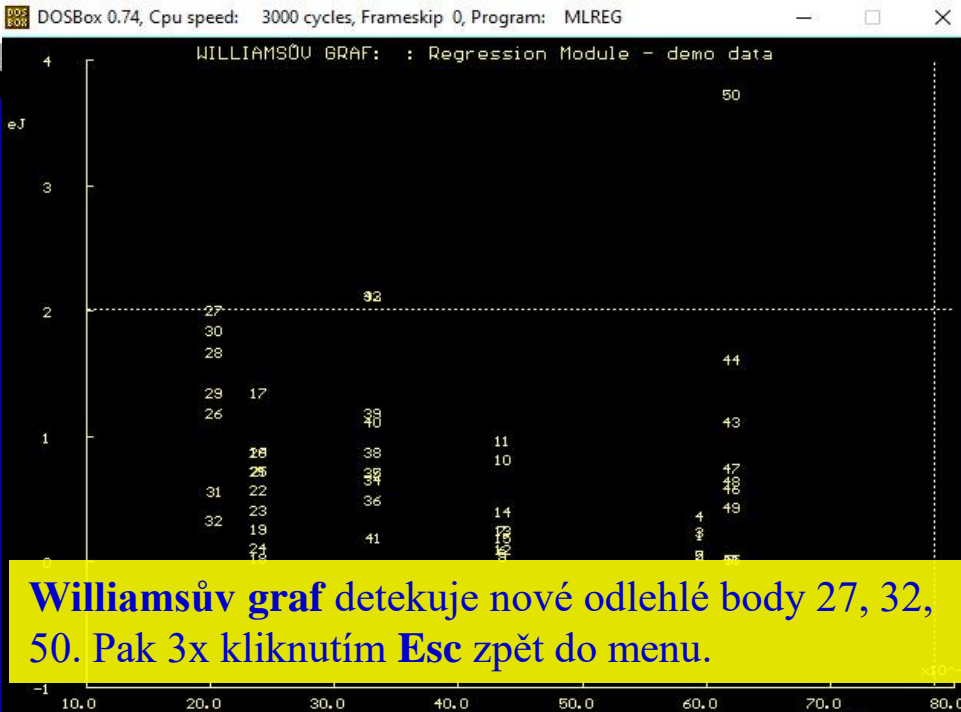
Napověda-F1 Řádek: 155 - 177 Celkem: 388 Délka: 24070

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

F1=nápoředa F2=uloženi F3=čteni F4=numerickej mód F5=zoom ESC=konec

Řádek	22	Sloupec	1	NUM	U604
1	0.00000	2	0.05000		
2	2.00000	2	0.04000		
3	2.00000	2	0.03000		
4	2.00000	2	0.01000		
5	2.00000	2	0.05000		
6	3.00000	2	0.93000		
7	3.00000	2	0.88000		
8	3.00000	2	0.90000		
9	3.00000	2	0.92000		
10	3.00000	2	0.89000		
11	3.00000	2	0.97000		
12	3.00000	3	0.03000		
13	4.00000	3	0.88000		
14	4.00000	3	0.97000		

Z dat byly vypuřtěny odlehlé body 28, 29, 35, 54 a opakován výpočet.



Williamsův graf detekuje nové odlehlé body 27, 32, 50. Pak 3x kliknutím Esc zpět do menu.

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

U V Y S L E D K Y

(3) ODHADY PARAMETRŮ A TESTY VÝZNAMNOSTI:

Parametr	Odhad	Směrodatná	Test H0: B j = 0 vs. HA: B j <> 0	t-kriterium	hypoteza H0 je	Hlad. výz.
B 0	-1.2051E-03	1.5728E-02	-7.6622E-02	akceptována		0.939
B 1	1.0025E+00	5.0269E-03	1.9943E+02	Zamítnuta		0.000

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R	: 9.9938E-01
Koeficient determinace, R ²	: 9.9877E-01
Predikovaný korelační koeficient, Rp ²	: 9.9934E-01
Střední kvadratická chyba predikce, MEP	: 3.2067E-03
Akaikeho informační kritérium, AIC	: -2.9245E+02

(5) ANALÝZA KLASICKÝCH REZIDUÍ:

Bod	Meřená hodnota	Predikovaná hodnota	Směrodatná odchylka	Klasické reziduum	Relativní reziduum
i	yexp[i]	yvypl[i]	s(yvypl[i])	e[i]	er[i]
1	4.8000E-01	5.0006E-01	1.3604E-02	-2.0064E-02	-4.1799E+00
2	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
3	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
4	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
5	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
6	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
7	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
8	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
9	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
10	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
11	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
12	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
13	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
14	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
15	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
16	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
17	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
18	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
19	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
20	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
21	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
22	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
23	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
24	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
25	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
26	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
27	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
28	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
29	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
30	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
31	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
32	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
33	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
34	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
35	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
36	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
37	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
38	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
39	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
40	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
41	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
42	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
43	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
44	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
45	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
46	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
47	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
48	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
49	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
50	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00
51	1.4000E-01	5.0006E-01	1.3604E-02	1.0000E-02	2.0538E+00

Napověda-F1 Řádek: 96 - 118 Celkem: 372 Délka: 22772

Sledujte Výsledky v bloku (3) a (4).

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

U V Y S L E D K Y

37	4.0600E+00	4.0089E+00	1.0134E-02	5.1055E-02	1.2575E+00
38	4.0700E+00	4.0089E+00	1.0134E-02	6.1055E-02	1.5001E+00
39	4.0900E+00	4.0089E+00	1.0134E-02	8.1055E-02	1.9818E+00
40	3.9400E+00	4.0089E+00	1.0134E-02	-6.8945E-02	-1.7499E+00
41	4.0000E+00	4.0089E+00	1.0134E-02	-8.9454E-03	-2.2363E-01
42	3.8800E+00	4.0089E+00	1.0134E-02	-1.2895E-01	-3.3233E+00
43	5.0900E+00	5.0115E+00	1.3888E-02	7.8517E-02	1.5426E+00
44	5.1200E+00	5.0115E+00	1.3888E-02	1.0852E-01	2.1195E+00
45	5.0200E+00	5.0115E+00	1.3888E-02	8.5170E-03	1.6966E-01
46	4.8800E+00	5.0115E+00	1.3888E-02	-2.1483E-02	-4.3053E-01
47	4.8800E+00	5.0115E+00	1.3888E-02	-2.1483E-02	-4.3053E-01
48	4.8800E+00	5.0115E+00	1.3888E-02	-2.1483E-02	-4.3053E-01
49	4.9900E+00	5.0115E+00	1.3888E-02	-2.1483E-02	-4.3053E-01
50	5.0600E+00	5.0115E+00	1.3888E-02	4.8517E-02	9.5883E-01
51	5.0200E+00	5.0115E+00	1.3888E-02	8.5170E-03	1.6966E-01

Rezidualní součet čtverců, RSC : 1.5246E-01
Průměr absolutních hodnot reziduí, Me : 4.2887E-02
Průměr relativních reziduí, Mer : 1.8905E+00
Odhad reziduálního rozptylu, s²(e) : 3.1114E-03
Odhad směrodatné odchylky reziduí, s(e) : 5.5780E-02
Odhad šikmosti reziduí, g1(e) : -5.0422E-01
Odhad špičatosti reziduí, g2(e) : 2.9647E+00

Napověda-F1 Řádek: 151 - 173 Celkem: 372 Délka: 22772

Proložení v bloku (5) Analýza klasických reziduí

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	50	Sloupec	5	Insert	NUM	U604
3.000000			2.920000			
3.000000			2.890000			
3.000000			2.970000			
4.000000			3.880000			
4.000000			3.970000			
4.000000			4.060000			
4.000000			3.980000			
4.000000			4.060000			
4.000000			4.070000			
4.000000			4.090000			
4.000000			3.940000			
4.000000			4.000000			
4.000000			3.880000			

Z dat byly vypuštěny odlehlé body 27, 32, 50 a opakován celý regresní výpočet.

5.000000			4.970000			
5.000000			5.060000			
5.000000			4.990000			
5.000000			5.020000			

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

V Ý S L E D K Y

34	3.9800E+00	4.0055E+00	1.0815E-02	-2.5467E-02	-6.3987E-01
35	4.0600E+00	4.0055E+00	1.0815E-02	5.4533E-02	1.3432E+00
36	4.0700E+00	4.0055E+00	1.0815E-02	6.4534E-02	1.5856E+00
37	4.0900E+00	4.0055E+00	1.0815E-02	8.4533E-02	2.0668E+00
38	3.9400E+00	4.0055E+00	1.0815E-02	-6.5467E-02	-1.6616E+00
39	4.0000E+00	4.0055E+00	1.0815E-02	-5.4667E-03	-1.3667E-01
40	3.8800E+00	4.0055E+00	1.0815E-02	-1.2547E-01	-3.2337E+00
41	5.0900E+00	5.0069E+00	1.4775E-02	8.3101E-02	1.6326E+00
42	5.1200E+00	5.0069E+00	1.4775E-02	1.1210E-01	2.2098E+00
43	5.1300E+00	5.0069E+00	1.4775E-02	1.1210E-01	2.2098E+00
44	5.1300E+00	5.0069E+00	1.4775E-02	1.1210E-01	2.2098E+00
45	5.1300E+00	5.0069E+00	1.4775E-02	1.1210E-01	2.2098E+00
46	5.1300E+00	5.0069E+00	1.4775E-02	1.1210E-01	2.2098E+00
47	4.9900E+00	5.0069E+00	1.4775E-02	-1.6899E-02	-3.3867E-01
48	5.0200E+00	5.0069E+00	1.4775E-02	1.3101E-02	2.6097E-01

Sledujte těsnost proložení v bloku (5) Analýza klasických reziduí

Rezidualní součet čtverců, RSC	: 1.4866E-01
Průměr absolutních hodnot reziduí, Me	: 4.3594E-02
Průměr relativních reziduí, Mer	: 1.9615E+00
Odhad reziduálního rozptylu, s ² (e)	: 3.2317E-03
Odhad směrodatné odchylky reziduí, s(e)	: 5.6848E-02
Odhad šikmosti reziduí, g1(e)	: -3.7160E-01
Odhad špičatosti reziduí, g2(e)	: 2.8329E+00

Napověda-F1 Řádek: 144 - 166 Celkem: 356 Délka: 21757

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[i,j] = 0 vs. HA: B[i,j] <> 0	t-kriterium	hypoteza H0 je	Hlad. úzj.
BI 01	-2.6357E-04	1.6163E-02	-1.6306E-02		Akceptována	0.987
BI 11	1.0014E+00	5.2426E-03	1.9102E+02		Zamítnuta	0.000

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R	: 9.9937E-01
Koeficient determinace, R ²	: 9.9874E-01
Predikovaný korelační koeficient, Rp ²	: 9.9932E-01
Střední kvadratická chyba predikce, MEP	: 3.3418E-03
Akaikeho informační kritérium, AIC	: -2.7331E+02

(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	4.8000E-01	5.0045E-01	1.3968E-02	-2.0453E-02	-4.2610E+00
2	4.9000E-01	5.0045E-01	1.3968E-02	-1.0453E-02	-2.1332E+00

Sledujte Výsledky v bloku (3) a (4).

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

V Ý S L E D K Y

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

Fisher-Snedocorův test významnosti regrese, F	: 3.6487E+04
Tabulkový kvantil, F(1-alpha, m-1, n-m)	: 4.0517E+00
Závěr: Navržený model je přijat jako významný.	
Spočtená hladina významnosti	: 0.000
Scottovo kritérium multikolinearity, M	: -1.5853E-14
Závěr: Navržený model je korektní.	
Cook-Weisbergův test heteroskedasticity, Sf	: 5.2900E+02
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)	: 1.1606E+00
Tabulkový kvantil, Chi ² (1-alpha, 2)	: 5.9915E+00
Závěr: Normalita je přijata.	
Spočtená hladina významnosti	: 0.560
Maldřův test autokorelace, Ma	: 5.7403E+00
Spočtená hladina významnosti	: 0.017

Sledujte regresní diagnostiky v bloku (6) Testování regresního tripletu.

Celkem: 356 Délka: 21757

Úloha V6.06 *Ověření stanovení železa spektrofotometrickou metodou*

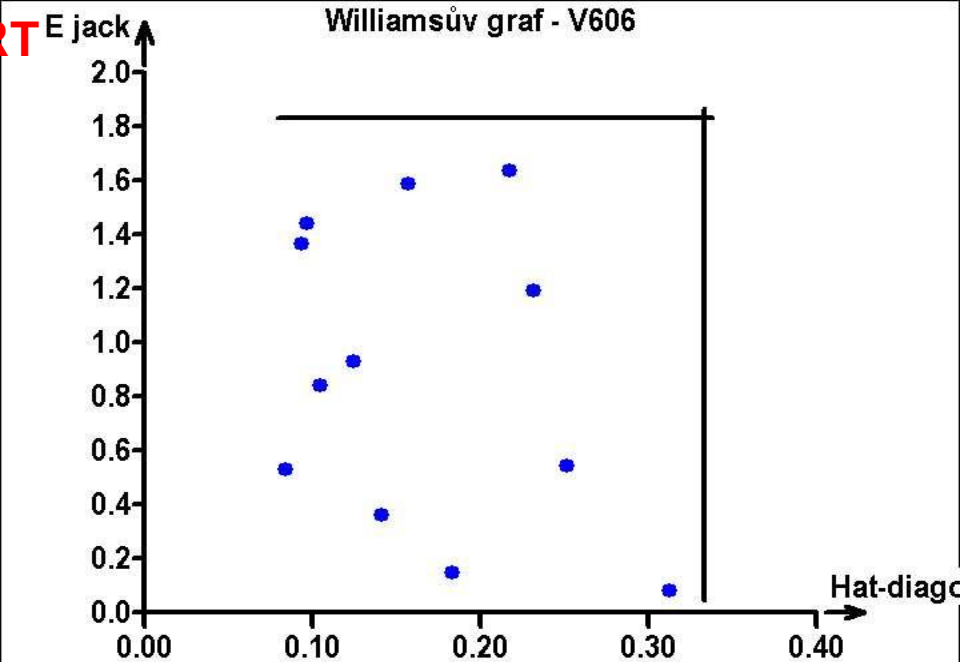
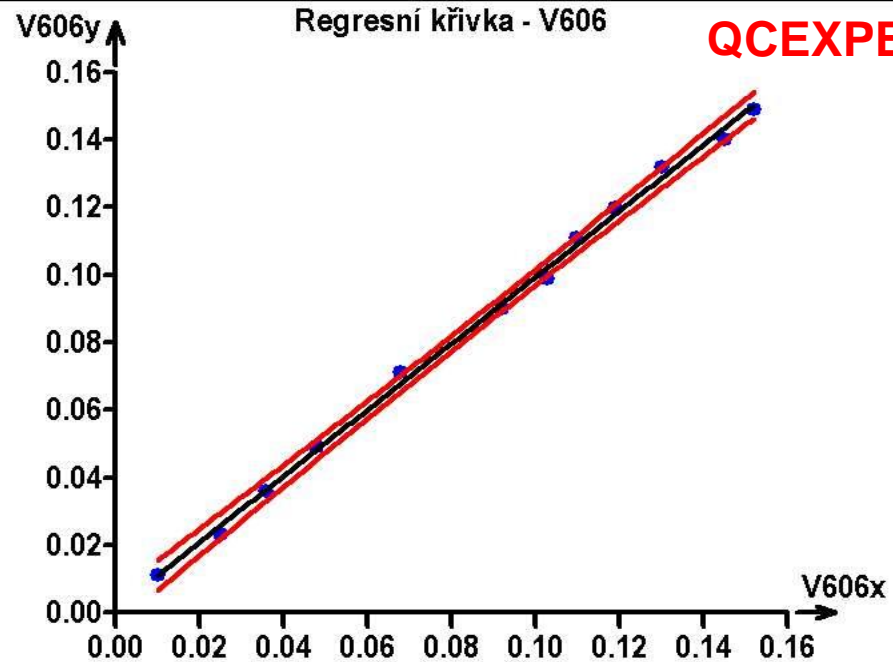
Zadání: Ověřte stanovení obsahu železa y v CoSO_4 spektrofotometricky SFM y porovnáním výsledků standardního stanovení obsahu x metodou AAS, u které je předpokládána zanedbatelná náhodná chyba.

Úkoly:

- (1) Vedou obě metody ke shodným výsledkům?
- (2) Jsou v datech odlehlé hodnoty? Užijte pět grafů indikace vlivných bodů.

Data: Obsah železa v CoSO_4 [%], když je AAS x [%], SFM y [%]:

Dáno x	Stanoveno y
0.010	0.011
...	...
0.152	0.149



Odhady parametrů

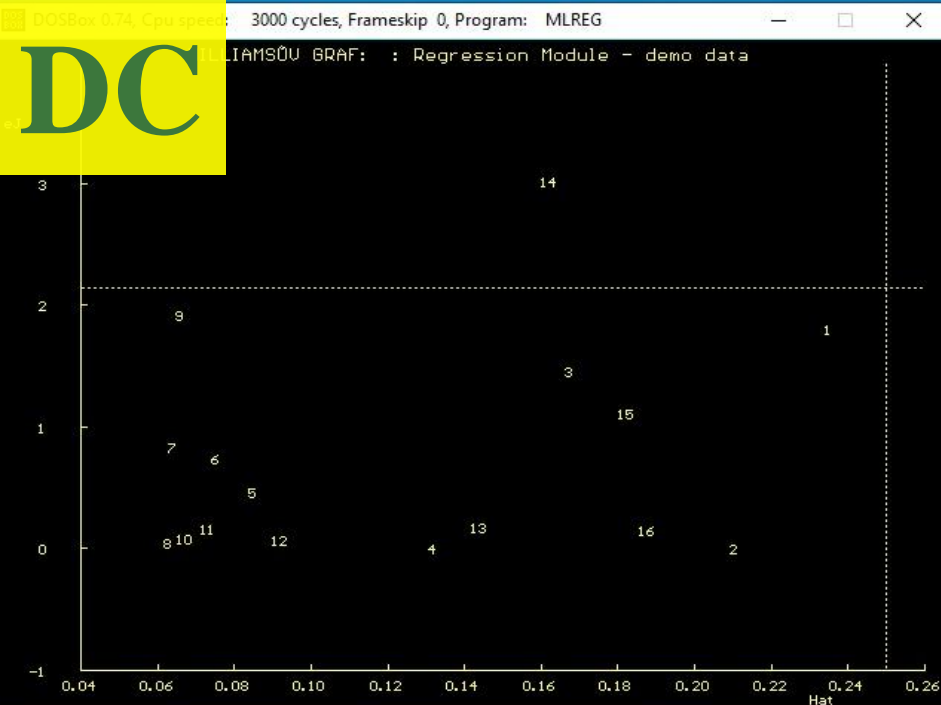
Proměnná	Odhad	Směr.Odch.	Závěr	Pravděpodobnost
Abs	0.001011	0.001531	Nevýznamný	0.5238
V606x	0.981567	0.015615	Významný	2.5313E-014

Spodní mez
-0.0023997
0.9467727

Horní mez
0.0044219
1.0163614

Statistické charakteristiky regrese

Vícenásobný korelační koeficient R : 0.9987368986
 Koeficient determinace R² : 0.9974753927
 Predikovaný korelační koeficient R_p : 0.9929318535
 Střední kvadratická chyba predikce MEP : 7.273899498E-006
 Akaikeho informační kritérium : -142.032264



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	2.0336E+00	1.1087E+00	1.8341E+00	1.8341E+00	Akceptována	0.088
BI 11	8.6790E-01	6.5378E-02	1.3275E+01	1.3275E+01	Zamítnuta	0.000

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R : 9.6250E-01
 Koeficient determinace, R² : 9.2640E-01
 Predikovaný korelační koeficient, Rp² : 9.4751E-01
 Střední kvadratická chyba predikce, MEP : 8.9916E-03
 Akaikeho informační kritérium, AIC : -7.6641E+01

(5) ANALÝZA KLASICKÝCH REZIDUÍ:

Bod	Meřená hodnota	Predikovaná hodnota	Směrodatná odchylka	Klasické reziduum	Relativní reziduum
i	yexp[i]	yyp[i]	s(yyp[i])	e[i]	er[i]
1	1.6150E+01	1.6276E+01	4.1652E-02	-1.2583E-01	-7.7911E-01
2	1.6310E+01	1.6311E+01	3.9435E-02	-5.4383E-04	-3.3343E-03
3	1.6490E+01	1.6380E+01	3.5167E-02	1.1002E-01	6.6722E-01
4	1.6450E+01	1.6449E+01	3.1192E-02	5.9359E-04	3.6085E-03

Napověda-F1 Řádek: 68 - 90 Celkem: 204 Délka: 11131

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

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

Řádek 14 Sloupec 5 Insert NUM U606.txt

1.641000 E+01	1.615000 E+01
1.645000 E+01	1.631000 E+01
1.653000 E+01	1.649000 E+01
1.661000 E+01	1.645000 E+01
1.676000 E+01	1.654000 E+01
1.681000 E+01	1.656000 E+01
1.691000 E+01	1.678000 E+01
1.694000 E+01	1.674000 E+01
1.703000 E+01	1.696000 E+01
1.704000 E+01	1.683000 E+01
1.709000 E+01	1.688000 E+01
1.718000 E+01	1.695000 E+01
1.733000 E+01	1.706000 E+01
1.741000 E+01	1.723000 E+01
1.742000 E+01	1.714000 E+01

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	0.0000E+00					
BI 11	9.8875E-01	1.0619E-03	9.3114E+02	9.3114E+02	Zamítnuta	0.000

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R : 9.7503E-01
 Koeficient determinace, R² : 9.5068E-01
 Predikovaný korelační koeficient, Rp² : 9.7132E-01
 Střední kvadratická chyba predikce, MEP : 5.1879E-03
 Akaikeho informační kritérium, AIC : -7.8972E+01

(5) ANALÝZA KLASICKÝCH REZIDUÍ:

Bod	Meřená hodnota	Predikovaná hodnota	Směrodatná odchylka	Klasické reziduum	Relativní reziduum
i	yexp[i]	yyp[i]	s(yyp[i])	e[i]	er[i]
1	1.6150E+01	1.6225E+01	1.7425E-02	-7.5419E-02	-4.6699E-01
2	1.6310E+01	1.6265E+01	1.7468E-02	4.5030E-02	2.7609E-01
3	1.6490E+01	1.6344E+01	1.7553E-02	1.4593E-01	8.8496E-01
4	1.6450E+01	1.6423E+01	1.7638E-02	2.6831E-02	1.6311E-01

Napověda-F1 Řádek: 64 - 86 Celkem: 196 Délka: 10768

Závěry a vysvětlení výstupu úlohy V606

Úkoly

Odpovědi, závěry a vysvětlení

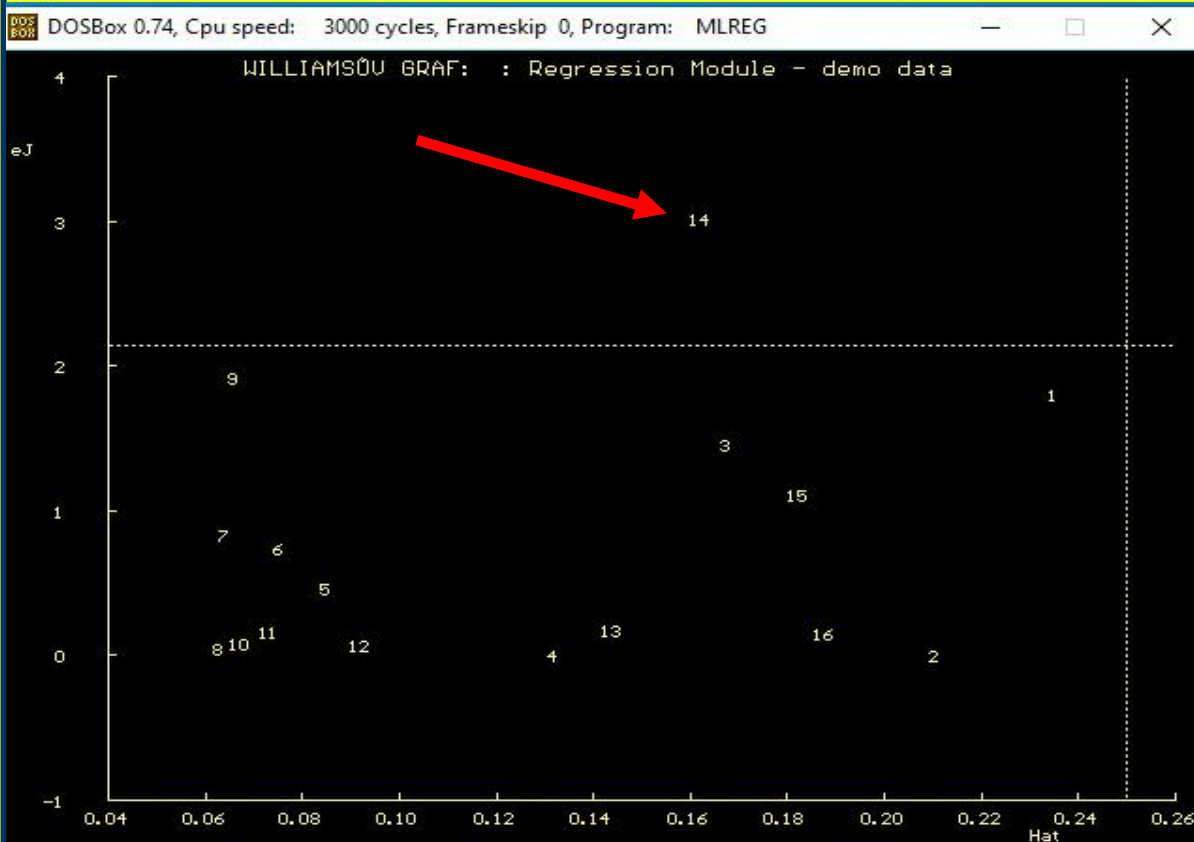
(1) Vedou obě metody ke shodným výsledkům?

Přímka prochází počátkem. Metoda mírně podhodnocuje.

(3) ODHADY PARAMETRŮ A TESTY VÝZNAMNOSTI:

Parametr	Odhad	Směrodatná odchylka	Test $H_0: \beta[j] = 0$ vs. $H_A: \beta[j] \neq 0$	t-kriterium	hypoteza H_0 je	Hlad. výz.
$\beta[0]$	0.0000E+00	-----	-----	-----	-----	-----
$\beta[1]$	9.8875E-01	1.0619E-03	9.3114E+02	Zamítnuta	0.000	

(2) Jsou v datech odlehlé hodnoty? Užijte více grafů indikace vlivných bodů.



Úloha V6.07 *Ověření stanovení dusičnanů v pitné a povrchové vodě*

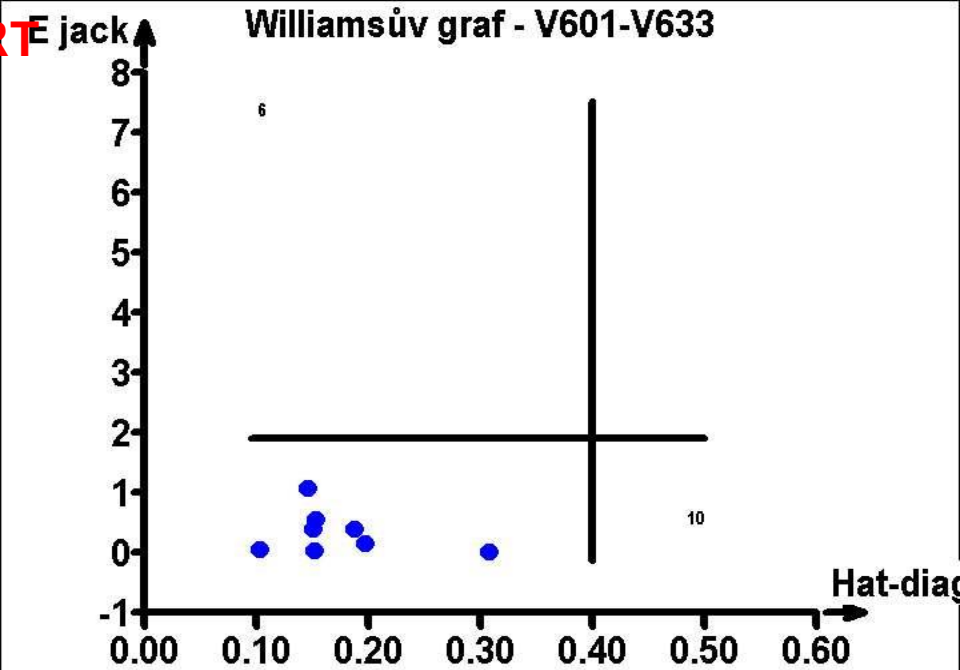
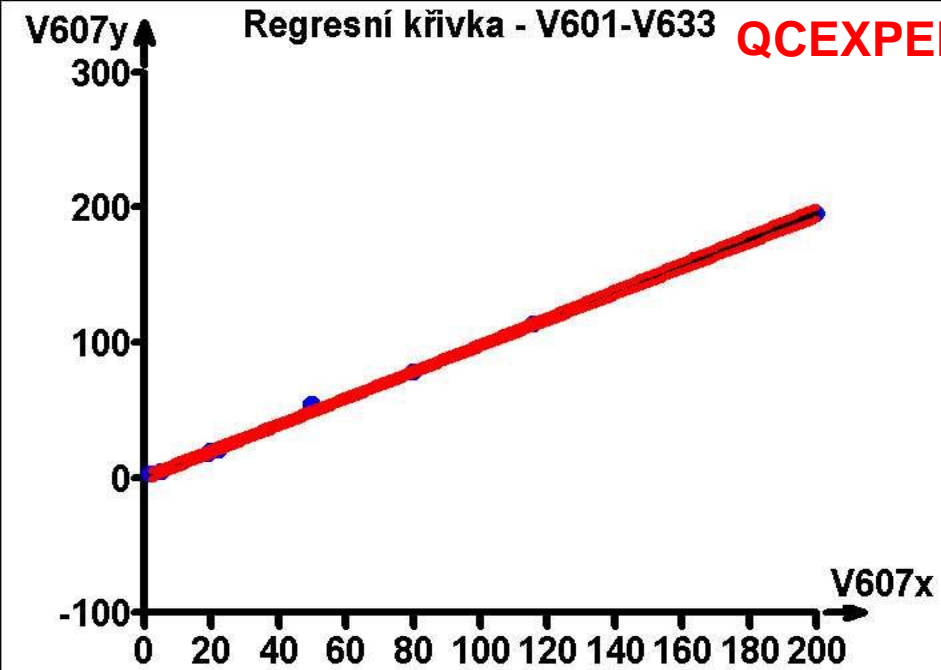
Zadání: V chemických laboratořích geochemické firmy se zavedla nová metoda stanovení obsahu dusičnanů y v pitných ale také povrchových vodách pomocí iontově párové chromatografie.

Úkoly:

- (1) Validujte novou metodu vůči deklarovaným obsahům NO_3^- [mg/l] x .
- (2) Odhadněte regresní parametry metodu ortogonální regrese.
- (3) Vede nová metoda ke správným výsledkům?
- (4) Proveďte simultánní test významnosti úseku a významnosti směrnice, zda je rovna jedné.

Data: Pro obsah dusičnanů NO_3^- [mg/l] je dáno x , nalezeno y .

Dáno x	Stanoveno y
2.10	2.20
...	...
200.00	195.00



Odhady parametrů

Proměnná	Odhad	Směr.Odch.	Závěr	Pravděpodobnost
Abs	-0.14284	0.9389	Nevýznamný	0.88284
V607x	0.9801479567	0.0098	Významný	1.179E-013

Spodní mez

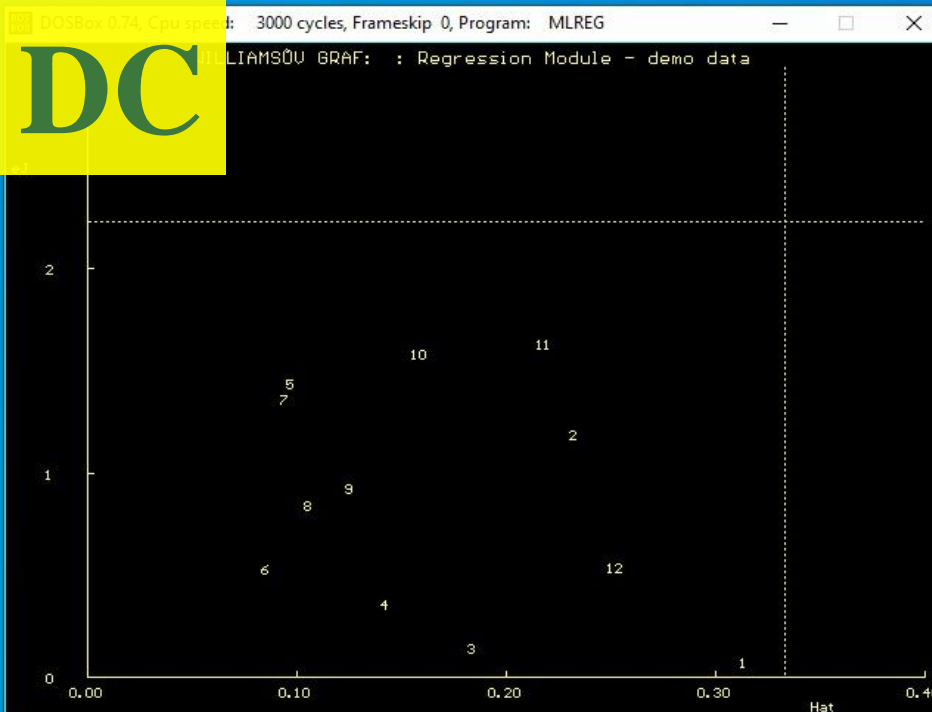
-2.3079
0.95739

Horní mez

2.0222
1.0029

Statistické charakteristiky regrese

Vícenásobný korelační koeficient R :	0.9995948631
Koeficient determinace R^2 :	0.9991898903
Predikovaný korelační koeficient Rp :	0.9978300448
Střední kvadratická chyba predikce MEP :	4.643796689
Akaikého informační kritérium :	16.4284433



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	1.0111E-03	1.5308E-03	6.6052E-01		Akceptována	0.524
BI 11	9.8157E-01	1.5616E-02	6.2857E+01		Zamítnuta	0.000

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R : 9.9874E-01
 Koeficient determinace, R^2 : 9.9748E-01
 Predikovaný korelační koeficient, Rp^2 : 9.9823E-01
 Střední kvadratická chyba predikce, MEP : 7.2739E-06
 Akaikeho informační kritérium, AIC :-1.4203E+02

(5) ANALÝZA KLASICKÝCH REZIDUÍ:

Bod	Meřená hodnota	Predikovaná hodnota	Směrodatná odchylka	Klasické reziduum	Relativní reziduum
i	yexpli	yvypil	s(yvypil)	e[i]	er[i]
1	1.1000E-02	1.0827E-02	1.3949E-03	1.7322E-04	1.5747E+00
2	2.3000E-02	2.5550E-02	1.2004E-03	-2.5503E-03	-1.1088E+01
3	3.6000E-02	3.6348E-02	1.0680E-03	-3.4753E-04	-9.6535E-01
4	4.9000E-02	4.8126E-02	9.3816E-04	8.7367E-04	1.7830E+00

Napověda-F1 Řádek: 64 - 86 Celkem: 184 Délka: 9793

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

ČEKÁM ...

REGRESNÍ MODUL : Regresní diagnostika

Data Metoda **Zadání** Výpočet Výsledky Grafy Konec

Podmínky
Volby

PODMÍNKY VÝPOČTU

Název: Regression Module - demo data

Počáteční podmínky:

Omezení	: 1.000000E-34	Absolutní člen	: Ne
Transformace	: Ne		
Váhy	: Ne		
Hladina význam.	: 0.050		

Zadání podmínek pro výpočet

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	0.0000E+00					
BI 11	9.9067E-01	7.1563E-03	1.3843E+02		Zamítnuta	0.000

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R : 9.9868E-01
 Koeficient determinace, R^2 : 9.9737E-01
 Predikovaný korelační koeficient, Rp^2 : 9.9832E-01
 Střední kvadratická chyba predikce, MEP : 6.8833E-06
 Akaikeho informační kritérium, AIC :-1.4363E+02

(5) ANALÝZA KLASICKÝCH REZIDUÍ:

Bod	Meřená hodnota	Predikovaná hodnota	Směrodatná odchylka	Klasické reziduum	Relativní reziduum
i	yexpli	yvypil	s(yvypil)	e[i]	er[i]
1	1.1000E-02	9.9067E-03	7.1563E-05	1.0933E-03	9.9392E+00
2	2.3000E-02	2.4767E-02	1.7891E-04	-1.7667E-03	-7.6814E+00
3	3.6000E-02	3.5664E-02	2.5763E-04	3.3592E-04	9.3312E-01
4	4.9000E-02	4.7552E-02	3.4350E-04	1.4479E-03	2.9549E+00

Napověda-F1 Řádek: 64 - 86 Celkem: 184 Délka: 9793

Závěry a vysvětlení výstupu úlohy V607

Úkoly

Odpovědi, závěry a vysvětlení

(1) Validujte novou metodu vůči deklarovaným obsahům NO_3^- [mg/l] x .

(3) ODHADY PARAMETRU 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. úz.
BI 01	1.0111E-03	1.5308E-03		6.6052E-01	Akceptována	0.524
BI 11	9.8157E-01	1.5616E-02		6.2857E+01	Zamítnuta	0.000

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R	: 9.9874E-01
Koeficient determinace, R^2	: 9.9748E-01
Predikovaný korelační koeficient, R_p^2	: 9.9823E-01
Střední kvadratická chyba predikce, MEP	: 7.2739E-06
Akaikeho informační kritérium, AIC	: -1.4203E+02

(2) Odhadněte regresní parametry metodu ortogonální regrese.

(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. úz.
BI 01	0.0000E+00	-----		-----	-----	-----
BI 11	9.9067E-01	7.1563E-03		1.3843E+02	Zamítnuta	0.000

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R	: 9.9868E-01
Koeficient determinace, R^2	: 9.9737E-01
Predikovaný korelační koeficient, R_p^2	: 9.9832E-01
Střední kvadratická chyba predikce, MEP	: 6.8833E-06
Akaikeho informační kritérium, AIC	: -1.4363E+02

(3) Vede nová metoda ke správným výsledkům?

Ano, přímka prochází počátkem a R je blízké jedné. Směrnice je mírně pod 1, čili nová metoda slabě podhodnocuje. MEP a AIC rozhoduje o nejlepším modelu a o b_0 rovném nule.

(4) Proved'te simultánní test významnosti úseku a významnosti směrnice, zda je rovna jedné.

V Adstatu nelze tento test.

Úloha V6.20 *Validace nové metody stanovení arsenu v odpadní vodě*

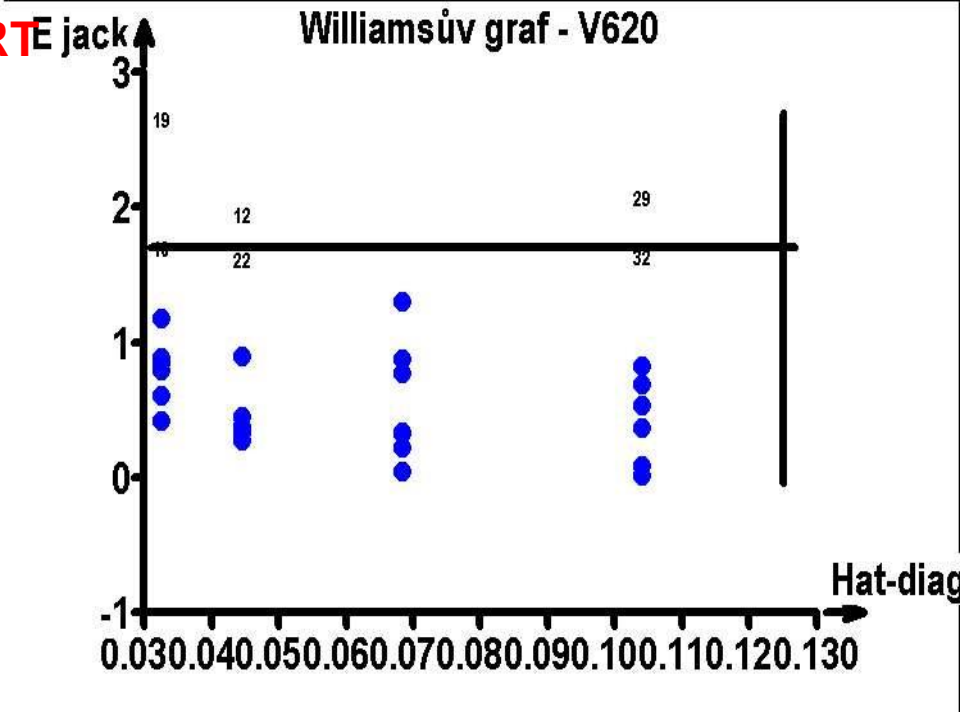
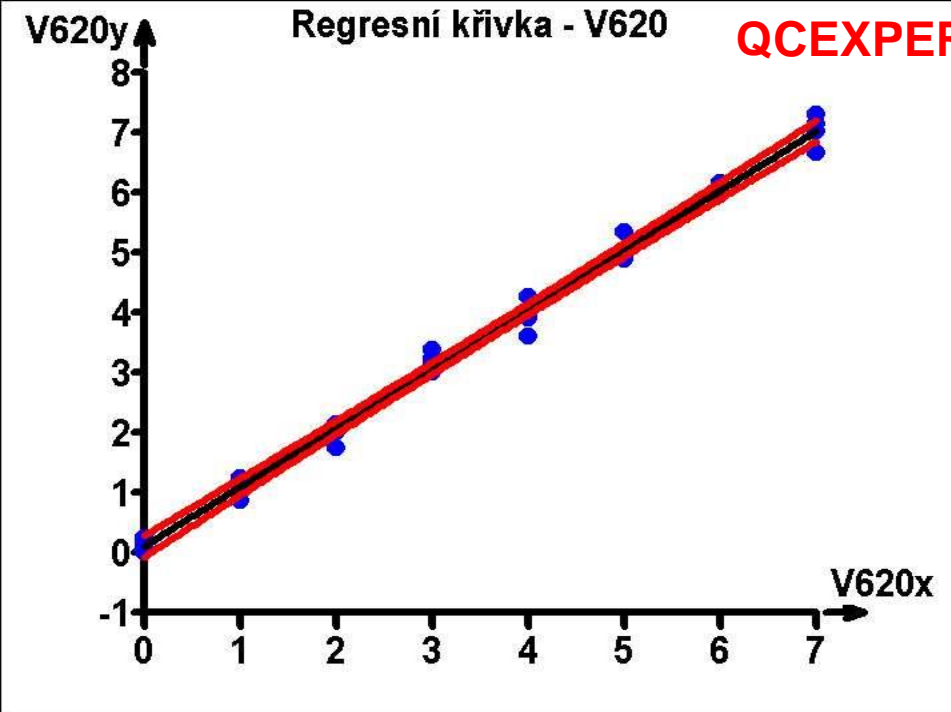
Zadání: Je třeba validovat nové jednodušší stanovení arsenu v odpadní vodě. Mezi naměřenou koncentrací arsenu y a známou koncentrací x v $\mu\text{g/ml}$ je předpokládán lineární regresní model $y = \beta_0 + \beta_1 x$.

Úkoly:

- (1) Užitím ortogonální regrese ověřte správnost nové metody.
- (2) K jakému výsledku dospěje nová metoda, když standard arsen vůbec neobsahuje čili absolutní člen je nulový, $\beta_0 = 0$?
- (3) Vyšetřete, zda nová metoda nadhodnocuje či podhodnocuje?
- (4) Jakou modifikaci MNČ je třeba použít, když jsou všechny proměnné zatíženy náhodnými chybami?

Data: Koncentrace arsenu daná x [$\mu\text{g. cm}^{-3}$], nalezená y [$\mu\text{g. cm}^{-3}$].

Dáno x	Stanoveno y
0	0.17
...	...
7.0	7.30



Odhady parametrů

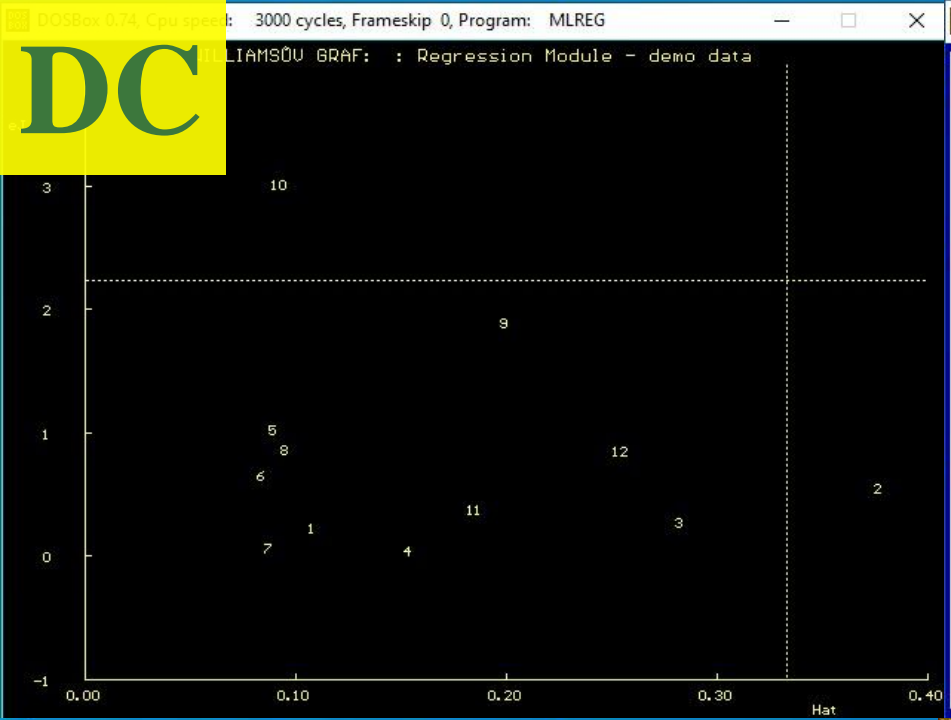
Proměnná	Odhad	Směr.Odch.	Závěr	Pravděpodobnost
Abs	0.10458	0.06051	Nevýznamný	0.0942
V620x	0.98770	0.01446	Významný	0

Spodní mez
-0.01899
0.958168

Horní mez
0.228167
1.017248

Statistické charakteristiky regrese

Vícenásobný korelační koeficient R : 0.99679
 Koeficient determinace R² : 0.99360
 Predikovaný korelační koeficient R_p : 0.98563
 Střední kvadratická chyba predikce MEP : 0.03715
 Akaikeho informační kritérium : -105.20



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. $H_A: B[j] <> 0$	t-kriterium	hypoteza H0 je	Hlad. úz.
BI 01	-3.5864E+00	6.0594E+00		-5.9188E-01	Akceptována	0.567
BI 11	1.0482E+00	8.0404E-02		1.3037E+01	Zamítuta	0.000

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R : 9.7182E-01
 Koeficient determinace, R^2 : 9.4443E-01
 Predikovaný korelační koeficient, R_p^2 : 9.6145E-01
 Střední kvadratická chyba predikce, MEP : 2.5582E-02
 Akaikeho informační kritérium, AIC :-4.3685E+01

(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	7.5072E+01	7.5108E+01	4.9107E-02	-3.5590E-02	-4.7408E-02
2	7.4418E+01	7.4349E+01	9.2107E-02	6.9330E-02	9.3163E-02
3	7.4497E+01	7.4535E+01	7.9765E-02	-3.8254E-02	-5.1350E-02
4	7.4898E+01	7.4891E+01	5.8785E-02	7.3989E-03	9.8787E-03

Napověda-F1 Řádek: 64 - 86 Celkem: 184 Délka: 9793

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

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

Řádek	Sloupec	NUM	U620
7.507300 E+01	7.507200 E+01		
7.434900 E+01	7.441800 E+01		
7.452700 E+01	7.449700 E+01		
7.486600 E+01	7.489800 E+01		
7.549900 E+01	7.540700 E+01		
7.536600 E+01	7.531600 E+01		
7.546900 E+01	7.551100 E+01		
7.555800 E+01	7.549000 E+01		
7.599300 E+01	7.629900 E+01		
7.595400 E+01	7.597700 E+01		
7.613100 E+01	7.610300 E+01		

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. $H_A: B[j] <> 0$	t-kriterium	hypoteza H0 je	Hlad. úz.
BI 01	0.0000E+00					
BI 11	1.0002E+00	4.2949E-04		2.3289E+03	Zamítuta	0.000

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R : 9.8467E-01
 Koeficient determinace, R^2 : 9.6958E-01
 Predikovaný korelační koeficient, R_p^2 : 9.8138E-01
 Střední kvadratická chyba predikce, MEP : 1.2703E-02
 Akaikeho informační kritérium, AIC :-4.8149E+01

(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	7.5072E+01	7.5092E+01	3.2243E-02	-1.9516E-02	-2.5996E-02
2	7.4418E+01	7.4367E+01	3.1932E-02	5.0662E-02	6.8078E-02
3	7.4497E+01	7.4545E+01	3.2009E-02	-4.8381E-02	-6.4943E-02
4	7.4898E+01	7.4884E+01	3.2154E-02	1.3540E-02	1.8078E-02

Napověda-F1 Řádek: 64 - 86 Celkem: 180 Délka: 9470

Závěry a vysvětlení výstupu úlohy V620

Úkoly

Odpovědi, závěry a vysvětlení

(1) Užitím ortogonální regrese ověřte správnost nové metody.

(3) ODHADY PARAMETRŮ A TESTY VÝZNAMNOSTI:

Parametr	Odhad	Směrodatná odchylka	Test H ₀ : B _l j = 0 vs. H _A : B _l j <> 0	t-kriterium	hypoteza H ₀ je	Hlad. výz.
BI 01	-3.5864E+00	6.0594E+00		-5.9188E-01	Ákceptována	0.567
BI 11	1.0482E+00	8.0404E-02		1.3037E+01	Zamítnuta	0.000

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R	: 9.7182E-01
Koeficient determinace, R ²	: 9.4443E-01
Predikovaný korelační koeficient, R _p ²	: 9.6145E-01
Střední kvadratická chyba predikce, MEP	: 2.5582E-02
Ákaikeho informační kritérium, AIC	: -4.3685E+01

(2) K jakému výsledku dospěje nová metoda, když standard arsen vůbec neobsahuje čili absolutní člen je nulový, $\beta_0 = 0$?

(3) ODHADY PARAMETRŮ A TESTY VÝZNAMNOSTI:

Parametr	Odhad	Směrodatná odchylka	Test H ₀ : B _l j = 0 vs. H _A : B _l j <> 0	t-kriterium	hypoteza H ₀ je	Hlad. výz.
BI 01	0.0000E+00					
BI 11	1.0002E+00	4.2949E-04		2.3289E+03	Zamítnuta	0.000

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R	: 9.8467E-01
Koeficient determinace, R ²	: 9.6958E-01
Predikovaný korelační koeficient, R _p ²	: 9.8138E-01
Střední kvadratická chyba predikce, MEP	: 1.2703E-02
Ákaikeho informační kritérium, AIC	: -4.8149E+01

(3) Vyšetřete, zda nová metoda nad- či pod-hodnocuje?

Protože je směrnice rovna jedné ani nepod- ani nenad-hodnocuje, ale je správná. MEP a AIC rozhoduje o nejlepším modelu a o b_0 rovném nule.

(4) Jakou modifikaci MNČ je třeba použít, když jsou všechny proměnné zatíženy náhodnými chybami?

Orthogonální regresi.

Úloha V6.22 *Validace navržené titrační metody ke stanovení modré báze MB H-3R*

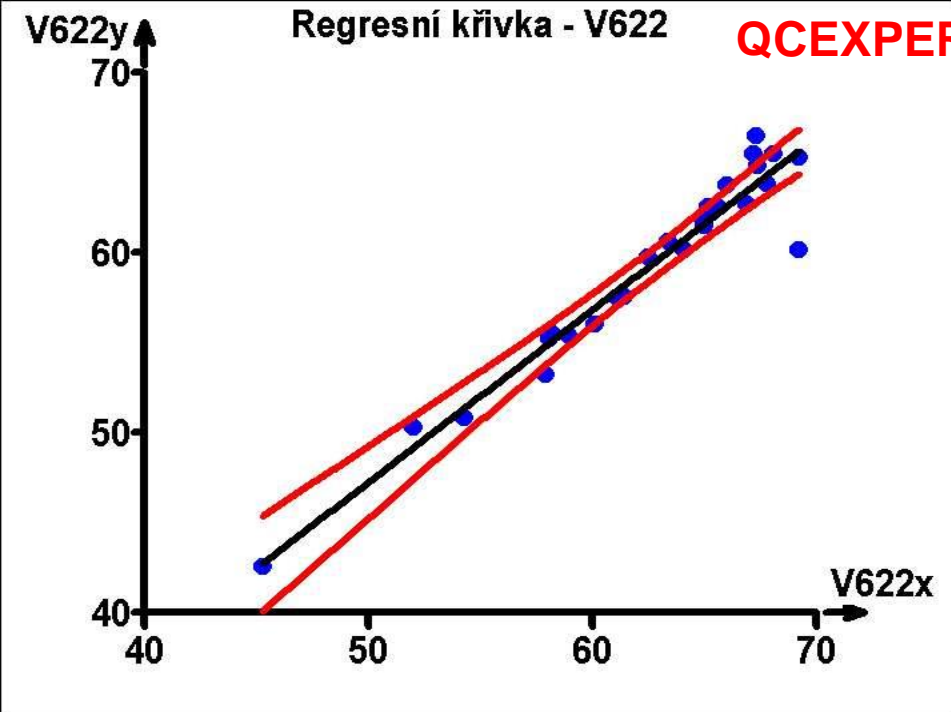
Zadání: Při výrobě modré báze MB H-3R byl stanovován její obsah v pastě z kalolisu titračně dusitanem v kyselém prostředí y a standardně spektrofotometricky x . Za základ byla vzata titrační metoda. Rozptyl této metody se považuje za zanedbatelný vůči rozptylu spektrofotometrické metody.

Úkoly:

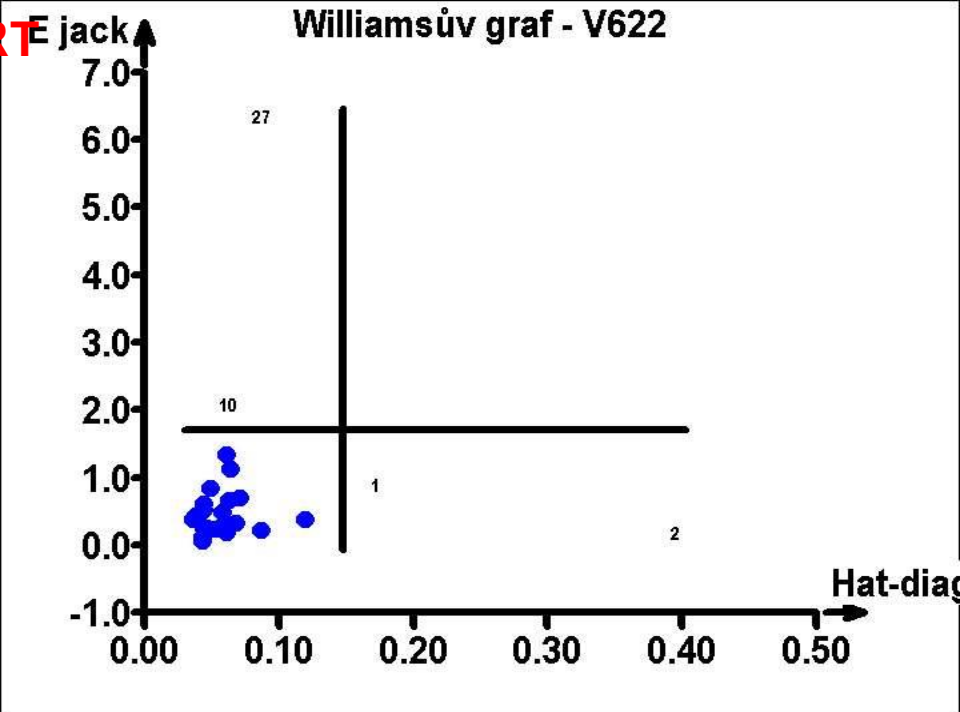
- (1) Popište test významnosti absolutního členu.
- (2) Vysvětlete test shodnosti odhadu parametru β s předepsanou $\beta_0 = 0$.

Data: Koncentrace modré báze spektrofotometrickou metodou x a titrační metodou y .

Dáno x	Stanoveno y
52.0	50.3
...	...
69.2	60.1



QCEXPERT



Odhady parametrů

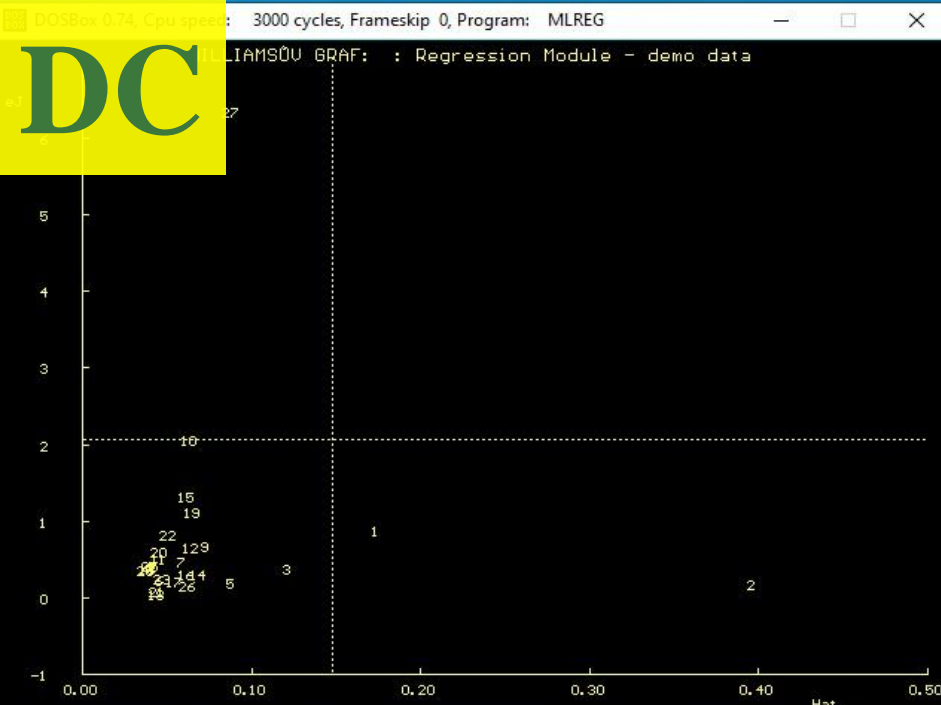
Proměnná	Odhad	Směr.Odch.	Závěr	Pravděpodobnost
Abs	-0.65650	3.14318	Nevýznamný	0.83624
V622x	0.957240	0.04994	Významný	2.220E-016

Spodní mez
-7.1300
0.85438

Horní mez
5.8170
1.0600

Statistické charakteristiky regrese

Vícenásobný korelační koeficient R :	0.967619281
Koeficient determinace R^2 :	0.936287073
Predikovaný korelační koeficient Rp :	0.8551935541
Střední kvadratická chyba predikce MEP :	2.300430311
Akaikého informační kritérium :	22.00604283



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	-6.5651E-01	3.1432E+00	-2.0887E-01		Akceptována	0.836
BI 11	9.5724E-01	4.9941E-02	1.9167E+01		Zamítnuta	0.000

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R	: 9.6762E-01
Koeficient determinace, R ²	: 9.3629E-01
Predikovaný korelační koeficient, Rp ²	: 9.6165E-01
Střední kvadratická chyba predikce, MEP	: 2.3004E+00
Akaikeho informační kritérium, AIC	: 2.2006E+01

(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	5.0300E+01	4.9120E+01	6.0240E-01	1.1800E+00	2.3459E+00
2	4.2500E+01	4.2706E+01	9.1219E-01	-2.0650E-01	-4.8587E-01
3	5.0800E+01	5.1322E+01	5.0344E-01	-5.2167E-01	-1.0269E+00
4	5.7500E+01	5.8118E+01	2.8648E-01	-6.1808E-01	-1.0749E+00

Napověda-F1 Řádek: 76 - 98 Celkem: 256 Délka: 14784

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

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

Řádek 27	Sloupec 5	Insert	NUM	U622
6.920000 E+01	6.530001 E+01			
6.530001 E+01	6.220000 E+01			
6.690000 E+01	6.270000 E+01			
6.120000 E+01	5.740000 E+01			
6.810000 E+01	6.550000 E+01			
6.520000 E+01	6.250000 E+01			
6.740000 E+01	6.480001 E+01			
6.340000 E+01	6.060000 E+01			
6.780000 E+01	6.380000 E+01			
6.720000 E+01	6.550000 E+01			
5.820000 E+01	5.550000 E+01			
5.890000 E+01	5.540000 E+01			
6.500000 E+01	6.150000 E+01			
5.790000 E+01	5.320000 E+01			
6.010000 E+01	5.600000 E+01			
6.500000 E+01	6.170000 E+01			
6.600000 E+01	6.370000 E+01			
6.560000 E+01	6.250000 E+01			
6.250000 E+01	5.970000 E+01			
6.410000 E+01	6.010000 E+01			
5.810000 E+01	5.520000 E+01			

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	0.0000E+00	-----	-----	-----	-----	-----
BI 11	9.4878E-01	2.5485E-03	3.7229E+02		Zamítnuta	0.000

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R	: 9.9009E-01
Koeficient determinace, R ²	: 9.8027E-01
Predikovaný korelační koeficient, Rp ²	: 9.8924E-01
Střední kvadratická chyba predikce, MEP	: 6.6072E-01
Akaikeho informační kritérium, AIC	: -1.0425E+01

(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	5.0300E+01	4.9337E+01	1.3252E-01	9.6345E-01	1.9154E+00
2	4.2500E+01	4.2980E+01	1.1545E-01	-4.7972E-01	-1.1288E+00
3	5.0800E+01	5.1519E+01	1.3838E-01	-7.1874E-01	-1.4148E+00
4	5.7500E+01	5.8255E+01	1.5648E-01	-7.5508E-01	-1.3132E+00

Napověda-F1 Řádek: 72 - 94 Celkem: 244 Délka: 14094

Závěry a vysvětlení výstupu úlohy V622

Úkoly

Odpovědi, závěry a vysvětlení

(1) Popište test významnosti absolutního členu.

Protože úsek β_0 vyšel jako statisticky nevýznamný, čili roven nule, přímka prochází počátkem a nová metoda není zatížena systematickou chybou.

(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.
BI 01	-6.5651E-01	3.1432E+00	-2.0887E-01	akceptována	0.836
BI 11	9.5724E-01	4.9941E-02	1.9167E+01	Zamítnuta	0.000

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R	: 9.6762E-01
Koeficient determinace, R ²	: 9.3629E-01
Predikovaný korelační koeficient, Rp ²	: 9.6165E-01
Střední kvadratická chyba predikce, MEP	: 2.3004E+00
Akaikeho informační kritérium, AIC	: 2.2006E+01

(2) Vysvětlete test shodnosti odhadu parametru β s předepsanou $\beta_0 = 0$.

Přímka prochází počátkem a R je blízké jedné. Směrnice je mírně pod 1, čili nová metoda slabě podhodnocuje. MEP a AIC rozhoduje o nejlepším modelu a o b_0 rovném nule.

(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.
BI 01	0.0000E+00				
BI 11	9.4878E-01	2.5485E-03	3.7229E+02	Zamítnuta	0.000

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R	: 9.9009E-01
Koeficient determinace, R ²	: 9.8027E-01
Predikovaný korelační koeficient, Rp ²	: 9.8924E-01
Střední kvadratická chyba predikce, MEP	: 6.6072E-01
Akaikeho informační kritérium, AIC	: -1.0425E+01