

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

ZÁVAŽNÉ PŘEDPOKLADY MNČ - kdy je určitě nutné korigovat klasickou MNČ

- 1) **Regresní parametry β** mohou teoreticky nabývat libovolných hodnot.
- 2) Regresní model je **lineární v parametrech**.
- 3) Jednotlivé nezávislé proměnné jsou skutečně **vzájemně nezávislé**, tedy mezi nimi nedochází k tzv. **multikolinearitě**.
- 4) **Podmíněný rozptyl $D(y/x) = \sigma^2$** je konstantní (tzv. podmínka homoskedasticity).
- 5) **Náhodné chyby** mají nulovou střední hodnotu $E(\varepsilon_i) = 0$, mají konečný rozptyl $E(\varepsilon_i^2) = \sigma^2$ a jsou nekorelované.

Kritika metody

Ověření předpokladů
metody nejmenších čtverců

Metoda:

Vyšetření heteroskedasticity

Vyšetření autokorelace

Vyšetření multikolinearity

Vyšetření normality náhodných chyb

Vyšetření omezení parametrů

Vyšetření trendů reziduí

REGRESNÍ DIAGNOSTIKA testy vybraných předpokladů klasické MNČ

Multikolinearita: VIF diagnostika indikuje

Heteroskedasticita: testy heteroskedasticity (např. Cook-Weisberg)

Autokorelace reziduí: test významnosti autokorelačního koeficientu

Normalita reziduí: testy normality

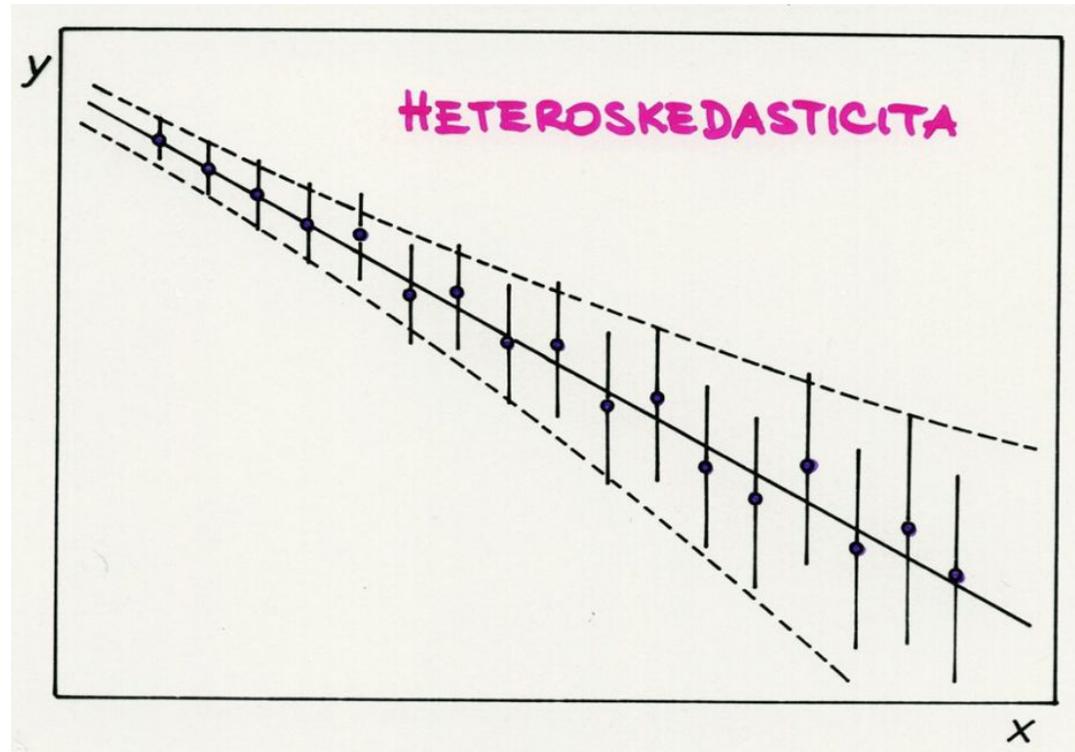
Ověření předpokladů MNČ

1. Hetero/homoskedasticita (nekonstantnost rozptylu)

Rozptyl veličiny y_i v itém bodě je popsán

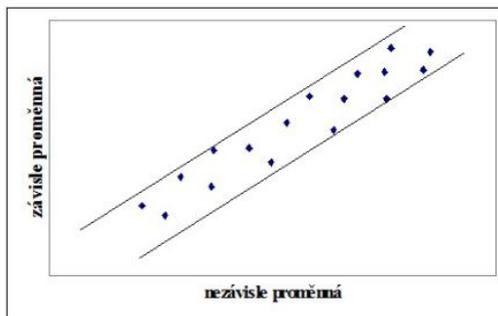
$$\sigma_i^2 = \sigma^2 \exp(\lambda \mathbf{x}_i \boldsymbol{\beta}),$$

kde \mathbf{x}_i je i -tý řádek matice \mathbf{X} .

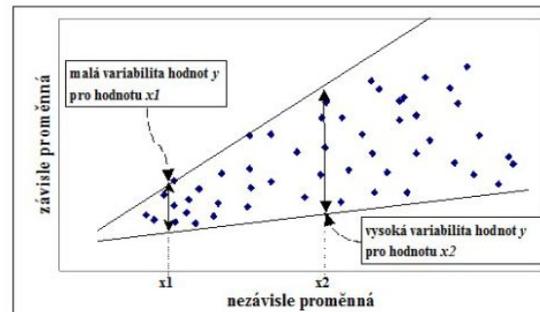


HOMOSKEDASTICITA vs. ETEROSKEDASTICITA Grafické vysvětlení principu

Homoskedasticita znamená, že hodnoty závisle proměnné y mají pro všechny hodnoty nezávisle proměnné x **konstantní rozptyl** (variabilitu, proměnlivost).



Homoskedasticita



Heteroskedasticita

Testování homoskedasticity testem trendu reziduí

$$D = \sum_{i=1}^n \left[R(|\hat{\epsilon}_i|) - i \right]^2$$

$$\rho_s = 1 - \frac{6}{n^3 - n} \cdot D$$

Testujeme významnost Spearmanova korelačního koeficientu ρ_s

$$t_R = \frac{\rho_s \cdot \sqrt{n-2}}{\sqrt{1-\rho_s^2}}$$

Cookův - Weisbergův test homoskedasticity

Vycházíme z předpokladu, že rozptyl naměřené hodnoty y_i je určitou funkcí proměnné x_i β (např. exponenciální funkcí).

$$S_f = \frac{\left[\sum_{i=1}^n (y'_i - \bar{y}')^2 e_i^2 \right]^2}{2 \cdot \sigma^4 \sum_{i=1}^n (y'_i - \bar{y}')^2}$$

Test: Pokud v datech není heteroskedasticita potom platí, že $S_f < \chi^2(1)$.

HETEROSKEDASTICITA Vhodné řešení v lineární regresi při nesplnění jednoho předpokladu klasické MNČ

Nejčastějším způsobem je užití metody **vážených nejmenších čtverců**,

kdy se účelová funkce čili suma čtverců reziduí násobí

vhodně zvolenými váhami

$$U(\mathbf{b}) = \sum_{i=1}^n \left[y_i V_{ii} - \sum_{j=1}^m V_{ij} x_{ij} b_j \right]^2$$

V běžných případech je možné jako váhy volit hodnoty $1/y_i$ nebo $1/y_i^2$.

Test: ověření $H_0: \lambda = 0$ (homoskedasticita).
Cook-Weisbergovo testační kritérium

$$S_f = \frac{\left[\sum_{i=1}^n (\hat{y}_i - \hat{y}_p) \hat{e}_i^2 \right]^2}{2 \hat{\sigma}^4 \sum_{i=1}^n (\hat{y}_i - \hat{y}_p)^2}$$

kde

$$\hat{y}_p = \frac{1}{n} \sum_{i=1}^n \hat{y}_i$$

Testování:

1. Je-li $S_f < \chi^2(1)$, H_0 (homoskedasticita) je přijata.
2. Pro homoskedasticitu tvoří diagnostický graf

\hat{e}_{Si}^2 na $(1 - H_{ii}) \hat{y}_i$
náhodný mrak bodů. Pro heteroskedasticitu vznikne
typický klínový obrazec

2. Autokorelace

Data časových řad mají chyby ε_i vzájemně korelované.

Nejčastější je případ autokorelace prvního řádu

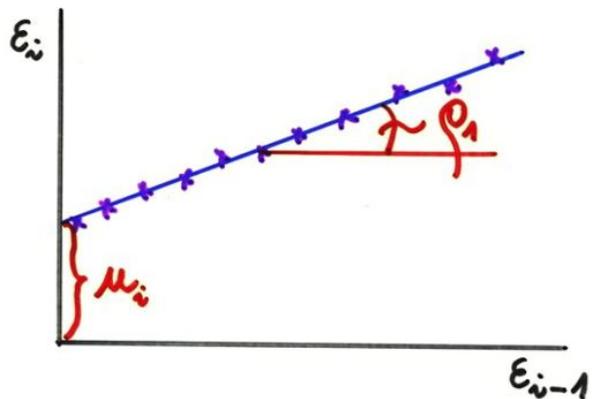
$$\varepsilon_i = \rho_1 \varepsilon_{i-1} + u_i$$

kde $u_i \sim N(0, \sigma^2)$.

- a) Pro $\rho_1 = 1$ případ kumulativních chyb, který se v chemii vyskytuje často.
- b) Pro $\rho_1 \leq 1$ jde o autokorelační koeficient 1. řádu.

Test:

1. Grafická indikace autokorelace:



2. vzorová úloha na výstavbu lineárního regresního modelu

P638

Test autokorelace

2. Waldův test pro ρ_1 : $H_0: \rho_1 = 0$ vs. $H_A: \rho_1 \neq 0$

Je-li Waldovo kritérium $W_a = \frac{n \hat{\rho}_1^2}{1 - \hat{\rho}_1^2} < \chi^2(1)$, H_0 je přijata.

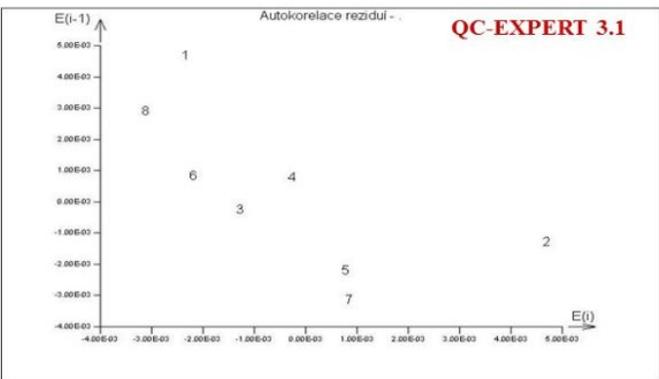
Příklad 6.38 Test autokorelace u kinetických dat

Byla sledována kinetika inverze sacharidu v 1M HCl při 30 °C. Lze očekávat, že mezi chybami ϵ_i existuje autokorelace jako důsledek postupu odebírání vzorků. Určete graficky významnost autokorelace v datech.

Data: x je čas [min]; y je desetinásobek logaritmu podílu zbylého nezreagovaného sacharidu v reakční směsi.

x	y
0	1.000
10	0.954
20	0.895
30	0.843
40	0.791
50	0.735
60	0.685
70	0.628
80	0.581

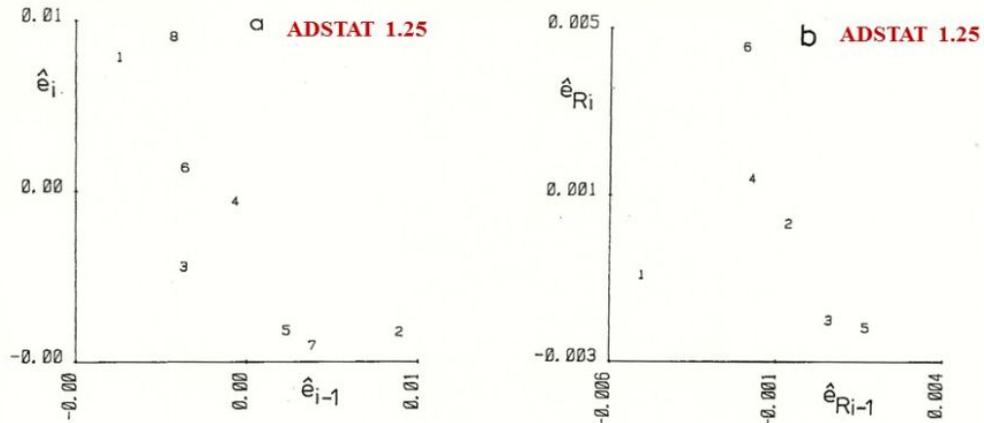
Waldův test autokorelace		Durbin-Watsonův test autokorelace	
Hodnota kritéria WA :	3.454228	Hodnota kritéria DW :	-1.0
Kvantil $\chi^2(1-\alpha, 1)$:	3.841459	Kritické hodnoty DW :	0.0 2.0
Pravděpodobnost :	0.06309	Závěr :	Negativní autokorelace reziduí není prokázána.
Závěr :	Autokorelace je nevýznamná		



14

Řešení:

- Graf závislosti $\hat{\epsilon}_i$ na $\hat{\epsilon}_{i-1}$ a graf závislosti $\hat{\epsilon}_{Ri}$ na $\hat{\epsilon}_{Ri-1}$ ukazují výraznou negativní autokorelaci.
- Waldovo testovací kritérium autokorelace $W_a = 9.424$ je větší než $\chi_{0.95}^2(2) = 5.99$, což indikuje autokorelaci reziduí.



Závěr: Grafy umožňují odhadnout znaménko a velikost autokorelačního koeficientu 1. řádu ρ_1 .

Normalita chyb

1. **Rankitový (Q-Q) graf:** $\hat{e}_{(i)}$ nebo $\hat{e}_{R(i)}$ na u_{P_i}
pro $P_i = i / (n + 1)$

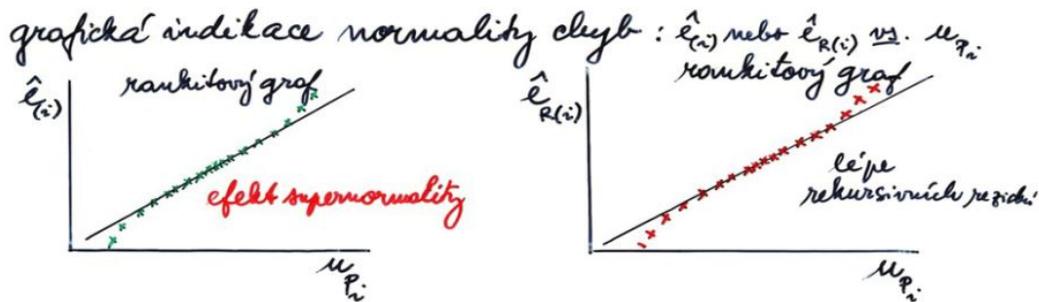
2. **Test normality:** H_0 : normalita vs. H_A : nenormalita
Jarque-Berrova testační statistika

$$L(\hat{e}) = n \left[\frac{\hat{u}_3^2}{6 \hat{u}_2^3} + \frac{\hat{u}_4 - 3}{24} \right] + n \left[\frac{3 \hat{u}_1^2}{2 \hat{u}_2} - \frac{\hat{u}_3 \hat{u}_1}{\hat{u}_2^2} \right]$$

kde \hat{u}_j je j-tý výběrový moment reziduí

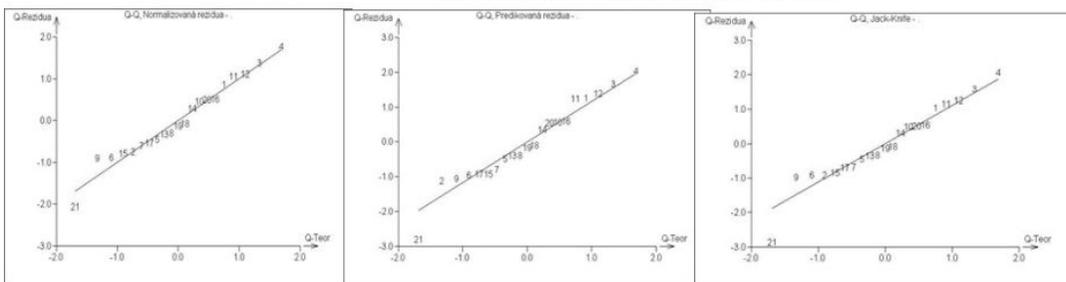
$$\hat{u}_j = \frac{\sum_{i=1}^n \hat{e}_i^j}{n}$$

NORMALITA CHYB



Příklad P6.51

Jarque-Berrův test normality, Hodnota kritéria JB : 0.080308
Kvantil $\chi^2(1-\alpha, 2)$: 5.991465, Pravděpodobnost: 0.960641,
Závěr: Rezidua mají normální rozdělení.



Test: je-li $L(\hat{e}) > \chi^2_{1-\alpha}(2) = 5.99$, je H_0 (normalita) zamítnuta.

Např. pro lineární modely s absolutním členem je $\hat{u}_1 = 0$ a $L(\hat{e})$ je

$$L(\hat{e}) = n \left(\frac{\hat{g}_1}{6} + \frac{(\hat{g}_2 - 3)^2}{24} \right)$$

kde

$$\hat{g}_1 = \frac{\hat{u}_3^2}{\hat{u}_2^3} \quad \hat{g}_2 = \frac{\hat{u}_4}{\hat{u}_2^2}$$

(test není vhodný pro malé výběry)

Vyšetření předpokladů u MNČ při testování regresního tripletu

Fisher-Snedecorův test významnosti modelu, Hodnota kritéria F : 59.57416

Kvantil F (1-alfa, m-1, n-m) : 3.196777, Pravděpodobnost : 0.0, Závěr : Model je významný

Scottovo kritérium multikolinearity, Hodnota kritéria SC : 0.626359

Závěr : Model vykazuje multikolinearitu!

Cook-Weisbergův test heteroskedasticity, Hodnota kritéria CW : 1.059618

Kvantil $\chi^2(1-\alpha, 1)$: 3.841459, Pravděpodobnost: 0.303302, Závěr: Rezidua ukazují homoskedasticitu.

Jarque-Berrův test normality, Hodnota kritéria JB : 0.080308

Kvantil $\chi^2(1-\alpha, 2)$: 5.991465, Pravděpodobnost: 0.960641, Závěr: Rezidua mají normální rozdělení.

Waldův test autokorelace, Hodnota kritéria WA : 0.001722

Kvantil $\chi^2(1-\alpha, 1)$: 3.841459, Pravděpodobnost: 0.966903, Závěr: Autokorelace je nevýznamná.

Durbin-Watsonův test autokorelace, Hodnota kritéria DW: -1.0,

Kritické hodnoty DW 0.93 1.81, Závěr: Pozitivní autokorelace reziduí není prokázána.

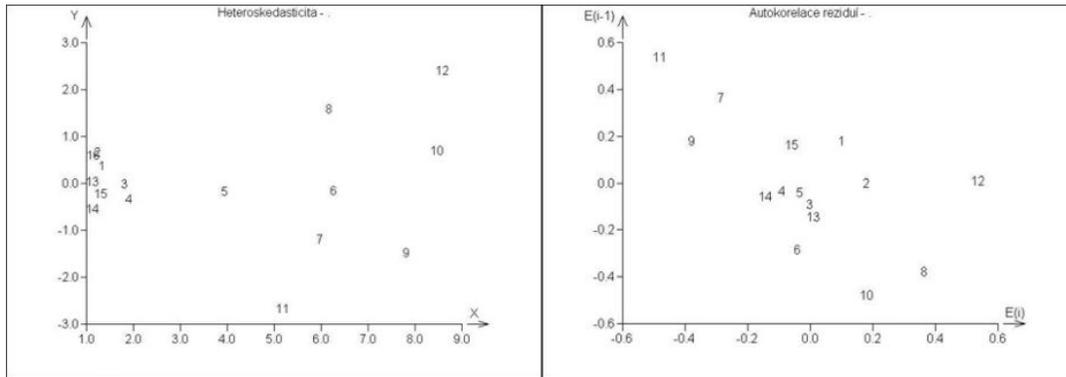
Znaménkový test reziduí: Hodnota kritéria Sg: 0.555662,

Kvantil $N(1-\alpha/2)$: 1.959964, Pravděpodobnost: 0.578442, Závěr: V reziduích není trend.

3. vzorová úloha na hledání metody výstavby lineárního regresního modelu

M619

21



Cook-Weisbergův test heteroskedasticity

Hodnota kritéria CW : 5.373857
Kvantil $\chi^2(1-\alpha, 1)$: 3.841459
Pravděpodobnost : 0.020441

Závěr: Rezidua vykazují heteroskedasticitu!

Waldův test autokorelace

Hodnota kritéria WA : 5.003828
Kvantil $\chi^2(1-\alpha, 1)$: 3.841459
Pravděpodobnost : 0.025291

Závěr: Autokorelace je významná

Durbin-Watsonův test autokorelace

Hodnota kritéria DW : -1.0
Kritické hodnoty DW : 0.74 1.93

Závěr: Negativní autokorelace reziduí není prokázána.

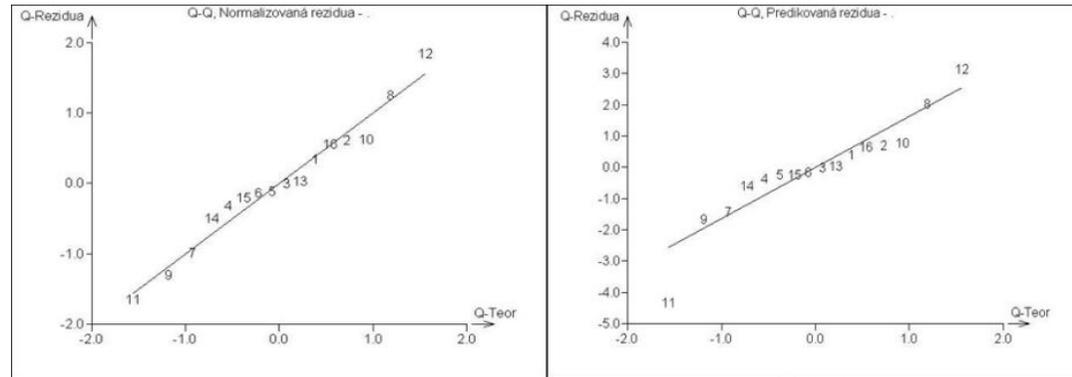
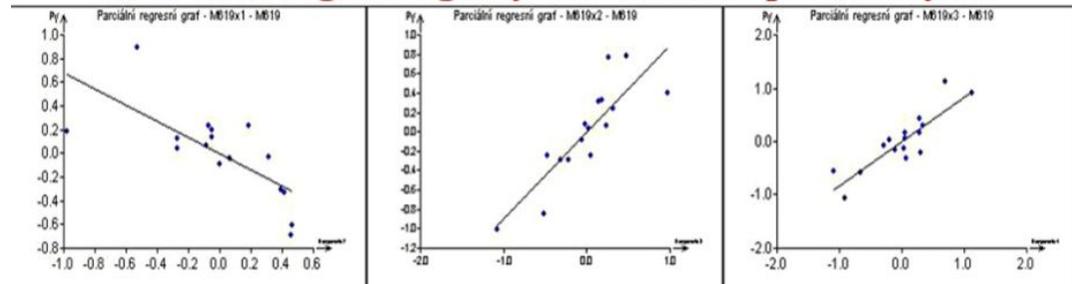
23

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

Výstavba lineárního regresního modelu:

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

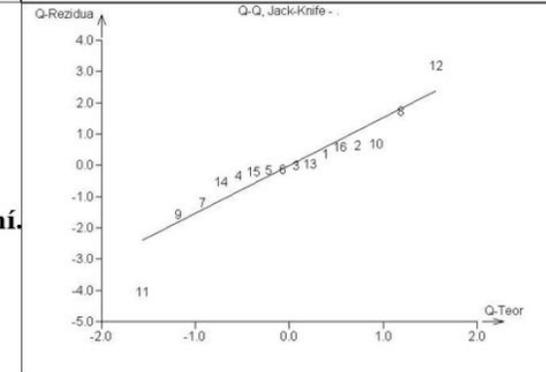
Parciální regresní grafy tří nezávisle proměnných



Jarque-Berrův test normality

Hodnota kritéria JB : 0.033085
Kvantil $\chi^2(1-\alpha, 2)$: 5.991465
Pravděpodobnost : 0.983593

Závěr: Rezidua mají normální rozdělení.



24

Vyšetření předpokladů u MNČ při testování regresního tripletu

Fisher-Snedecorův test významnosti modelu, Hodnota kritéria F: 1404.982774
Kvantil F (1-alfa, m-1, n-m) : 3.490295, Pravděpodobnost: 0.0, **Závěr: Model je významný**

Scottovo kritérium multikolinearity, Hodnota kritéria SC: 0.970655, **Závěr: Model je nekorektní!**

Cook-Weisbergův test heteroskedasticity, Hodnota kritéria CW: 5.373857,
Kvantil $\chi^2(1-\alpha, 1)$: 3.841459, Pravděpodobnost: 0.020441,
Závěr: Rezidua vykazují heteroskedasticitu!

Jarque-Berrův test normality, Hodnota kritéria JB: 0.033085, Kvantil $\chi^2(1-\alpha, 2)$: 5.991465,
Pravděpodobnost: 0.983593, **Závěr: Rezidua mají normální rozdělení.**

Waldův test autokorelace, Hodnota kritéria WA: 5.003828,
Kvantil $\chi^2(1-\alpha, 1)$: 3.841459, Pravděpodobnost: 0.025291, **Závěr: Autokorelace je významná**

Durbin-Watsonův test autokorelace, Hodnota kritéria DW: -1.0, Kritické hodnoty DW 0.74, 1.93,
Závěr: Negativní autokorelace reziduí není prokázána.

Znaménkový test reziduí, Hodnota kritéria Sg: 0.328976, Kvantil $N(1-\alpha/2)$: 1.959964,
Pravděpodobnost: 0.742174, **Závěr: V reziduích není trend.**

25

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

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

Úkoly: (1) Vyšetřete regresní triplet.

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

(3) Proveďte Cookův-Weisbergův test heteroskedasticity v datech.

(4) Proveďte autokorelační test.

(6) Proveďte Jarque-Berrův test normality v datech.

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

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

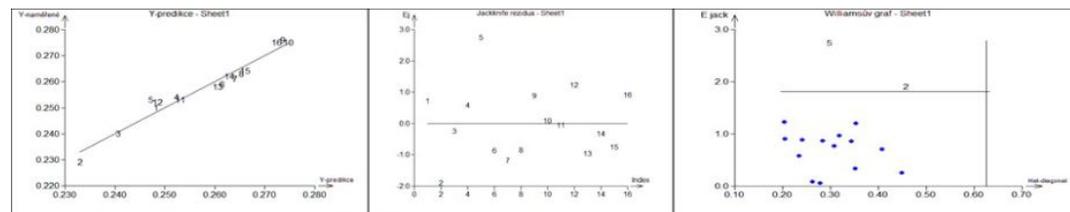
27

Úlohy na výstavbu lineárního regresního modelu

Kritika metody

Software QC-EXPERT 3.1, ADSTAT 1.25

26



Odhady parametrů ze všech bodů a se všemi parametry

Prom	Odhad	Směr. Odch.	Závěr	Pravděpod	Spodní mez	Horní mez
Abs	0,4618543406	0,02249566045	Významný	4,034304002E-010	0,4123417258	0,5113669555
M604x1	-0,009726835	0,002207118652	Významný	0,001050871796	-0,0145846707	-0,004868999907
M604x2	-0,002128889	0,002625238302	Nevýznamný	0,4345998	-0,00790700051	0,003649220578
M604x3	0,0018900840	0,000522373664	Významný	0,004038736893	0,0007403473462	0,003039820715
M604x4	-0,003415052	0,000392450423	Významný	2,910919431E-006	-0,0042788304	-0,002551275286

Statistické charakteristiky regrese

Vícenásobný korelační koeficient R :
Koeficient determinace R² : 0,9515141007
Predikovaný korelační koeficient Rp :
Střední kvadratická chyba predikce MEP :
Akaikeho informační kritérium : -179,8510432

Testování regresního tripletu

Fisher-Snedecorův test významnosti modelu
Hodnota kritéria F : 33,96752076
Kvantil F (1-alfa, m-1, n-m) : 3,35669021
Pravděpodobnost : 3,677942646E-007
Závěr : Model je významný

Scottovo kritérium multikolinearity
Hodnota kritéria SC : -0,325613384
Závěr : Model je korektní.

Cook-Weisbergův test heteroskedasticity
Hodnota kritéria CW : 0,6529596655
Kvantil $\chi^2(1-\alpha, 1)$: 3,841458829
Pravděpodobnost : 0,4190563984

Jarque-Berrův test normality
Hodnota kritéria JB : 0,8427678152
Kvantil $\chi^2(1-\alpha, 2)$: 5,991464547
Pravděpodobnost : 0,6561381566
Závěr : Rezidua mají normální rozdělení.

Reziduální součet čtverců : 0,0001124388005
Průměr absolutních reziduí : 0,002246979059
Reziduální směr. odchylka : 0,003197140775
Reziduální rozptyl : 1,022170913E-005
Šířmost reziduí : 0,2085152978
Špičatost reziduí : 2,344187121

Waldův test autokorelace

Hodnota kritéria WA : 0,09080762796
Kvantil $\chi^2(1-\alpha, 1)$: 3,841458829
Pravděpodobnost : 0,7631529238
Závěr : Autokorelace je nevýznamná

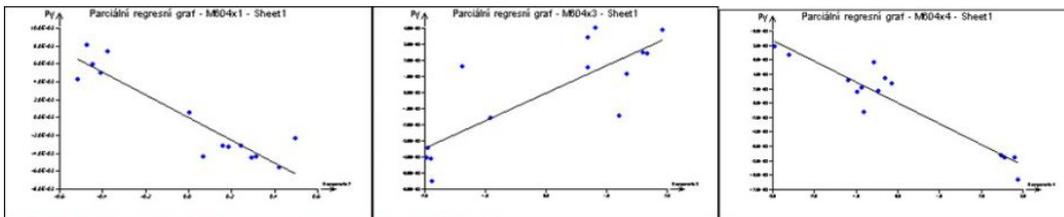
Durbin-Watsonův test autokorelace

Hodnota kritéria DW : -1
Kritické hodnoty DW : 0,62
Závěr : Negativní autokorelace reziduí není prokázána.

Znaménkový test reziduí

Hodnota kritéria Sg : 0,3289758475
Kvantil $N(1-\alpha/2)$: 1,959963999
Pravděpodobnost : 0,7421739438
Závěr : V reziduích není trend.

28



Odhady parametrů bez nevýznamného M604x2

Proměnná	Odhad	Směr.Odch.	Závěr	Pravděpodobnost	Spodní mez	Horní mez
Abs	0,4670272783	0,01613459391	Významný	5,64468472E-011	0,4310771628	0,5029773939
M604x1	-0,01266489713	0,001681884114	Významný	1,992614648E-005	-0,01641236847	-0,008917425794
M604x3	0,001716894733	0,0003944326548	Významný	0,001437212579	0,0008380440105	0,002595745456
M604x4	-0,002891844086	0,0003166485876	Významný	3,627229233E-006	-0,003597381106	-0,002186307066

Statistické charakteristiky regrese

Vicenasobný korelační koeficient R :	0,9822222557	Reziduální součet čtverců :	4,897275363E-005
Koeficient determinace R ² :	0,9647605597	Průměr absolutních reziduí :	0,001378843487
Předikovaný korelační koeficient Rp :	0,8750470594	Reziduální směr. odchylka :	0,002212978844
Střední kvadratická chyba predikce MEP :	6,408617E-006	Reziduální rozptyl :	4,897275363E-006
Akaikeho informační kritérium :	-167,8862531	Šikmost reziduí :	0,1094674598
		Špičatost reziduí :	3,015790211

Testování regresního tripletu bez outlierů

Fisher-Snedecorův test významnosti modelu

Hodnota kritéria F :	91,25765061
Kvantil F (1-alfa, m-1, n-m) :	3,708264819
Pravděpodobnost :	1,44932095E-007
Závěr :	Model je významný

Scottovo kritérium multikolinearity

Hodnota kritéria SC :	-0,4639553173
Závěr :	Model je korektní.

Cook-Weisbergův test heteroskedasticity

Hodnota kritéria CW :	0,000848921306
Kvantil Chi ² (1-alfa,1) :	3,841458829
Pravděpodobnost :	0,9767559214
Závěr :	Rezidua vykazují homoskedasticitu.

Jarque-Berrův test normality

Hodnota kritéria JB :	0,2555695158
Kvantil Chi ² (1-alfa,2) :	5,991464547
Pravděpodobnost :	0,880042781
Závěr :	Rezidua mají normální rozdělení.

Waldův test autokorelace

Hodnota kritéria WA :	0,261862158
Kvantil Chi ² (1-alfa,1) :	3,841458829
Pravděpodobnost :	0,6088437041
Závěr :	Autokorelace je nevýznamná

Durbin-Watsonův test autokorelace

Hodnota kritéria DW :	-1
Kritické hodnoty DW :	0
Závěr :	Negativní autokorelace reziduí není prokázána.

29

Úloha M6.21 Vliv pěti parametrů na obsah manganu ve vodárenské vodě – předpoklady MNČ

Zadání: Naleznete závislost mezi obsahem rozpuštěného manganu y a pěti parametry: teplotou vody x_1 , obsahem kyslíku x_2 , pH x_3 , průtokem vody vypouštěné x_4 a časem x_5 .

Úkoly:

- (1) Jsou všechny regresní parametry statisticky významné?
- (2) Vyšetřete regresní triplet a odhalte vlivné body.
- (3) Vyšetření porušení předpokladů MNČ.

Data: Teplota vody x_1 [°C], obsah kyslíku x_2 [mg/l], pH x_3 , průtok vody vypouštěné dnovými výpusťmi x_4 [%] a časem x_5 [týden], obsah rozpuštěného manganu y [µg/l]:

x_1	x_2	x_3	x_4	x_5	y
3.4	9.9	7.7	16.26	17	10
...
4.9	5	6.9	6.55	49	410

30

Úloha M6.40 Závislost koncentrace kreatininu na hmotnosti, věku a výšce mužů – předpoklady MNČ

Zadání: Zjistěte, zda existuje závislost celkové koncentrace kreatininu y u zdravých mužů na jejich hmotnosti x_1 , věku x_2 a výšce x_3 . Výběr je tvořen muži, u kterých nebylo prokázáno renální onemocnění.

Úkoly:

- (1) Postavte lineární regresní model a vyšetřete regresní triplet.
- (2) Vyšetřete statistickou významnost jednotlivých regresních parametrů.
- (3) Užijte také parciální regresní grafy a komentujte fyzikální význam jednotlivých regresních parametrů.
- (4) Vyšetřete porušení předpokladů MNČ.

Data: Věk x_1 [roky], hmotnost x_2 [kg], výška x_3 [cm], koncentrace kreatininu y [µmol/l] :

x_1	x_2	x_3	y
20	70	180	66
...
75	84	179	119

31

32

Úloha M6.45 Vliv čtyř parametrů na koncentraci TSH v kontrolním séru – předpoklady MNČ

Zadání: Souprava na stanovení TSH je dlouhodobě sledována i pomocí interních kontrolních sér. Data obsahují kromě koncentrace TSH v séru také šarže složek použitých v jednotlivých analýzách a datum provedení analýzy.

Úkoly:

- (1) Vysvětlete všech sedm předpokladů MNČ a regresní triplet.
- (2) Jak se řeší úloha v případě porušení předpokladů MNČ, a to především při heteroskedasticitě, autokorelaci a nenormalitě chyb?

Data: Šarže radioindikátoru x_1 , šarže potahovaných zkumavek x_2 , šarže kalibrátorů x_3 , datum analýzy (formát ymmdd) x_4 , koncentrace TSH v séru (mIU/L) y :

x_1	x_2	x_3	x_4	y
313	1142	1057	10904	0.72
...
330	1151	1060	2050	0.7

Úloha M6.54 Vliv glukosy, proteinů, albuminů a globulinů na přírůstek hmotnosti kuřat – předpoklady MNČ

Zadání: Před začátkem studie byl stanoven přírůstek váhy a biochemické parametry testovaných kuřat. Zjistěte jestli glukosa, celkové množství proteinů, obsah albuminů a globulinů má vliv na přírůstek hmotnosti kuřat.

Úkoly:

- (1) Vyšetřením regresního tripletu nalezněte nejlepší model.
- (2) Využijte regresní diagnostiku a parciální regresních a parciální reziduální grafy ke statistické významnosti jednotlivých parametrů.

Data: Obsah glukosy x_1 [mmol.L⁻¹], celkový obsah proteinů x_2 [g.L⁻¹], obsah albuminů x_3 [g.L⁻¹], obsah globulinů x_4 [g.L⁻¹], přírůstek váhy v kg za den y [kg]:

x_1	x_2	x_3	x_4	y
13.1	33	28	5	0.016
...
13.6	29	25	4	0.079

