

# **SEMESTRÁLNÍ PRÁCE**

**Licenční studium**

## **STATISTICKÉ ZPRACOVÁNÍ DAT PŘI KONTROLE A ŘÍZENÍ JAKOSTI**

**Předmět**

### **TVORBA LINEÁRNÍCH REGRESNÍCH MODELŮ PŘI ANALÝZE DAT**

**(QC EXPERT)**

**MEDICAMENTA, a.s.**

**Ing. Martina Krátka**

## Příklad 1 – Porovnání dvou regresních přímek u jednoduchého lineárního regresního modelu

### Porovnání 2 instrumentálních technik

**Zadání:** Obsah paracetamolu v přípravku Ataralgin se stanovuje pomocí UV nebo HPLC. Pro porovnání obou technik byly naměřeny kalibrační přímky. Porovnejte tyto kalibrační přímky včetně testování jejich úseku a směrnice, s vyšetřením vlivných bodů a posouzením míry spolehlivosti navrženého modelu. Proveďte test dvou přímek, test jejich rovnoběžnosti a test společného úseku. Jaké jsou míry přesnosti obou kalibračních přímek, jsou shodné?

**Data:**

Paracetamol [mg/tbl]	HPLC	UV
50	50,600	49,781
70	70,267	71,353
80	81,267	81,727
100	100,133	101,230
120	120,067	119,382
130	130,467	128,954
150	150,000	149,866

### Výpočet:

V S T U P

#### ZVOLENA STRATEGIE REGRESNÍ ANALÝZY:

Omezení, P : 1.0000E-34  
Transformace : Ne  
Váhy : Ne  
Absolutní člen zahrnut : Ano

#### PODMÍNKY A KVANTILY PRO STATISTICKÉ TESTY:

Hladina významnosti, alfa : 0.050  
Počet bodů, n : 7  
Počet parametrů, m : 2  
Kvantil Studentova rozdělení  $t(1-\alpha/2, n-m)$  : 2.571  
Kvantil  $F(1-\alpha, m, n-m)$  : 6.608  
Kvantil rozd. Chí-kvadrát  $\text{Chi-square}(1-\alpha, m)$  : 5.991

# I. Řešení dat získaných při HPLC stanovení

## 1. Návrh modelu

$$Y = \beta_0 + \beta_1 x$$

## 2. Předběžná analýza dat

### Charakteristiky proměnných

Proměnná	Průměr	Směr.Odch.	Kor.vs.Y	Významnost
y	100,40	35,37	1,0000	-----
x	100	35,59	0,99994	4,67E-011

### Indikace multikolinearity

Proměnná	Vlas. čísla kor. m.	Podmíněnost kappa	VI faktor	Vícenás. kor. $\beta_0$
	1	1	1	0
$B_1$	1	1	1	0

Maximální číslo podmíněnosti K : 1.0000E+00  
( $K_{jj}$ ),  $K > 1000$  indikuje silnou multikolinearitu)  
( $VIF_{jj} > 10$  indikuje silnou multikolinearitu)

## 3. Odhadování parametrů

Test:  $H_0: \beta_j$  vs.  $H_1: \beta_j \neq 0$

Proměnná	Odhad	Směr.Odch.	Závěr	Pravděp.	Spodní mez	Horní mez
$\beta_0$	1,032	0,4988	Nevýzn.	0,093	-0,251	2,314
$\beta_1$	0,994	0,00474	Význam.	4,7E-011	0,982	1,006

Úsek  $\beta_0$  je statisticky nevýznamný, směrnice  $\beta_1$  je statisticky významná.

## 4. Základní statistické charakteristiky

Vícenásobný korelační koeficient R : 0,99994

Koeficient determinace  $R^2$  : 0,99989

Predikovaný korelační koeficient  $R_p$  : 0,99963

Střední kvadratická chyba predikce MEP : 0,19933

Akaikeho informační kritérium : -10,7353

Absolutní hodnota párového korelačního koeficientu R ukazuje, že navržený lineární regresní model je statisticky významný. Vysoká hodnota koeficient determinance  $R^2$  představuje % variability vysvětlené modelem. Predikovaný korelační koeficient  $R_p^2$  ukazuje na predikční schopnost modelu. Střední kvadratická chyba predikce MEP a Akaikeho informační kritérium AIC se užívají k rozlišení mezi několika navrženými modely. Za optimální se považuje model, pro který dosahuje MEP a AIC minimální hodnotu.

## 5. Regresní diagnostika

### a. Analýza klasických reziduí

Bod	naměř.Y	vyp.Y	Směr.odch.Y	Reziduuum	Reziduuum [%Y]
1	50,600	50,7159	0,28369	-0,1159	-0,2291
2	70,267	70,5896	0,21111	-0,3226	-0,4591
3	81,267	80,5265	0,18261	0,74054	0,9112
4	100,133	100,4001	0,15610	-0,26714	-0,2668
5	120,067	120,2738	0,18261	-0,20683	-0,1723
6	130,467	130,2107	0,21111	0,256331	0,1965
7	150,00	150,0844	0,28369	-0,08435	-0,0562

Reziduální součet čtverců	: 0,8529
Průměr absolutních reziduí	: 0,2848
Reziduální směr. odchylka	: 0,4130
Reziduální rozptyl	: 0,1706
Šikmost reziduí	: 1,4567
Špičatost reziduí	: 3,1087

### b. Analýza ostatních reziduí

Indikace vlivných dat

#### A. Analýza reziduí

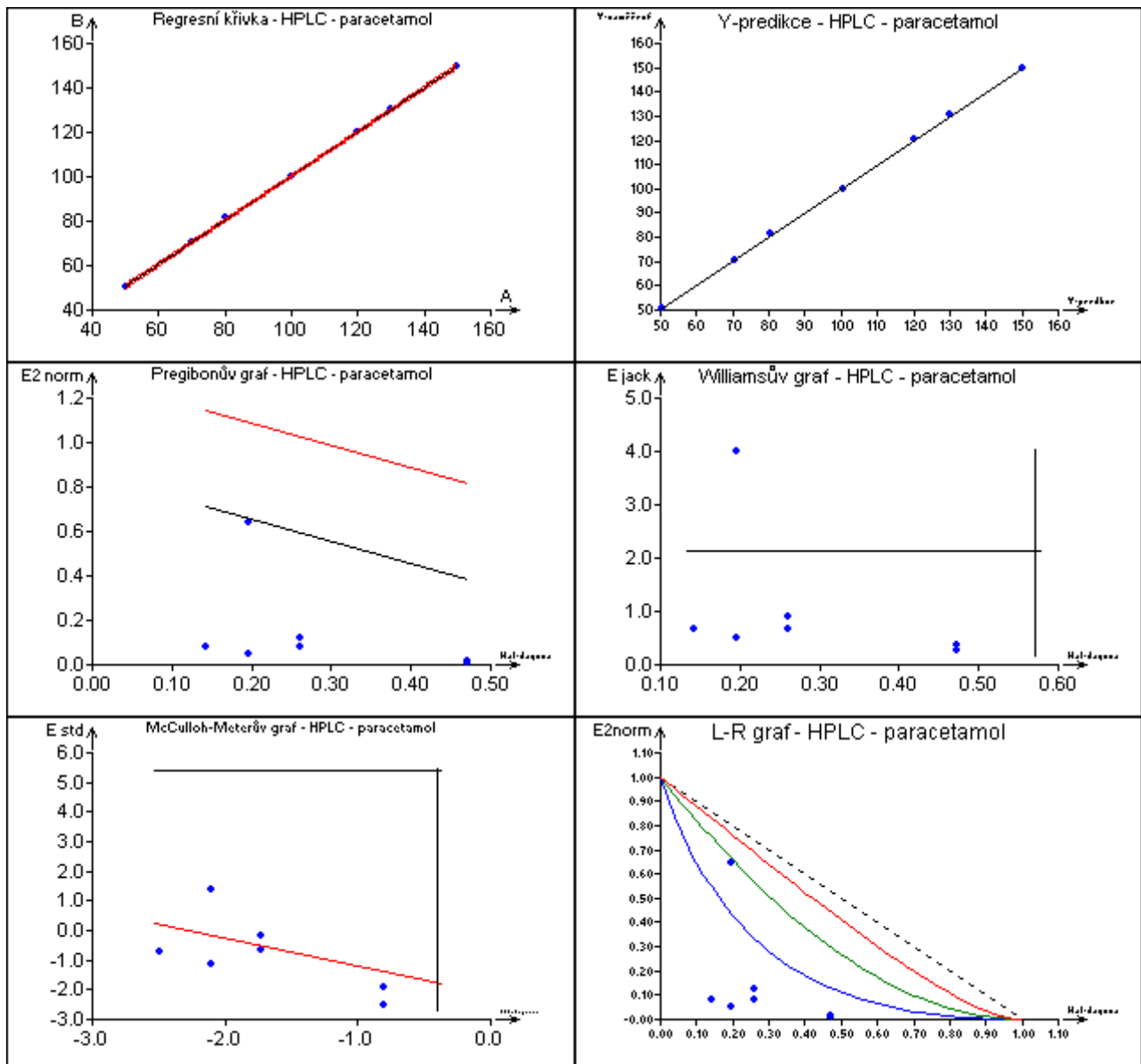
Bod	Standardní	Jackknife	Predikované	Diag(Hii)	Diag(H*ii)	Cook.vzdál.
1	-0,3862	-0,3507	-0,2195	0,4718	0,4876	-0,1725
2	-0,9088	-0,8897	-0,4367	0,2613	0,3833	-0,1607
3	1,9990	3,99047	0,92049	0,1955	0,8385	0,2429
4	-0,6986	-0,6578	-0,3117	0,1429	0,2265	-0,0582
5	-0,5583	-0,5157	-0,2571	0,1955	0,2456	-0,0678
6	0,7221	0,6824	0,34699	0,2613	0,3383	0,1277
7	-0,2810	-0,2534	-0,1597	0,4718	0,4801	-0,1255

#### B. Analýza vlivu

Index	Atkinsonova vzdál.	Andrews-Pregibon st.	Vliv na $Y^{\wedge}$
1	0,5241098695	0,5124368877	-0,3314761864
2	0,836603026	0,6166875405	-0,5291142119
3	<b>3,110207301</b>	0,1615169305	<b>1,967067813</b>
4	0,424620605	0,7734678694	-0,2685536507
5	0,4019467005	0,7543552343	-0,2542134143
6	0,6417136485	0,6616828817	0,405855347
7	0,3786171974	0,5198526559	-0,239458541

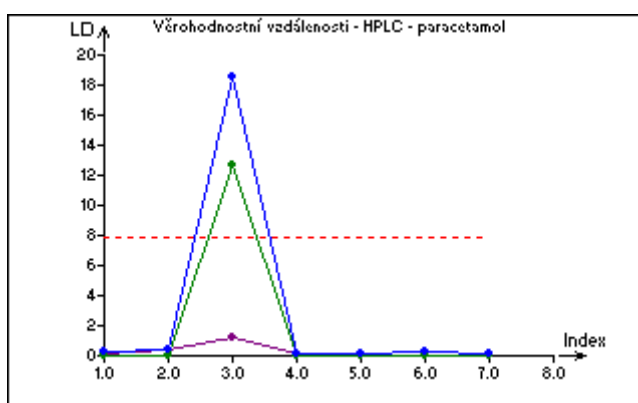
Index	Vliv na parametry LD(b)	Vliv na rozptyl LD(s)	Celkový vliv LD(b,s)
1	0,1841046591	0,05154610927	0,2163608025
2	0,3974953858	0,002462512712	0,4224052865
3	1,242375805	<b>12,72551054</b>	<b>18,52954421</b>
4	0,1129743734	0,009101043711	0,1172826386
5	0,1052470145	0,02738535559	0,1243220456
6	0,253548185	0,006682430504	0,2537602745
7	0,09807052304	0,06389820776	0,1499087971

### c. Graf vlivných bodů

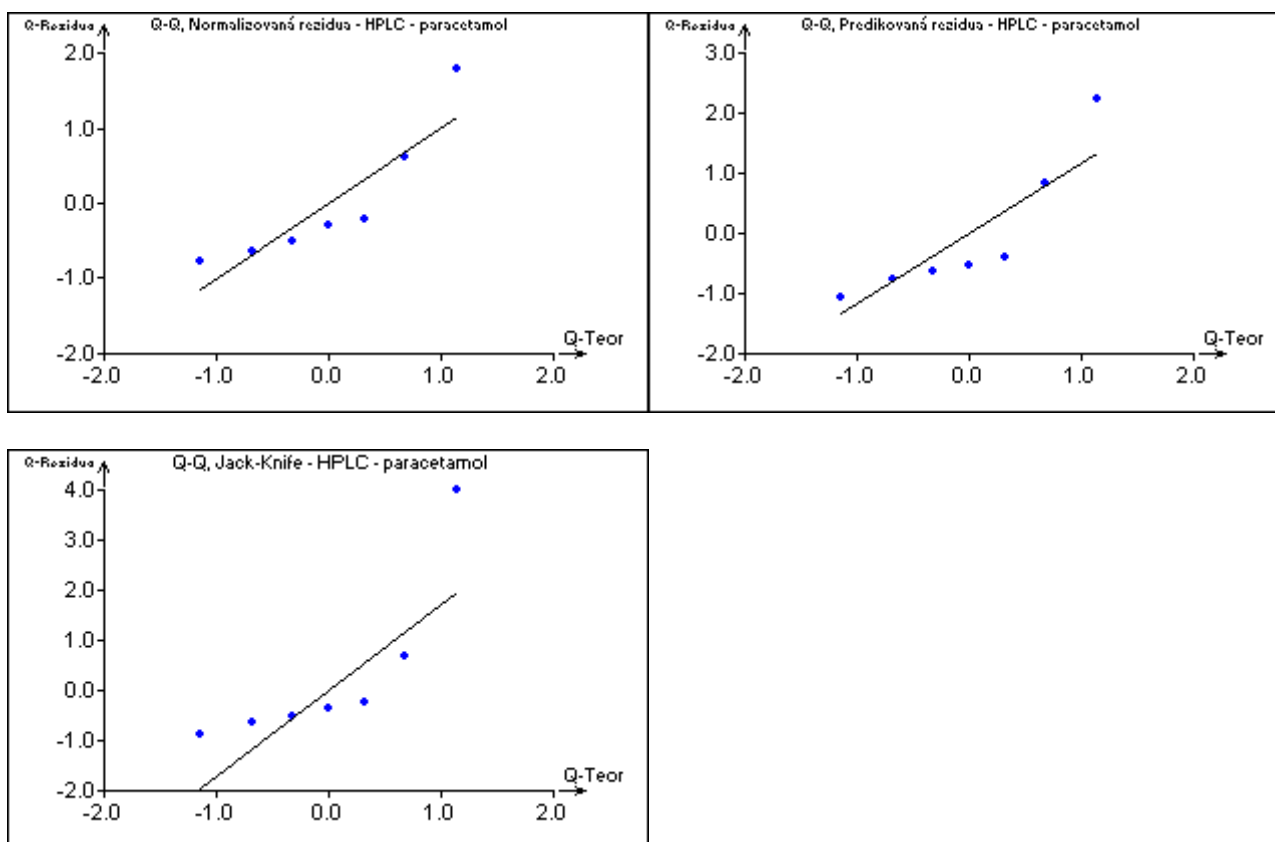


Grafy indikovaly bod č.3 jako odlehlý a proto je vhodné ho z výběru vyloučit.

### d. Indexové grafy



## e. Rankitové grafy



## 6. Testování regresního tripletu

### Fisher-Snedecorův test významnosti modelu

Hodnota kritéria F : 43993,53894  
Kvantil F (1-alfa, m-1, n-m) : 6,607890974  
Pravděpodobnost : 4,67439976E-011

**Závěr : Model je významný**

### Scottovo kritérium multikolinearity

Hodnota kritéria SC : 0,333290115

**Závěr : Model vykazuje multikolinearitu!**

### Cook-Weisbergův test heteroskedasticity

Hodnota kritéria CW : 0,3032009713

Kvantil  $\chi^2(1-\text{alfa}, 1)$  : 3,841458829

Pravděpodobnost : 0,5818826234

**Závěr : Rezidua vykazují homoskedasticitu.**

### Jarque-Berrův test normality

Hodnota kritéria JB : 1,7028749

Kvantil  $\chi^2(1-\text{alfa}, 2)$  : 5,991464547

Pravděpodobnost : 0,4268009858

**Závěr : Rezidua mají normální rozdělení.**

### Waldův test autokorelace

Hodnota kritéria WA : 1,407236262  
Kvantil  $\chi^2(1-\alpha,1)$  : 3,841458829  
Pravděpodobnost : 0,2355157318

**Závěr : Autokorelace je nevýznamná**

### Durbin-Watsonův test autokorelace

Hodnota kritéria DW : -1  
Kritické hodnoty DW 0 2

**Závěr : Negativní autokorelace reziduí není prokázána.**

### Znaménkový test reziduí

Hodnota kritéria Sg : 1,746975408  
Kvantil  $N(1-\alpha/2)$  : 1,959963999  
Pravděpodobnost : 0,08064160394

**Závěr : V reziduích není trend.**

## 7. Konstrukce zpřesněného modelu

Po odstranění bodu č.3 byly nalezeny nové parametry:

Proměnná	Odhad	Směr.Odch.	Závěr	Pravděp.	Spodní mez	Horní mez
$\beta_0$	0,6578	0,2669	Nevýzn.	0,0693	-0,0829	1,3989
$\beta_1$	0,9961	0,00245	Význam.	2, 2E-010	0,9893	1,0029

$$y = 0,9961x$$

### Statistické charakteristiky regrese

Vícenásobný korelační koeficient R : 0,99999  
Koeficient determinace  $R^2$  : 0,99998  
Predikovaný korelační koeficient  $R_p$  : 0,99989  
Střední kvadratická chyba predikce MEP : 0,06409  
Akaikého informační kritérium : -17,3391

R,  $R^2$ ,  $R_p$  dosáhly vysokých hodnot. MEP je vyšší a AIC má nižší hodnotu než u předešlého modelu, což dokazuje, že zpřesněný model je lepší.

## 8. Zhodnocení kvality modelu

$$y = 0,9961x$$

## II. Řešení dat získaných při UV stanovení

### 1. Návrh modelu

$$Y = \beta_0 + \beta_1 x$$

### 2. Předběžná analýza dat

#### Charakteristiky proměnných

Proměnná	Průměr	Směr.Odch.	Kor.vs.Y	Významnost
y	100,33	35,590	1,0000	-----
x	100	35,066	0,99963	4,90E-009

#### Indikace multikolinearity

Proměnná	Vlas. čísla kor. m.	Podmíněnost kappa	VI faktor	Vícenás. kor. $\beta_0$
	1	1	1	0
B <sub>1</sub>	1	1	1	0

### 3. Odhadování parametrů

Test:  $H_0: \beta_j$  vs.  $H_1: \beta_j \neq 0$

Proměnná	Odhad	Směr.Odch.	Závěr	Pravděp.	Spodní mez	Horní mez
$\beta_0$	1,8357	1,2538	Nevýzn.	0,2030	-1,3872	5,0586
$\beta_1$	0,9849	0,0119	Význ.	4,9E-009	0,9543	1,0155

Úsek  $\beta_0$  je statisticky nevýznamný, směrnice  $\beta_1$  je statisticky významná.

### 4. Základní statistické charakteristiky

Vícenásobný korelační koeficient R :	0,9996
Koeficient determinace $R^2$ :	0,9993
Predikovaný korelační koeficient $R_p$ :	0,9967
Střední kvadratická chyba predikce MEP :	1,7340
Akaikeho informační kritérium :	2,1681

### 5. Regresní diagnostika

#### a. Analýza klasických reziduí

Index	Y naměřené	Y vypočítané	Směr.odch.Y	Reziduum	Reziduum [%Y]
1	49,781	51,0817	0,7130	-1,3007	-2,6128
2	71,353	70,7800	0,5306	0,57298	0,8031
3	81,727	80,6292	0,4590	1,09780	1,3433
4	101,23	100,3276	0,3924	0,90243	0,8915
5	119,382	120,0259	0,4590	-0,64394	-0,5394
6	128,954	129,8751	0,5306	-0,92112	-0,7143
7	149,866	149,5735	0,7130	0,29251	0,1952

Reziduální součet čtverců :	5,388223135
Průměr absolutních reziduí :	0,8187755102
Reziduální směr. odchylka :	1,038096637
Reziduální rozptyl :	1,077644627
Šikmost reziduí :	0,04278048289
Špičatost reziduí :	1,442867462



## b. Analýza ostatních reziduí

### A. Analýza reziduí

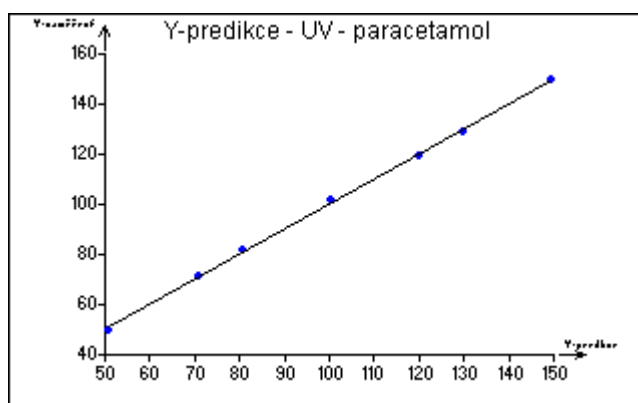
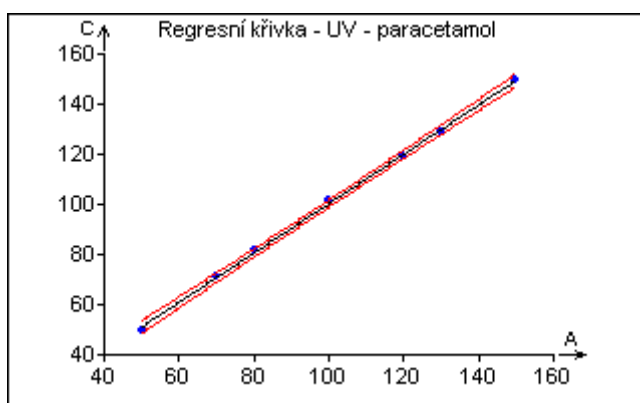
Bod	Standardní	Jackknife	Predikované	Diag(H <sub>ii</sub> )	Diag(H* <sub>ii</sub> )	Cook. vzdál.
1	-1,72395	-2,4212	-2,4624	0,4718	0,78577	-0,7700
2	0,64219	0,5997	0,77564	0,2613	0,32221	0,1136
3	1,17901	1,24108	1,36455	0,1955	0,41916	0,1432
4	0,93897	0,92538	1,05284	0,1426	0,29400	0,0782
5	-0,69158	-0,65046	-0,80041	0,1955	0,27246	-0,0840
6	-1,03238	-1,04098	-1,24692	0,2613	0,41875	-0,1826
7	0,38771	0,352107	0,55379	0,4718	0,48768	0,1732

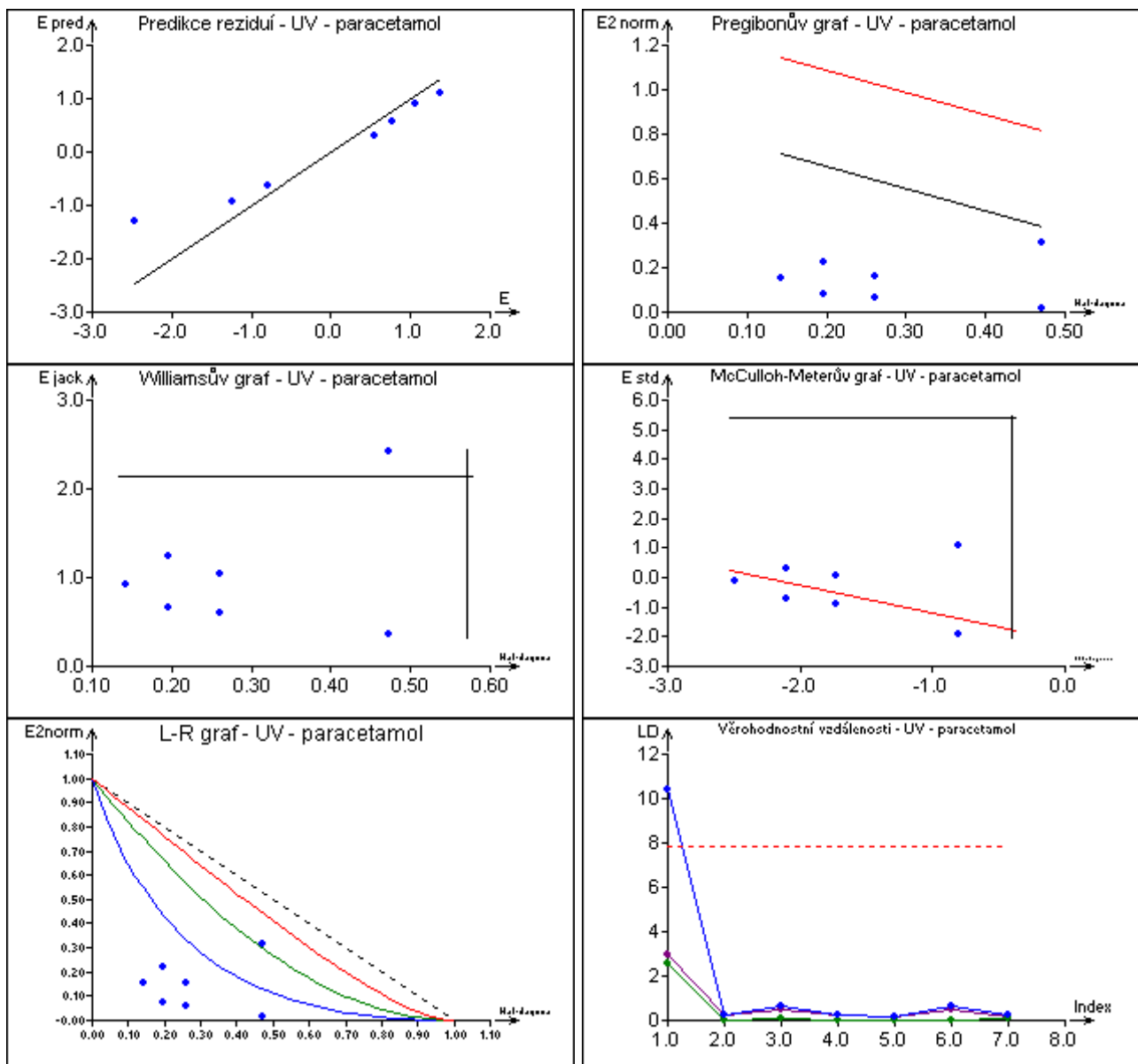
### B. Analýza vlivu

Index	Atkinsonova vzdál.	Andrews-Pregibon st.	Vliv na Y <sup>^</sup>
1	<b>3,618067736</b>	0,2142345859	-2,288266955
2	0,5638722692	0,6777912434	0,356624136
3	0,9673051968	0,5808460352	0,6117775229
4	0,5973257945	0,7060026192	0,3777820031
5	0,5069725412	0,7275548272	-0,3206375883
6	0,9788662165	0,5812543959	-0,6190893538
7	0,5261733703	0,5123162932	0,3327812588

Index	Vliv na parametry LD(b)	Vliv na rozptyl LD(s)	Celkový vliv LD(b,s)
1	2,981189344	2,555294103	<b>10,40954259</b>
2	0,201286876	0,01585490696	0,2066260685
3	0,4575953144	0,1092181431	0,6706257493
4	0,2027537375	0,00562837407	0,2197072369
5	0,1608427327	0,009879932681	0,1640926272
6	0,508804349	0,02638231751	0,6012897594
7	0,1854950372	0,05135434524	0,2174693946

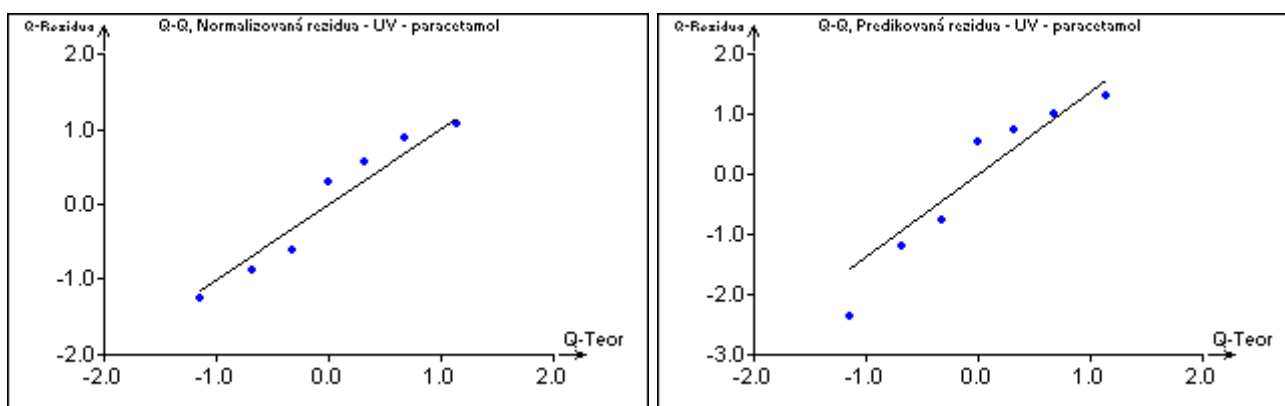
## c. Grafy vlivných bodů

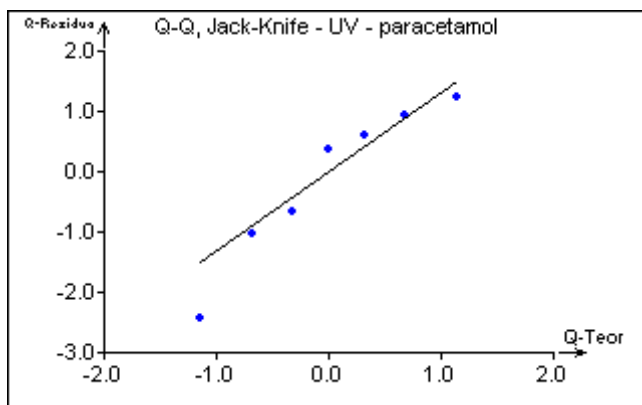




Bod č.1 byl indikován jako vlivný ve 3 grafech ( Williamsův, McCulloch-Meterův a L-R graf) a proto bude z dat vyloučen.

#### d. Rankitové grafy





## 6. Testování regresního tripletu

### Fisher-Snedecorův test významnosti modelu

Hodnota kritéria F : 6841,298574

Kvantil F (1-alfa, m-1, n-m) : 6,607890974

Pravděpodobnost : 4,895280521E-009

**Závěr : Model je významný**

### Scottovo kritérium multikolinearity

Hodnota kritéria SC : 0,3331940737

**Závěr : Model vykazuje multikolinearitu!**

### Cook-Weisbergův test heteroskedasticity

Hodnota kritéria CW : 0,3672171372

Kvantil  $\chi^2(1-\text{alfa}, 1)$  : 3,841458829

Pravděpodobnost : 0,5445250881

**Závěr : Rezidua vykazují homoskedasticitu.**

### Jarque-Berrův test normality

Hodnota kritéria JB : 0,7571035707

Kvantil  $\chi^2(1-\text{alfa}, 2)$  : 5,991464547

Pravděpodobnost : 0,6848525048

**Závěr : Rezidua mají normální rozdělení.**

### Waldův test autokorelace

Hodnota kritéria WA : 0,0643377585

Kvantil  $\chi^2(1-\text{alfa}, 1)$  : 3,841458829

Pravděpodobnost : 0,7997668241

**Závěr : Autokorelace je nevýznamná**

### Durbin-Watsonův test autokorelace

Hodnota kritéria DW : -1

Kritické hodnoty DW 0 2

**Závěr : Pozitivní autokorelace reziduí není prokázána.**

### Znaménkový test reziduí

Hodnota kritéria Sg : 0,06063390626

Kvantil  $N(1-\text{alfa}/2)$  : 1,959963999

Pravděpodobnost : 0,9516507699

**Závěr : V reziduích není trend.**

## 7. Konstrukce zpřesněného modelu

Po odstranění bodu č.1 byly nalezeny nové parametry:

Proměnná	Odhad	Směr.Odch.	Závěr	Pravděp.	Spodní mez	Horní mez
$\beta_0$	3,8075	1,2084	Význam.	0,0344	0,4525	7,1625
$\beta_1$	0,9687	0,0108	Význam.	9,3E-008	0,9887	0,9987

### Statistické charakteristiky regrese

Vícenásobný korelační koeficient R	: 0,99975
Koeficient determinace R <sup>2</sup>	: 0,99950
Predikovaný korelační koeficient Rp	: 0,99745
Střední kvadratická chyba predikce MEP	: 0,93364
Akaikeho informační kritérium	: -2,05963

R, R<sup>2</sup>, Rp dosáhly nízkých hodnot. MEP a AIC mají vyšší hodnotu než u předešlého modelu, což dokazuje, že zpřesněný model je horší a nebylo nutné bod č.1 eliminovat. Z tohoto důvodu zůstává v platnosti předchodzí model.

## 8. Zhodnocení kvality modelu

$$y = 0,985x$$

### III. Řešení dat získaných HPLC a UV stanovení

#### 1. Návrh modelu

$$Y = \beta_0 + \beta_1 x$$

#### 2. Předběžná analýza dat

##### Charakteristiky proměnných

Proměnná	Průměr	Směr.Odch.	Kor.vs.Y	Významnost
y	100,3276	35,37	1,0000	-----
x	100,400	35,367	0,99963	4,95E-009

##### Indikace multikolinearity

Proměnná	Vlas. čísla kor. m.	Podmíněnost kappa	VI faktor	Vícenás. kor.
$\beta_0$	1	1	1	0
$B_1$	1	1	1	0

#### 3. Odhadování parametrů

Test:  $H_0: \beta_j$  vs.  $H_1: \beta_j \neq 0$

Proměnná	Odhad	Směr.Odch.	Závěr	Pravděp.	Spodní mez	Horní mez
$\beta_0$	0,8189	1,2683	Nevýzn.	0,5469	-2,4414	4,0792
$\beta_1$	0,9911	0,0120	Význ.	4,9E-009	0,9602	1,0220

Úsek  $\beta_0$  je statisticky nevýznamný, směrnice  $\beta_1$  je statisticky významná.

#### 4. Základní statistické charakteristiky

Vícenásobný korelační koeficient R :	0,9996330992
Koeficient determinace $R^2$ :	0,9992663331
Predikovaný korelační koeficient $R_p$ :	0,9967608517
Střední kvadratická chyba predikce MEP :	1,708387227
Akaikeho informační kritérium :	2,200129033

#### 5. Regresní diagnostika

##### a. Analýza klasických reziduí

Index	Y naměřené	Y vypočítané	Směr.odch.Y	Reziduum	Reziduum [%Y]
1	49,781	50,9696	0,7158139948	-1,18863	-2,38771
2	71,353	70,4620	0,5344433683	0,891005	1,248728
3	81,727	81,36432	0,4554768673	0,362679	0,44377
4	101,23	100,0628	0,3932743228	1,167199	1,153017
5	119,382	119,8198	0,4587441742	-0,437797	-0,36672
6	128,954	130,1275	0,5339045972	-1,173450	-0,8
7	149,866	149,4870	0,7138053	0,3789925	0,252888

Reziduální součet čtverců :	5,41290396
Průměr absolutních reziduí :	0,7999644473
Reziduální směr. odchylka :	1,040471428
Reziduální rozptyl :	1,082580792
Šikmost reziduí :	0,04234054266
Špičatost reziduí :	1,541729669

## b. Analýza ostatních reziduí

### A. Analýza reziduí

Bod	Standardní	Jackknife	Predikované	Diag(Hii)	Diag(H*ii)	Cook. vzdál.
1	-1,57411	-1,98234	-2,25676	0,473303	0,734318	-0,7073
2	0,99808	0,997598	1,210343	0,263841	0,410508	0,1789
3	0,38769	0,352096	0,448657	0,191634	0,215934	0,0460
4	1,21169	1,289501	1,361748	0,142867	0,394553	0,10098
5	-0,46879	-0,42883	-0,54344	0,194393	0,229802	-0,05656
6	-1,31399	-1,45251	-1,59287	0,263310	0,517699	-0,23482
7	0,5006	0,45945	0,715960	0,47065	0,497187	0,2226

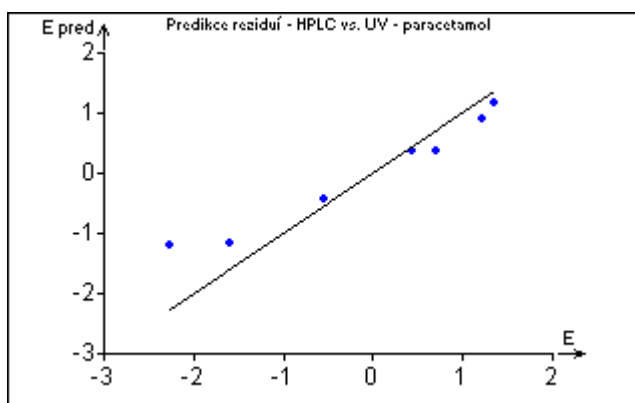
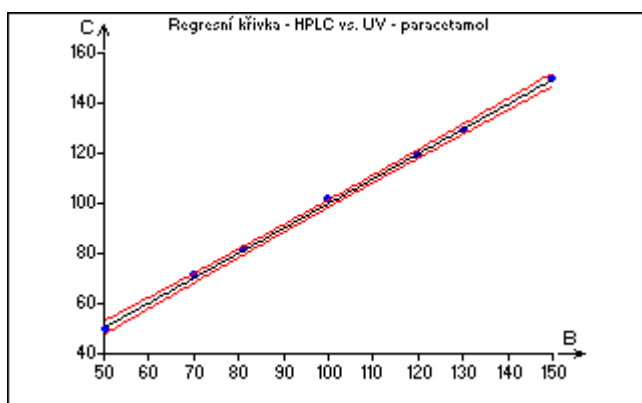
### B. Analýza vlivu

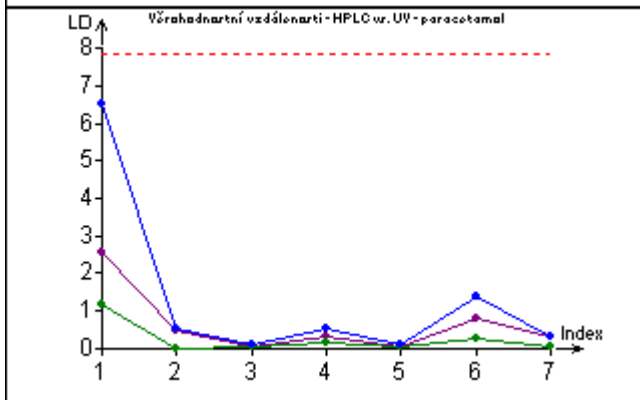
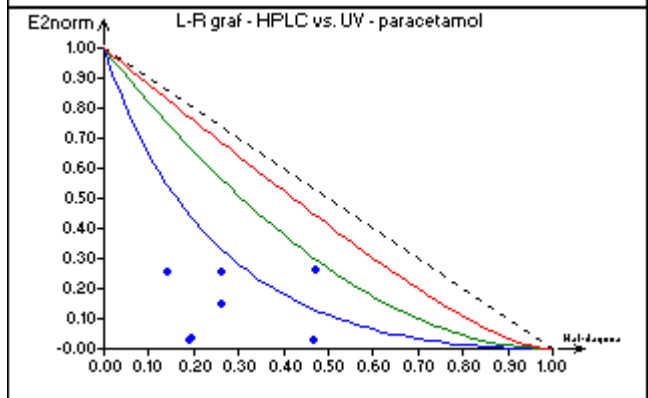
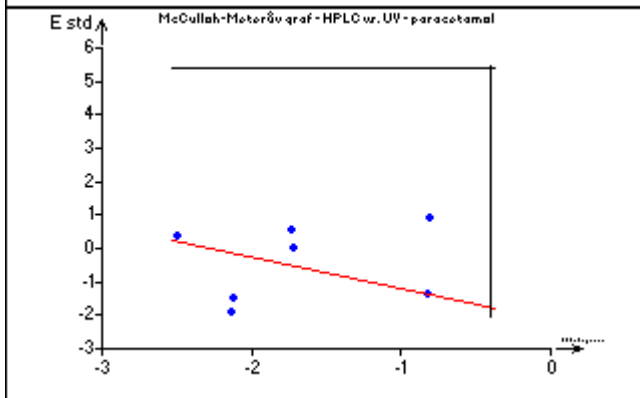
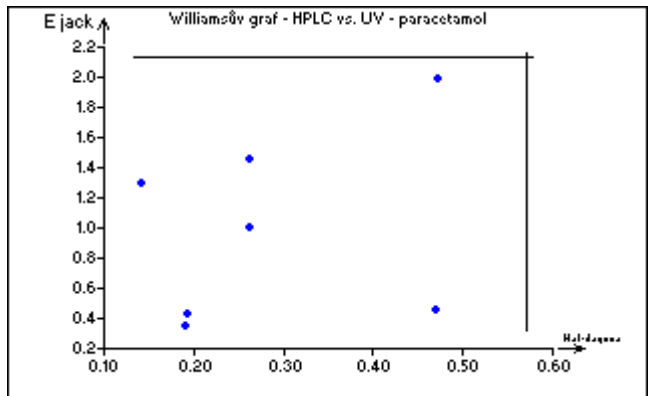
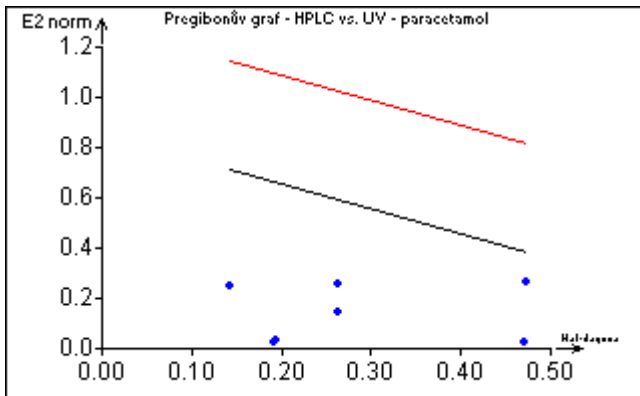
Index	Atkinsonova vzdál.	Andrews-Pregibon st.	Vliv na Y <sup>^</sup>
1	<b>2,971245847</b>	0,2656834064	-1,879180873
2	0,9443034127	0,5894924565	0,5972299173
3	0,271058578	0,7840656423	0,1714324972
4	0,8324018954	0,6054469514	0,5264571836
5	0,3330698589	0,770197783	-0,2106518748
6	1,373035203	0,4823007265	-0,86838371
7	0,6850005168	0,5028130338	0,4332323663

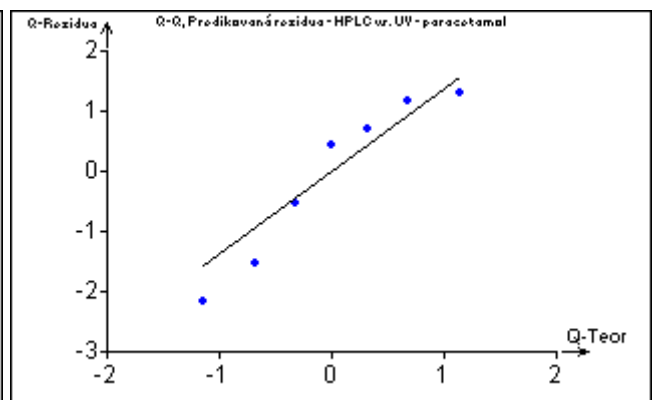
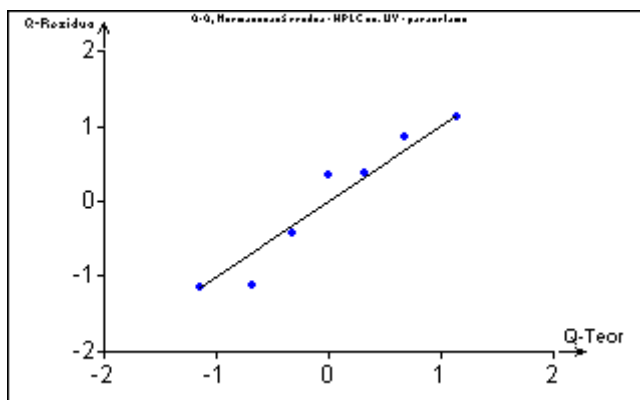
Index	Vliv na parametry LD(b)	Vliv na rozptyl LD(s)	Celkový vliv LD(b,s)
1	2,57836132	1,183346835	6,480327967
2	0,4827976044	0,01657335783	0,5515987189
3	0,04970794131	0,05135593737	0,095439589
4	0,3344833692	0,1398896594	0,5556254086
5	0,0738507386	0,04025914234	0,1068204609
6	0,8146623863	0,2785569141	1,409692316
7	0,305239828	0,03569830851	0,3172337333

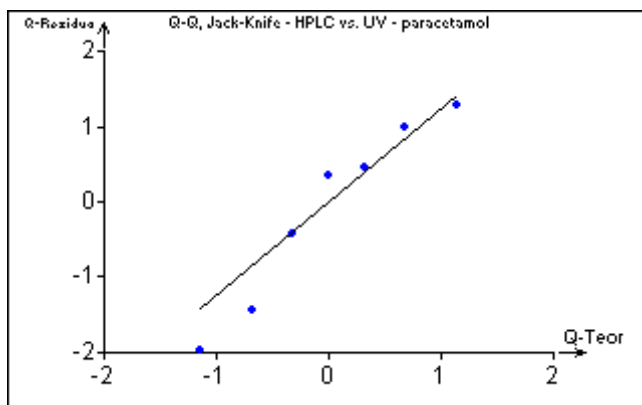
## c. Grafy vlivných bodů





#### d. Rankitové grafy





## 7. Testování regresního tripletu

### Fisher-Snedecorův test významnosti modelu

Hodnota kritéria F : 6810,082004  
 Kvantil F (1-alfa, m-1, n-m) : 6,607890974  
 Pravděpodobnost : 4,951536402E-009

**Závěr : Model je významný**

### Scottovo kritérium multikolinearity

Hodnota kritéria SC : 0,3333061262

**Závěr : Model vykazuje multikolinearitu!**

### Cook-Weisbergův test heteroskedasticity

Hodnota kritéria CW : 0,1144314895  
 Kvantil  $\chi^2(1-\text{alfa}, 1)$  : 3,841458829  
 Pravděpodobnost : 0,7351542754

**Závěr : Rezidua vykazují homoskedasticitu.**

### Jarque-Berrův test normality

Hodnota kritéria JB : 0,6696417374  
 Kvantil  $\chi^2(1-\text{alfa}, 2)$  : 5,991464547  
 Pravděpodobnost : 0,7154662373

**Závěr : Rezidua mají normální rozdělení.**

### Waldův test autokorelace

Hodnota kritéria WA : 0,09576515706  
 Kvantil  $\chi^2(1-\text{alfa}, 1)$  : 3,841458829  
 Pravděpodobnost : 0,7569720463

**Závěr : Autokorelace je nevýznamná**

### Durbin-Watsonův test autokorelace

Hodnota kritéria DW : -1  
 Kritické hodnoty DW : 0

**Závěr : Pozitivní autokorelace reziduí není prokázána.**



## Znaménkový test reziduí

Hodnota kritéria Sg : 0,06063390626

Kvantil N(1-alfa/2) : 1,959963999

Pravděpodobnost : 0,9516507699

**Závěr : V reziduích není trend.**

## Porovnání regresních přímek

### 1. Testování úseku a směrnice

Metodou nejmenších čtverců byly určeny odhady parametrů úseků a směrníc a rovněž jejich 95%ní intervaly spolehlivosti pro oba dva modely regresních přímek viz tab. 1.

Tabulka č.1

95%ní intervaly	Úsek		Směrnice	
	L1	L2	L1	L2
HPLC	-0,0829	1,3980	0,9893	1,0029
UV	-1,3872	5,0586	0,9543	1,0155

Intervaly spolehlivosti úseků obou regresních přímek zahrnují nulu, tudíž úseky lze považovat za nulové.

Intervaly spolehlivosti směrníc obou regresních přímek zahrnují jednotku, tudíž směrnice lze považovat za jednotkové.

### 2. Identifikace vlivných bodů a extrémů

a) Williamsův graf

HPLC: graf identifikuje jeden odlehlý a souběžně vlivný bod č. 3.

UV: graf identifikuje jeden vlivný bod č. 1.

b) L - R graf

HPLC: graf identifikuje jeden vlivný bod č. 3.

UV: graf identifikuje jeden vlivný bod č. 1.

c) McCulloh-Meterův graf

HPLC: graf identifikuje jeden odlehlý a souběžně vlivný bod č. 2 a 3.

UV: graf identifikuje jeden vlivný bod č. 1, 3 a 6.

Testováním regresního tripletu byly testy přijatelné a proto lze potvrdit, že navržené modely regresních přímek jsou správné. Vzhledem k nízkému počtu dat nebudeme vlivné body ze souborů vylučovat. Pro další výpočet homoskedasticit reziduí a test shodnosti modelů byly modely znovu propočteny jenom s významným z parametrů (směrníc).

### 3. Test shodnosti dvou přímek

a) Test shody rozptylů

$$\text{Testační statistika } F = \frac{\max [s_1^2, s_2^2]}{\min [s_1^2, s_2^2]} = \frac{1,07764}{0,17058} = 6,31$$

Tuto statistiku porovnáme s tabulkovou hodnotou  $F_{0.95}(5,5) = 5,05$

**Jelikož je testační statistika vyšší než tabelovaná hodnota F-rozdělení, reziduální rozptyly nejsou shodné.**

#### b) Chowův test

Pro účely Chowova testu vypočítáme ještě souhrnný jednoparametrový regresní model. Dále jsou uvedené údaje nezbytně potřebné pro výpočet Chowova testu.

RSC= 9,36999

$$F_c = \frac{(RSC - RSC_1 - RSC_2)(n-2*m)}{(RSC_1 + RSC_2)*m} = \frac{(6,55153 - 0,85289 - 5,38822)(14 - 2*2)}{(0,85289 + 5,38822)*2} = 0,25$$

$$r = \frac{((n_1-m)*s_1^2 + (n_2-m)*s_2^2)^2}{(n_1-m)*(s_1^2)^2 + (n_2-m)*(s_2^2)^2} = \frac{((7-2)*0,17058 + (7-2)*1,07764)^2}{(7-2)*(0,17058)^2 + (7-2)*(1,07764)^2} = 6.54$$

r=7

Hodnotu testačního kritéria  $F_c = 0,25$  pro neshodné reziduální rozptyly porovnáme s tabulkovou hodnotou  $F_{0.95}(m,r) = F_{0.95}(1,7)=5.59$  Jelikož je hodnota testačního kritéria  $F_c$  nižší než tabulková hodnota F-rozdělení, jsou lineární závislosti shodné.

#### **Závěr:**

Výslední rovnice lineárních regresních přímk z měření oběma metodami jsou ve tvaru:

HPLC:  **$Y = 0,994 (\pm 0,005).x$**

UV:  **$Y = 0,985(\pm 0,012).x$**

Pomocí grafů vlivných bodů byl nalezen jeden vlivný bod č.3 u HPLC stanovení a bod č.1 u UV stanovení. Vzhledem k malému počtu dat tyto body nebyly vyloučeny.

Oba lineární modely vykazují vysoké hodnoty vícenásobného korelačního koeficientu, koeficientu determinace a predikovaného korelačního koeficientu.

Koeficient determinace dat obou metod ukazuje, že model vystihuje data v 99,93 %.

Ukázalo se, že 95 % interval spolehlivosti úseků obou modelů obsahují nulu, proto byly úseky v obou modelech zanedbány jako statisticky nevýznamné.

Testováním shody reziduálních rozptylů obou regresních přímk bylo zjištěno, že tyto rozptyly nejsou shodné. Na základě testování dle Chowa lze konstatovat, že stanovené obsahy paracetamolu pomocí HPLC a UV je možné považovat za shodné.

## Příklad 2 – Určení stupně polynomu metodou MNČ a RH křivkové závislosti

### Závislost HPLC signálu na koncentraci guaifenesinu

**Zadání:** Využitím střední kvadratické chyby predikce MEP určete stupeň polynomu n experimentální závislosti signálu HPLC y na koncentraci x léčivé látky guaifenesin v přípravku Ataralgin.

**Data:**

Guaifenesin	
c (%)	A1
50	815138
70	1091326
80	1244022
100	1585449
120	1863204
130	2011784
150	2353643

### Vyšetření stupně polynomu metodou nejmenších čtverců

Pro nalezení optimálního stupně polynomu n byl při výpočtě měněn stupeň polynomu a sledovaná hodnota střední kvadratické chyby predikce MEP, která musí být pro optimální n minimální.

### I. stupeň polynomu 2

#### (1) ZVOLENA STRATEGIE REGRESNÍ ANALÝZY

Omezení, P: 1E-34  
Transformace: Polynom  
Váhy: Ne  
Absolutní člen zahrnut: Ano

#### (2) PODMÍNKY A KVANTILY PRO STATISTICKÉ TESTY

Hladina významnosti : 0,05  
Kvantil t(1-alfa/2,n-m) : 2,77644510519769  
Kvantil F(1-alfa,m,n-m) : 6,94427190984624  
Absolutní člen : Ano  
Počet platných řádků : 7  
Počet parametrů : 3  
Metoda : Nejmenší čtverce  
Transformace : Polynom 2. stupni

## Základní analýza

### (3) CHARAKTERISTIKY PROMĚNNÝCH

Proměnná	Průměr	Směr.Odch.	Kor.vs.Y	Významnost
F	100	35,590	0,99950	1,05624216E-008
F <sup>2</sup>	11085,71	7190,139	0,991736	1,187187073E-005

### Párové korelace (Xi, Xj)

Abs - F	0	1
Abs - F <sup>2</sup>	0	1
F - F <sup>2</sup>	<b>0,98997</b>	1,923E-005

### (4) INDIKACE MULTIKOLINEARITY

Proměnná	Vlas. čísla kor. m.	Podmíněnost kappa	VI faktor	Vícenás. kor.
Abs	0,01002575667	1	1	0
F	1	99,74309497	50,123	0,98997
F <sup>2</sup>	1,989974243	198,4861899	50,123	0,98997

Model vykazuje silnou multikolinearitu, protože VIF > 10.

### (5) ANALÝZA ROZPTYLU

Průměr Y :	1566366,571		
Zdroj	Součet čtverců	Průměrný čtverec	Rozptyl
Celková variabilita	1,800591393E+012	2,572273418E+011	3,000985654E+011
Var.vysv.modelem	1,799261113E+012	2,570373018E+011	2,998768521E+011
Reziduální variabilita	1330280000	190040000	221713333,3
Hodnota kritéria F :	2705,086317		
Kvantil F (1-alfa, m-1, n-m) :	6,94427191		
Pravděpodobnost :	5,458279668E-007		

**Závěr : Model je významný**

### (6) ODHADY PARAMETRŮ

Proměnná	Odhad	Směr.Odch.	Závěr	Pravděp.	Spodní mez	Horní mez
Abs	104623,10	68959,95	Nevýzn.	0,203827	-86840,42	296086,61
F	13663,45	1480,991	Významný	0,000767	9551,5596	17775,34
F <sup>2</sup>	8,6055332	7,3307149	Nevýznamný	0,3055576	-11,747794	28,95886

### (7) STATISTICKÉ CHARAKTERISTIKY REGRESE

Vícenásobný korelační koeficient R :	0,9996305309
Koeficient determinace R <sup>2</sup> :	0,9992611983
Predikovaný korelační koeficient Rp :	0,9952346828
Střední kvadratická chyba predikce MEP :	613616826,6
Akaikeho informační kritérium :	139,4392159

## (8) ANALÝZA KLASICKÝCH REZIDUÍ

Index	Y naměřené	Y vypočítané	Směr. odch. Y	Reziduum	Reziduum [%Y]
1	815138	809309,4252	16260,3123	5828,574792	0,71504147671
2	1091326	1103231,704	9420,549334	-11905,7036	-1,090939243
3	1244022	1252774,503	9501,692515	-8752,50277	-0,7035649506
4	1585449	1557023,421	10528,84926	28425,57895	1,792904026
5	1863204	1868156,766	9501,692515	-4952,765928	-0,265819842
6	2011784	2026305,098	9420,549334	-14521,09834	-0,7218020592
7	2353643	2347765,083	16260,3123	5877,916897	0,2497369778

Reziduální součet čtverců : 1330280000  
Průměr absolutních reziduí : 11466,3059  
Reziduální směr. odchylka : 18236,50186  
Reziduální rozptyl : 332570000  
Šikmost reziduí : 0,9450730666  
Špičatost reziduí : 2,872785949

## (9) TESTOVÁNÍ REGRESNÍHO TRIPLETU

### Fisher-Snedecorův test významnosti modelu

Hodnota kritéria F : 2705,086317  
Kvantil F (1-alfa, m-1, n-m) : 6,94427191  
Pravděpodobnost : 5,458279668E-007  
**Závěr : Model je významný**

### Scottovo kritérium multikolinearity

Hodnota kritéria SC : 0,9783529686  
**Závěr : Model je nekorektní!**

### Cook-Weisbergův test heteroskedasticity

Hodnota kritéria CW : 0,0001791848938  
Kvantil  $\chi^2(1-\alpha, 1)$  : 3,841458829  
Pravděpodobnost : 0,9893198393  
**Závěr : Rezidua vykazují homoskedasticitu.**

### Jarque-Berrův test normality

Hodnota kritéria JB : 1,107305407  
Kvantil  $\chi^2(1-\alpha, 2)$  : 5,991464547  
Pravděpodobnost : 0,574846228  
**Závěr : Rezidua mají normální rozdělení.**

### Waldův test autokorelace

Hodnota kritéria WA : 0,2467453462

Kvantil  $\chi^2(1-\alpha,1)$  : 3,841458829

Pravděpodobnost : 0,6193761604

**Závěr : Autokorelace je nevýznamná**

### Durbin-Watsonův test autokorelace

Hodnota kritéria DW : -1

Kritické hodnoty DW 0 2

**Závěr : Negativní autokorelace reziduí není prokázána.**

### Znaménkový test reziduí

Hodnota kritéria Sg : 0,9095085939

Kvantil  $N(1-\alpha/2)$  : 1,959963999

Pravděpodobnost : 0,3630817236

**Závěr : V reziduích není trend.**

## (10) INDIKACE VLIVNÝCH DAT

### A. Analýza reziduí

Index	Standardní	Jackknife	Predikované	Diag(Hii)	Diag(H*ii)	Cook.vzdál.
1	0,70592503	0,6534037	28433,99324	0,7950139	0,82055159	0,91261479
2	-0,7624607786	-0,7142505596	-16239,1398	0,2668513389	0,37340468	-0,09250679
3	-0,562297883	-0,5074319524	-12013,89164	0,271468144	0,3290547459	-0,06984182
4	1,9090329	5,544944499	42638,36842	0,3333333333	0,9407344358	0,31817215
5	-0,3181866798	-0,2791126533	-6798,28327	0,271468144	0,2899077886	-0,03952129
6	-0,9299549456	-0,909684998	-19806,48552	0,2668513389	0,4253610488	-0,11282829
7	0,711901073	0,6597338144	28674,7027	0,7950138504	0,8209857564	0,92034058

### B. Analýza vlivu

Index	Atkinson.vzdál.	Andrews-Pregibon st.	Vliv na $Y^{\wedge}$	Vliv na param.LD(b)	Vliv na rozptyl LD(s)	Celkový vliv LD(b,s)
1	1,485853751	0,1794484552	1,286787095	2,759310594	0,001546788597	3,313188843
2	0,4975750872	0,6265953203	-0,4309126658	0,3608351555	2,940055445E-005	0,3714008517
3	0,3576698583	0,6709452541	-0,3097511835	0,2031994681	0,0176224772	0,2095140685
4	<b>4,527428225</b>	0,05926556424	<b>3,920867857</b>	2,627690829	<b>44,63004159</b>	<b>75,3764511</b>
5	0,1967360997	0,7100922114	-0,1703784602	0,06571013729	0,05540689331	0,1134645327
6	0,6337224187	0,5746389512	-0,5488197136	0,5302568157	0,02886242532	0,6312685813
7	1,500248499	0,1790142436	1,299253312	2,797976814	0,001212905647	3,377331245

## II. stupeň polynomu 3

Hladina významnosti :	0,05
Kvantil $t(1-\alpha/2, n-m)$ :	3,18244630528356
Kvantil $F(1-\alpha, m, n-m)$ :	9,27662815314474
Absolutní člen :	Ano
Počet platných řádků :	7
Počet parametrů :	4
Metoda :	Nejmenší čtverce
Sloupce pro výpočet :G	
Abs	
F	
F <sup>2</sup>	
F <sup>3</sup>	
Transformace :	Polynom 3. stupně

## ZÁKLADNÍ ANALÝZA

### Charakteristiky proměnných

Proměnná	Průměr	Směr.Odch.	Kor.vs.Y	Významnost
F	100	35,59026084	0,9995032131	1,05624216E-008
F <sup>2</sup>	11085,71429	7190,13875	0,991735886	1,187187073E-005
F <sup>3</sup>	1325714,286	1175883,174	0,9691572712	0,0003155891109

### Párové korelace (Xi, Xj)

Abs - F	0	1
Abs - F <sup>2</sup>	0	1
Abs - F <sup>3</sup>	0	1
F - F <sup>2</sup>	<b>0,9899742433</b>	1,922693999E-005
F - F <sup>3</sup>	<b>0,9655120051</b>	0,0004164397994
F <sup>2</sup> - F <sup>3</sup>	<b>0,9924301526</b>	9,536921139E-006

### Indikace multikolinearity

Proměnná	Vlas. čísla kor. m.	Podmíněnost kappa	VI faktor	Vícenás. kor.
Abs	0,0001267682826	1	1	0
F	1	<b>7888,408517</b>	<b>1161,42577</b>	0,9995694019
F <sup>2</sup>	0,03454582067	272,511546	<b>5219,932331</b>	0,9999042087
F <sup>3</sup>	2,965327411	<b>23391,71401</b>	<b>1536,334709</b>	0,9996744971

## Analýza rozptylu

Průměr Y : 1566366,571

Zdroj	Součet čtverců	Průměrný čtverec	Rozptyl
Celková variabilita	1,800591393E+012	2,572273418E+011	3,000985654E+011
Variab.vysv.modelem	1,799261327E+012	2,570373325E+011	2,998768879E+011
Reziduální variabilita	1330065296	190009328	221677549,3
Hodnota kritéria F :	1352,761652		
Kvantil F (1-alfa, m-1, n-m) :	9,276628153		
Pravděpodobnost :	3,407526786E-005		

**Závěr : Model je významný**

## Odhady parametrů

Proměnná	Odhad	Směr.Odch.	Závěr	Pravděpodobnost	Spodní mez	Horní mez
Abs	109730,5956	245371,6649	Nevýznamný	0,6850454074	-671151,5529	890612,744
F	13486,26382	8231,230936	Nevýznamný	0,1998603545	-12709,18666	39681,7143
F <sup>2</sup>	10,49719991	86,37645852	Nevýznamný	0,9109556933	-264,3912414	285,3856412
F <sup>3</sup>	-0,006305555520,2865358547		Nevýznamný	0,9838248938	-0,9181905277	0,905579417

## Statistické charakteristiky regrese

Vícenásobný korelační koeficient R :	0,9996305905
Koeficient determinace R <sup>2</sup> :	0,9992613175
Predikovaný korelační koeficient Rp :	0,9452740666
Střední kvadratická chyba predikce MEP :	7137529058
Akaikeho informační kritérium :	141,4380861

## Analýza klasických reziduí

Index	Y naměřené	Y vypočítané	Směr. odch. Y	Reziduum	Reziduum [%Y]	Váhy
1	815138	809498,5919	20648,63096	5639,408126	0,6918347723	1
2	1091326	1103042,537	13863,70867	-11716,53693	-1,073605589	1
3	1244022	1252585,336	13937,33483	-8563,336101	-0,6883588957	1
4	1585449	1557023,421	12156,68676	28425,57895	1,792904026	1
5	1863204	1868345,933	13937,33483	-5141,932588	-0,2759726036	1
6	2011784	2026494,265	13863,70867	-14710,265	-0,7312049901	1
7	2353643	2347575,916	20648,63096	6067,083574	0,2577741643	1

Reziduální součet čtverců :	1330065296
Průměr absolutních reziduí :	11466,3059
Reziduální směr. odchylka :	21055,99911
Reziduální rozptyl :	443355098,6
Šikmost reziduí :	0,9442165943
Špičatost reziduí :	2,87664345



### **Testování regresního tripletu**

#### **Fisher-Snedecorův test významnosti modelu**

Hodnota kritéria F : 1352,761652

Kvantil F (1-alfa, m-1, n-m) : 9,276628153

Pravděpodobnost : 3,407526786E-005

**Závěr : Model je významný**

#### **Scottovo kritérium multikolinearity**

Hodnota kritéria SC : 0,9989288099

**Závěr : Model je nekorektní!**

#### **Cook-Weisbergův test heteroskedasticity**

Hodnota kritéria CW : 0,0004618465899

Kvantil  $\chi^2(1-\alpha, 1)$  : 3,841458829

Pravděpodobnost : 0,9828542888

**Závěr : Rezidua vykazují homoskedasticitu.**

#### **Jarque-Berrův test normality**

Hodnota kritéria JB : 1,106024271

Kvantil  $\chi^2(1-\alpha, 2)$  : 5,991464547

Pravděpodobnost : 0,575214574

**Závěr : Rezidua mají normální rozdělení.**

#### **Waldův test autokorelace**

Hodnota kritéria WA : 0,1372832521

Kvantil  $\chi^2(1-\alpha, 1)$  : 3,841458829

Pravděpodobnost : 0,7109969054

**Závěr : Autokorelace je nevýznamná**

#### **Durbin-Watsonův test autokorelace**

Hodnota kritéria DW : -1

Kritické hodnoty DW 0 2

**Závěr : Negativní autokorelace reziduí není prokázána.**

#### **Znaménkový test reziduí**

Hodnota kritéria Sg : 0,9095085939

Kvantil  $N(1-\alpha/2)$  : 1,959963999

Pravděpodobnost : 0,3630817236

**Závěr : V reziduích není trend.**

## Indikace vlivných dat

### A. Analýza reziduí

Index	Standardní	Jackknife	Predikované	Diag(H <sub>ii</sub> )	Diag(H <sup>*</sup> <sub>ii</sub> )	Cookova vzdál.
1	1,368194537	1,821796596	147168,1687	0,9616805171	0,9855913161	<b>8,584184402</b>
2	-0,7393164288	-0,6675136968	-20682,98207	0,4335180055	0,5367289066	-0,1414462361
3	-0,5425641238	-0,4664792564	-15240,90879	0,4381348107	0,4932679876	-0,1057710258
4	1,653404422	4,531543681	42638,36843	0,3333333333	0,9408324849	0,2066755527
5	-0,3257875338	-0,2708385719	-9151,541484	0,4381348107	0,4580131361	-0,0635111688
6	-0,928221423	-0,8976801166	-25967,75386	0,4335180055	0,5962106921	-0,177587595
7	1,471954223	2,280317824	158328,9521	0,9616805171	0,9893554765	<b>9,235182667</b>

### B. Analýza vlivu

Index	Atkinson.vzdál.	Andrews-Pregibon st.	Vliv na Y <sup>^</sup>	Vliv na par. LD(b)	Vliv na rozptyl LD(s)	Celk.vliv LD(b,s)
1	<b>7,903802995</b>	0,01440868386	<b>9,126525574</b>	<b>19,69098158</b>	3,188991554	<b>253,0693987</b>
2	0,5057097513	0,4632710934	-0,5839433221	0,9137053667	0,007848075033	1,030817485
3	0,3567391596	0,5067320124	-0,4119268997	0,5161141941	0,008905859111	0,5179571711
4	<b>2,774992441</b>	0,05916751508	<b>3,204285266</b>	2,628044472	<b>44,73029657</b>	<b>75,53262877</b>
5	0,2071233034	0,5419868639	-0,2391653899	0,190501495	0,04697409759	0,2185743488
6	0,6800843049	0,4037893079	-0,785293713	1,390736529	0,1266790488	1,976737696
7	<b>9,893081854</b>	0,01064452349	<b>11,42354694</b>	<b>20,65698253</b>	6,712212415	<b>398,2044867</b>

### III. stupeň polynomu 4

Hladina významnosti : 0,05  
Kvantil t(1-alfa/2,n-m) : 4,30265272972926  
Kvantil F(1-alfa,m,n-m) : 19,24679434481  
Absolutní èlen : Ano  
Poèet platných øádkù : 7  
Poèet parametrù : 5  
Metoda : Nejmenší ètverce  
Sloupce pro výpoèet : G

Abs

F

F<sup>2</sup>

F<sup>3</sup>

F<sup>4</sup>

Transformace : Polynom 4. stupně

## Základní analýza

Charakteristiky proměnných

Proměnná	Průměr	Směr.Odch.	Kor.vs.Y	Významnost
F	100	35,59026084	0,9995032131	1,05624216E-008
F <sup>2</sup>	11085,71429	7190,13875	0,991735886	1,187187073E-005
F <sup>3</sup>	1325714,286	1175883,174	0,9691572712	0,0003155891109
F <sup>4</sup>	167205714,3	181406225,6	0,9390719456	0,001702834286

Párové korelace (Xi, Xj)

Abs - F	0	1
Abs - F <sup>2</sup>	0	1
Abs - F <sup>3</sup>	0	1
Abs - F <sup>4</sup>	0	1
F - F <sup>2</sup>	<b>0,9899742433</b>	1,922693999E-005
F - F <sup>3</sup>	<b>0,9655120051</b>	0,0004164397994
F - F <sup>4</sup>	<b>0,9338689508</b>	0,00208403653
F <sup>2</sup> - F <sup>3</sup>	<b>0,9924301526</b>	9,536921139E-006
F <sup>2</sup> - F <sup>4</sup>	<b>0,9741448225</b>	0,0002036013029
F <sup>3</sup> - F <sup>4</sup>	<b>0,994406487</b>	4,48087527E-006

## Indikace multikolinearity

Proměnná	Vlas. čísla kor. m.	Podmíněnost kappa	VI faktor	Vícenás. kor.
Abs	8,048165545E-007	1	1	0
F	1	1242519,173	43557,39282	0,9999885208
F <sup>2</sup>	0,07400761615	91955,88203	482578,2239	0,9999989639
F <sup>3</sup>	0,0005986271464	743,8057071	625285,3468	0,9999992004
F <sup>4</sup>	3,925392952	4877376,005	92782,46601	0,999994611

## Analýza rozptylu

Průměr Y : 1566366,571

Zdroj	Součet čtverců	Průměrný čtverec	Rozptyl
Celková variabilita	1,800591393E+012	2,572273418E+011	3,000985654E+011
Variab.vysv.modelem	1,800415832E+012	2,572022615E+011	3,000693051E+011
Reziduální variabilita	175560040,7	25080005,82	29260006,79

Hodnota kritéria F : 5127,635609

Kvantil F (1-alfa, m-1, n-m) : 19,24679434

Pravděpodobnost : 0,000194993137

**Závěr : Model je významný**

## Odhady parametrů

Proměnná	Odhad	Směr.Odch.	Závěr	Pravděpod.	Spodní mez	Horní mez
Abs	1815193,8	482772,4047	Nevýznamný	0,06401805995	-262008,2051	3892395,805
F	-66765,287	22429,62335	Nevýznamný	0,09675797477	-163272,1671	29741,59313
F <sup>2</sup>	1343,427056	369,5462525	Nevýznamný	0,06803598463	-246,6021361	2933,456248
F <sup>3</sup>	-9,323026575	2,572155161	Nevýznamný	0,06839945285	-20,390117	1,744063849
F <sup>4</sup>	0,02329180255	0,006422483268	Nevýznamný	0,06833136283	-0,0043419126	0,050925518

## Statistické charakteristiky regrese

Vícenásobný korelační koeficient R :	0,9999512482
Koeficient determinace R <sup>2</sup> :	0,9999024987
Predikovaný korelační koeficient Rp :	0,9846678331
Střední kvadratická chyba predikce MEP :	1979543251
Akaikeho informační kritérium :	129,2630705

## Analýza klasických reziduí

Index	Y naměřené	Y vypočítané	Směr. odch. Y	Reziduum	Reziduum [%Y]	Váhy
1	815138	815692,5338	9345,229758	-554,5337819	-0,06802943574	1
2	1091326	1085854,348	7779,255629	5471,652051	0,5013764953	1
3	1244022	1252546,624	6201,580165	-8524,623869	-0,6852470349	1
4	1585449	1579089,339	8141,291906	6359,660941	0,4011268064	1
5	1863204	1868307,22	6201,580166	-5103,220085	-0,2738948652	1
6	2011784	2009306,076	7779,255629	2477,924407	0,1231704998	1
7	2353643	2353769,858	9345,229757	-126,8575623	-0,005389838743	1

Reziduální součet čtverců : 175560040,7

Průměr absolutních reziduí : 4088,353242

Reziduální směr. odchylka : 9369,099229

Reziduální rozptyl : 87780020,36

Šikmost reziduí : 0,129428629

Špičatost reziduí : 1,937068653

## Testování regresního tripletu

### Fisher-Snedecorův test významnosti modelu

Hodnota kritéria F : 5127,635609

Kvantil F (1-alfa, m-1, n-m) : 19,24679434

Pravděpodobnost : 0,000194993137

**Závěr : Model je významný**

### Scottovo kritérium multikolinearity

Hodnota kritéria SC : 0,9951360602

**Závěr : Model je nekorektní!**

### Cook-Weisbergův test heteroskedasticity

Hodnota kritéria CW : 0,02495169128

Kvantil  $\chi^2(1-\alpha,1)$  : 3,841458829

Pravděpodobnost : 0,874487496

**Závěr : Rezidua vykazují homoskedasticitu.**

### Jarque-Berrův test normality

Hodnota kritéria JB : 0,4805317894

Kvantil  $\chi^2(1-\alpha,2)$  : 5,991464547

Pravděpodobnost : 0,7864187287

**Závěr : Rezidua mají normální rozdělení.**

### Waldův test autokorelace

Hodnota kritéria WA : 0,6117076366

Kvantil  $\chi^2(1-\alpha,1)$  : 3,841458829

Pravděpodobnost : 0,4341455783

**Závěr : Autokorelace je nevýznamná**

### Durbin-Watsonův test autokorelace

Hodnota kritéria DW : -1

Kritické hodnoty DW 0 2

**Závěr : Negativní autokorelace reziduí není prokázána.**

### Znaménkový test reziduí

Hodnota kritéria Sg : 2,607257969

Kvantil  $N(1-\alpha/2)$  : 1,959963999

Pravděpodobnost : 0,00912705693

**Závěr : V reziduích je trend!**

### Indikace vlivných dat

#### A. Analýza reziduí

Index	Standardní	Jackknife	Predikované	Diag(Hii)	Diag(H*ii)	Cook.vzdál.
1	-0,8296969223	-0,7244664458	-108969,9188	0,9949111297	0,9966627112	-32,44235579
2	1,047924611	1,103473529	17617,21623	0,689414492	0,859948607	0,465220942
3	-1,213841065	-1,672728961	-15172,04434	0,4381361088	0,8520640836	-0,189308339
4	1,371579322	3,979855036	25965,91154	0,7550765383	0,9854552003	0,845690617
5	-0,7266594044	-0,5989381723	-9082,662483	0,4381361088	0,5864776969	-0,113328416
6	0,4745692791	0,3562270289	7978,235762	0,689414492	0,724388905	0,210682682
7	-0,1898050789	-0,1354378212	-24928,43268	0,9949111296	0,9950027954	-7,421654393

## B. Analýza vlivu

Index	Atkinson.vzdál.	Andrews-Pregibon st.	Vliv na Y <sup>^</sup>	Vliv na par.LD(b)	Vliv na rozptyl LD(s)	Celk.vliv LD(b,s)
1	<b>6,406631213</b>	0,003337288812	-10,12977338	<b>29,56668236</b>	0,2748847457	<b>615,9487373</b>
2	1,039780158	0,140051393	1,644036782	5,578737758	1,80982911	<b>18,02697075</b>
3	0,9342100475	0,1479359164	-1,477115782	3,177455809	7,525824655	<b>20,61705085</b>
4	<b>4,41954482</b>	0,0145447997	<b>6,987913926</b>	9,526529672	<b>75,34854571</b>	<b>368,3341919</b>
5	0,3345037191	0,4135223031	-0,5288968191	1,310454935	0,08557907909	1,763970151
6	0,335665321	0,275611095	0,5307334729	1,561774252	0,004161746776	1,694229862
7	<b>1,197709254</b>	0,004997204623	-1,893744609	10,56216859	0,06187483042	<b>21,57948669</b>

## IV. stupeň polynomu 5

Hladina významnosti :	0,05
Kvantil t(1-alfa/2,n-m) :	12,7062047508531
Kvantil F(1-alfa,m,n-m) :	230,161878110122
Absolutní člen :	Ano
Počet platných řádků :	7
Počet parametrů :	6
Metoda :	Nejmenší četverce
Sloupce pro výpočet :G	
	Abs
	F
	F <sup>2</sup>
	F <sup>3</sup>
	F <sup>4</sup>
	F <sup>5</sup>
Transformace :	Polynom 5. stupně

## Základní analýza

### Charakteristiky proměnných

Proměnná	Průměr	Směr.Odch.	Kor.vs.Y	Významnost
F	100	35,59026084	0,9995032131	1,05624216E-008
F <sup>2</sup>	11085,71429	7190,13875	0,991735886	1,187187073E-005
F <sup>3</sup>	1325714,286	1175883,174	0,9691572712	0,0003155891109
F <sup>4</sup>	167205714,3	181406225,6	0,9390719456	0,001702834286
F <sup>5</sup>	2,189E+010	2,746E+010	0,9064157896	0,004890264053

### Párové korelace (Xi, Xj)

Abs - F	0	1
Abs - F <sup>2</sup>	0	1
Abs - F <sup>3</sup>	0	1
Abs - F <sup>4</sup>	0	1

Abs - F <sup>5</sup>	0	1
F - F <sup>2</sup>	<b>0,9899742433</b>	1,922693999E-005
F - F <sup>3</sup>	<b>0,9655120051</b>	0,0004164397994
F - F <sup>4</sup>	<b>0,9338689508</b>	0,00208403653
F - F <sup>5</sup>	<b>0,8999078591</b>	0,005764478487
F <sup>2</sup> - F <sup>3</sup>	<b>0,9924301526</b>	9,536921139E-006
F <sup>2</sup> - F <sup>4</sup>	<b>0,9741448225</b>	0,0002036013029
F <sup>2</sup> - F <sup>5</sup>	<b>0,9503683362</b>	0,001026135896
F <sup>3</sup> - F <sup>4</sup>	<b>0,994406487</b>	4,48087527E-006
F <sup>3</sup> - F <sup>5</sup>	<b>0,980862621</b>	9,631477741E-005
F <sup>4</sup> - F <sup>5</sup>	<b>0,9958820802</b>	2,085402175E-006

### Indikace multikolinearity

Proměnná	Vlas. čísla kor. m.	Podmíněnost kappa	VI faktor	Vícenás. kor.
Abs	3,117029266E-009	1	1	0
F	1	<b>320818290,3</b>	<b>2582249,181</b>	0,9999998064
F <sup>2</sup>	0,001602954226	<b>514257,0342</b>	<b>51789887,01</b>	0,9999999903
F <sup>3</sup>	0,126851578	<b>40696306,36</b>	<b>157728709,3</b>	0,9999999968
F <sup>4</sup>	4,96129316E-006	<b>1591,673589</b>	<b>99782856,41</b>	0,999999995
F <sup>5</sup>	4,871540503	<b>1562879295</b>	<b>9136793,225</b>	0,9999999453

### Analýza rozptylu

Průměr Y : 1566366,571

Zdroj	Součet čtverců	Průměrný čtverec	Rozptyl
Celková variabilita	1,800591393E+012	2,572273418E+011	3,000985654E+011
Var.vysv.modelem	1,800426258E+012	2,572036124E+011	3,000708811E+011
Reziduální variabilita	165134383,1	23590626,15	27522397,14

Hodnota kritéria F : 2180,558918

Kvantil F (1-alfa, m-1, n-m) : 230,1618781

Pravděpodobnost : 0,01625698891

**Závěr : Model je významný**

### Odhady parametrů

Proměnná	Odhad	Směr.Odch.	Závěr	Pravděp.	Spodní mez	Horní mez
Abs	799036,8623	4098005,237	Nevýznamný	0,877409	-51271056,75	52869130,47
F	-7751,976837	236870,5523	Nevýznamný	0,9791730085	-3017477,713	3001973,76
F <sup>2</sup>	30,23421764	5250,835107	Nevýznamný	0,9963343949	-66687,95176	66748,4202
F <sup>3</sup>	4,727894828	56,03177348	Nevýznamný	0,9464097191	-707,2232915	716,6790812
F <sup>4</sup>	-0,04926028795	0,2888809798	Nevýznamný	0,8924771308	-3,719841165	3,62132059
F <sup>5</sup>	0,00014510418	0,0005774933	Nevýznamný	0,8432836695	-0,0071926437	0,007482852

### Statistické charakteristiky regrese

Vícenásobný korelační koeficient R : 0,9999541434  
Koeficient determinace  $R^2$  : 0,9999082888  
Predikovaný korelační koeficient  $R_p$  : 0,5424213453  
Střední kvadratická chyba predikce MEP : 6,778130884E+010  
Akaikeho informační kritérium : 130,83452

### Analýza klasických reziduí

Index	Y naměřené	Y vypočítané	Směr. odch.	YReziduum	Reziduum [%Y]	Váhy
1	815138	815478,6749	12845,94525	-340,6749133	-0,04179352617	1
2	1091326	1087351,16	12220,28846	3974,840486	0,3642211847	1
3	1244022	1250835,846	10895,17575	-6813,846166	-0,5477271436	1
4	1585449	1579089,199	11166,42727	6359,801406	0,4011356661	1
5	1863204	1870017,688	10895,17567	-6813,687796	-0,3656973577	1
6	2011784	2007808,918	12220,28819	3975,082238	0,1975899121	1
7	2353643	2353983,254	12845,94469	-340,2544678	-0,01445650287	1

Reziduální součet čtverců : 165134383,1  
Průměr absolutních reziduí : 4088,312496  
Reziduální směr. odchylka : 12850,46237  
Reziduální rozptyl : 165134383,1  
Šikmost reziduí : 0,0970956473  
Špičatost reziduí : 1,654755267

### Testování regresního tripletu

#### Fisher-Snedecorův test významnosti modelu

Hodnota kritéria F : 2180,558918  
Kvantil F (1-alfa, m-1, n-m) : 230,1618781  
Pravděpodobnost : 0,01625698891

**Závěr : Model je významný**

#### Scottovo kritérium multikolinearity

Hodnota kritéria SC : 0,9999788354

**Závěr : Model je nekorektní!**

#### Cook-Weisbergův test heteroskedasticity

Hodnota kritéria CW : 3,995462E-006  
Kvantil  $\chi^2(1-\alpha, 1)$  : 3,841458829  
Pravděpodobnost : 0,9984051374

**Závěr : Rezidua vykazují homoskedasticitu.**



### **Jarque-Berrův test normality**

Hodnota kritéria JB : 0,6411025776

Kvantil  $\text{Chi}^2(1-\alpha,2)$  : 5,991464547

Pravděpodobnost : 0,7257488296

**Závěr : Rezidua mají normální rozdělení.**

### **Waldův test autokorelace**

Hodnota kritéria WA : 0,1500731671

Kvantil  $\text{Chi}^2(1-\alpha,1)$  : 3,841458829

Pravděpodobnost : 0,698465447

**Závěr : Autokorelace je nevýznamná**

### **Durbin-Watsonův test autokorelace**

Hodnota kritéria DW : -1

Kritické hodnoty DW 0 2

**Závěr : Negativní autokorelace reziduí není prokázána.**

### **Znaménkový test reziduí**

Hodnota kritéria Sg : 2,607257969

Kvantil  $N(1-\alpha/2)$  : 1,959963999

Pravděpodobnost : 0,00912705693

**Závěr : V reziduích je trend!**

### **Závěr :**

Optimální stupeň polynomu je  $n=2$ , když MEP má minimální hodnotu. Model však vykazuje multikolinearitu a není možné akceptovat metodu nejmenších čtverců. Je nutné použít metodu racionálních hodnot (RH), která spočívá v nalezení optimálního filtru omezení P.

## **V. Metoda racionálních hodnotí**

Filtr omezení P = 1

Hladina významnosti :	0,05
Kvantil t(1-alfa/2,n-m) :	2,77644510519769
Kvantil F(1-alfa,m,n-m) :	6,94427190984624
Absolutní člen :	Ano
Počet platných řádků :	7
Počet parametrů :	3
Metoda :	Nejmenší ètverce
Transformace :	Polynom 2. stupni

### **Základní analýza**

#### **Charakteristiky proměnných**

Proměnná	Průměr	Směr.Odch.	Kor.vs.Y	Významnost
F	100	35,59026084	0,99950321311	0,05624216E-008
F <sup>2</sup>	11085,71429	7190,13875	0,991735886	1,187187073E-005

#### **Párové korelace (Xi, Xj)**

Abs - F	0	1
Abs - F <sup>2</sup>	0	1
F - F <sup>2</sup>	0,98997424331	9,22693999E-005

#### **Indikace multikolinearity**

Proměnná	Vlas. čísla kor. m.	Podmíněnost kappa	VI faktor	Vícenás. kor.
Abs	0,01002575667	1	1	0
F	1	99,74309497	50,12280702	0,9899742433
F <sup>2</sup>	1,989974243	198,4861899	50,12280702	0,9899742433

#### **Analýza rozptylu**

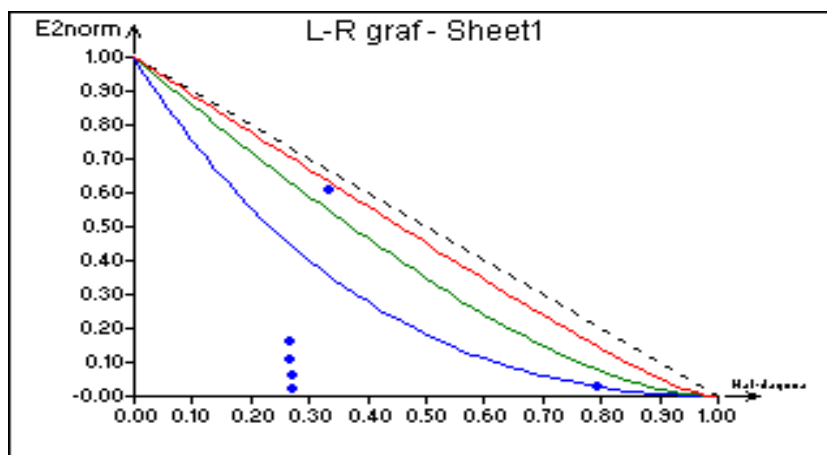
Průměr Y : 1566366,571

Zdroj	Součet ètvercù	Průměrný ètverec	Rozptyl
Celková variabilita	1,800591393E+012	2,572273418E+011	3,000985654E+011
Var.vysv. modelem	1,799261113E+012	2,570373018E+011	2,998768521E+011
Reziduální variabilita	1330280000	190040000	221713333,3
Hodnota kritéria F :	2705,086317		
Kvantil F (1-alfa, m-1, n-m) :	6,94427191		
Pravděpodobnost :	5,458279668E-007		

**Závěr : Model je významný**

## Odhady parametrů

Proměnná	Odhad	Směr.Odch.	Závěr	Pravděpod.	Spodní mez	Horní mez
Abs	104623,0956	68959,94943	Nevýznamný	0,2038265607	-86840,41849	296086,6096
F	13663,44993	1480,991041	Významný	0,00076709152	9551,559603	17775,34026
F <sup>2</sup>	8,605533241	7,330714927	Nevýznamný	0,3055576406	-11,74779434	28,95886082



Soubor dat obsahuje 1 odlehlý bod č.4, což je vidět z L-R grafu.

## Statistické charakteristiky regrese

Vícenásobný korelační koeficient R : 0,9996305309

Koeficient determinace R<sup>2</sup> : 0,9992611983

Predikovaný korelační koeficient Rp : 0,9952346828

Střední kvadratická chyba predikce MEP : 613616826,6

Akaikeho informační kritérium : 139,4392159

## Analýza klasických reziduí

Index	Y naměřené	Y vypočítané	Směr. odch. Y	Reziduum	Reziduum [%Y]	Váhy
1	815138	809309,4252	16260,3123	5828,574792	0,7150414767	1
2	1091326	1103231,704	9420,549334	-11905,7036	-1,090939243	1
3	1244022	1252774,503	9501,692515	-8752,50277	-0,7035649506	1
4	1585449	1557023,421	10528,84926	28425,57895	1,792904026	1
5	1863204	1868156,766	9501,692515	-4952,765928	-0,265819842	1
6	2011784	2026305,098	9420,549334	-14521,09834	-0,7218020592	1
7	2353643	2347765,083	16260,3123	5877,916897	0,2497369778	1

Reziduální součet čtverců : 1330280000

Průměr absolutních reziduí : 11466,3059

Reziduální směr. odchylka : 18236,50186

Reziduální rozptyl : 332570000

Šikmost reziduí : 0,9450730666

Špičatost reziduí : 2,872785949

### **Testování regresního tripletu**

#### **Fisher-Snedecorův test významnosti modelu**

Hodnota kritéria F : 2705,086317

Kvantil F (1-alfa, m-1, n-m) : 6,94427191

Pravděpodobnost : 5,458279668E-007

**Závěr : Model je významný**

#### **Scottovo kritérium multikolinearity**

Hodnota kritéria SC : 0,9783529686

**Závěr : Model je nekorektní!**

#### **Cook-Weisbergův test heteroskedasticity**

Hodnota kritéria CW : 0,0001791848938

Kvantil  $\chi^2(1-\text{alfa}, 1)$  : 3,841458829

Pravděpodobnost : 0,9893198393

**Závěr : Rezidua vykazují homoskedasticitu.**

#### **Jarque-Berrův test normality**

Hodnota kritéria JB : 1,107305407

Kvantil  $\chi^2(1-\text{alfa}, 2)$  : 5,991464547

Pravděpodobnost : 0,574846228

**Závěr : Rezidua mají normální rozdělení.**

#### **Waldův test autokorelace**

Hodnota kritéria WA : 0,2467453462

Kvantil  $\chi^2(1-\text{alfa}, 1)$  : 3,841458829

Pravděpodobnost : 0,6193761604

**Závěr : Autokorelace je nevýznamná**

#### **Durbin-Watsonův test autokorelace**

Hodnota kritéria DW : -1

Kritické hodnoty DW 0 2

**Závěr : Negativní autokorelace reziduí není prokázána.**

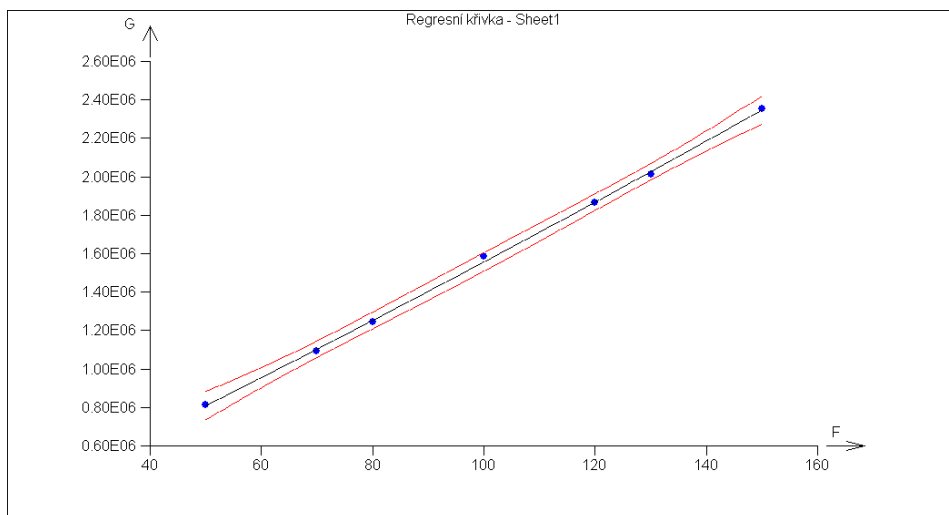
#### **Znaménkový test reziduí**

Hodnota kritéria Sg : 0,9095085939

Kvantil  $N(1-\text{alfa}/2)$  : 1,959963999

Pravděpodobnost : 0,3630817236

**Závěr : V reziduích není trend.**



## Indikace vlivných dat

### A. Analýza reziduí

Index	Standardní	Jackknife	Predikované	Diag(Hii)	Diag(H*ii)	Cook.vzdál.
1	0,7059250278	0,6534037283	28433,99324	0,7950138504	0,8205515448	0,912614788
2	-0,7624607786	-0,7142505596	-16239,1398	0,2668513389	0,3734046797	-0,092506786
3	-0,562297883	-0,5074319524	-12013,89164	0,271468144	0,3290547459	-0,069841816
4	1,9090329	5,544944499	42638,36842	0,3333333333	0,9407344358	0,318172150
5	-0,3181866798	-0,2791126533	-6798,28327	0,271468144	0,2899077886	-0,039521286
6	-0,9299549456	-0,909684998	-19806,48552	0,2668513389	0,4253610488	-0,112828287
7	0,711901073	0,6597338144	28674,7027	0,7950138504	0,8209857564	0,920340576

### B. Analýza vlivu

Index	Atkinson.vzd.	Andrews-Pregibon st.	Vliv na Y <sup>^</sup>	Vliv na param. LD(b)	Vliv na rozptyl LD(s)	Celk.vliv LD(b,s)
1	1,485853751	0,1794484552	1,286787095	2,759310594	0,001546788597	3,313188843
2	0,4975750872	0,6265953203	-0,4309126658	0,3608351555	2,940055445E-005	0,371400852
3	0,3576698583	0,6709452541	-0,3097511835	0,2031994681	0,0176224772	0,209514069
4	<b>4,527428225</b>	0,05926556424	<b>3,920867857</b>	2,627690829	<b>44,63004159</b>	<b>75,3764511</b>
5	0,1967360997	0,7100922114	-0,1703784602	0,06571013729	0,05540689331	0,11346453
6	0,6337224187	0,5746389512	-0,5488197136	0,5302568157	0,02886242532	0,631268581
7	1,500248499	0,1790142436	1,299253312	2,797976814	0,001212905647	3,377331245

### Závěr :

Výsledný regresní vztah modelu závislost HPLC signálu = F (obsah guaifenesinu) má tvar :

$$Y = 13663 (\pm 1480) \cdot x$$

kde :

Y - signál (plocha píku)

x - obsah guaifenesinu

### Příklad 3 – Validizace nové analytické metody

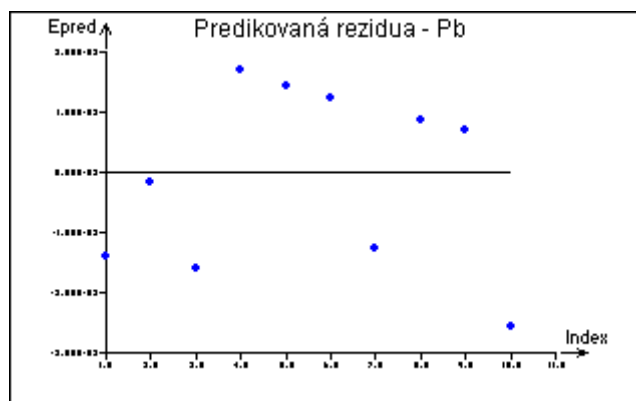
#### Stanovení Pb metodou AAS

**Zadání:** Byly porovnávány naměřené hodnoty Pb na AAS s atestovanými hodnotami. Na základě testování statistické významnosti regresních parametrů (úseku a směrnice) byla určena velikost systematické chyby metody, pro kterou by směrnice měla být 1.

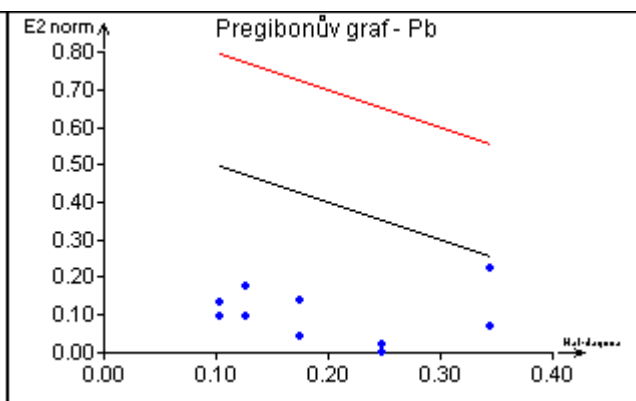
**Data:**

Olovo	
Atest x [%]	Neměřeno y [%]
0,005	0,003
0,010	0,009
0,015	0,013
0,020	0,021
0,025	0,026
0,030	0,031
0,035	0,034
0,040	0,041
0,045	0,046
0,050	0,049

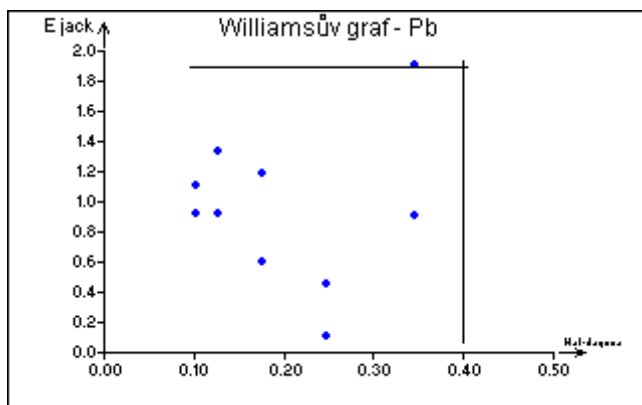
**Řešení:** Pomocí grafů vlivných bodů nebyly identifikovány významné odlehlé body a extrémny, jak je vidět z grafů č.1-5. Testováním regresního tripletu byla prokázána správnost navrženého modelu. Metodou nejmenších čtverců byl určen odhad parametru úseku  $b_0$  a odhad parametru směrnice  $b_1$ . 95-% intervaly spolehlivosti parametru úseku  $b_0$  a parametru směrnice  $b_1$  obsahují 0 a 1 a tudíž úsek nelze považovat za významně od nuly a směrnici lze považovat za jednotkovou (viz tab.č.1). Graf č.6 znázorňuje konstrukci 95% pásů spolehlivosti predikce pro daný regresní model.



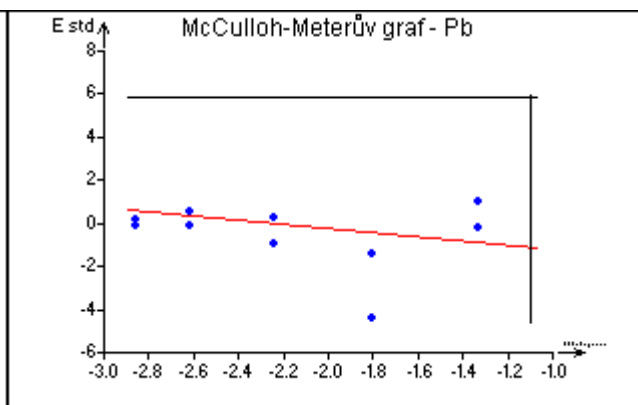
Graf č.1



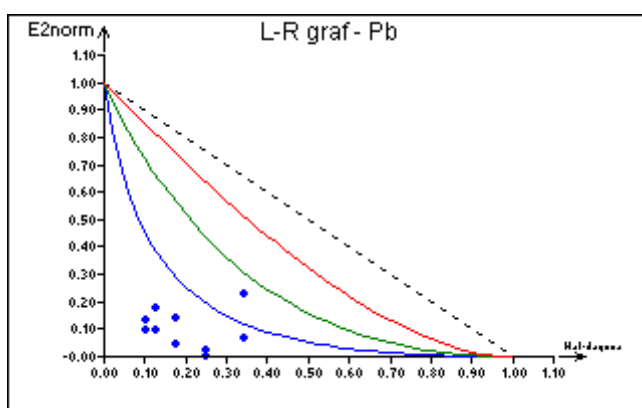
Graf č.2



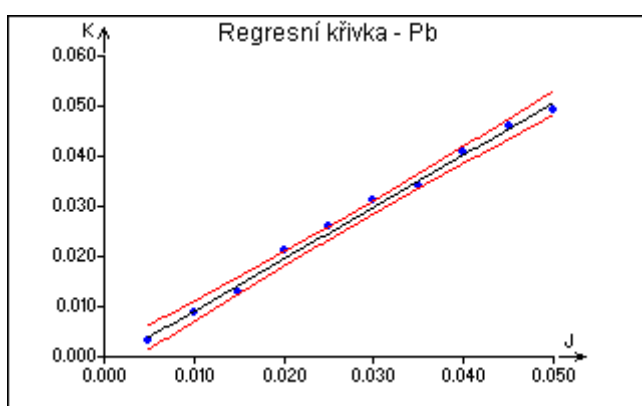
Graf č.3



Graf č.4



Graf č.5



Graf č.6

### (1) ZVOLENÁ STRATEGIE REGRESNÍ ANALÝZY

Hladina významnosti : 0,05  
 Kvantil  $t(1-\alpha/2, n-m)$  : 2,30600413519908  
 Kvantil  $F(1-\alpha, m, n-m)$  : 5,31765507155526  
 Absolutní člen : Ano  
 Počet platných řádků : 10  
 Počet parametrů : 2  
 Metoda : Nejmenší čtverce  
 Sloupce pro výpočet : K  
   Abs  
   J  
 Transformace : Bez transformace

## Základní analýza

### Charakteristiky proměnných

Proměnná	Průměr	Směr.Odch.	Kor.vs.Y	Významnost
J	0,0275	0,01513825177	0,9972042215	2,663982368E-010

### Párové korelace (Xi, Xj)

Abs - J0 1

### Indikace multikolinearity

Proměnná	Vlas. čísla kor. m.	Podmíněnost kappa	VI faktor	Vícenás. kor.
Abs	1	1	1	0
J	1	1	1	0

### Analýza rozptylu

Průměr Y : 0,0273

Zdroj	Součet čtverců	Průměrný čtverec	Rozptyl
Celková variabilita	0,0022381	0,00022381	0,0002486777778
Var.vysv.modelem	0,00222560303	0,000222560303	0,0002472892256
Reziduální variabilita	1,24969697E-005	1,24969697E-006	1,388552189E-006

Hodnota kritéria F : 1424,731329

Kvantil F (1-alfa, m-1, n-m) : 5,317655072

Pravděpodobnost : 2,663982069E-010

**Závěr : Model je významný**

### Odhady parametrů

Proměnná	Odhad	Směr.Odch.	Závěr	Pravd.	Spodní mez	Horní mez
Abs	-0,00127	0,00085381	Nevýzn.	0,176	-0,003236	0,000702
J	1,0387878	0,02752076	Význ.	2,66E-010	0,9753249	1,10225

### Statistické charakteristiky regrese

Vícenásobný korelační koeficient R : 0,9972042215

Koeficient determinace R<sup>2</sup> : 0,9944162595

Predikovaný korelační koeficient Rp : 0,9818290323

Střední kvadratická chyba predikce MEP : 2,042744348E-006

Akaikeho informační kritérium : -131,9260946

### Analýza klasických reziduí

Index	Y naměřené	Y vypočítané	Směr. odch. Y	Reziduum	Reziduum [%Y]	Váhy
1	0,003	0,003927272727	0,0007346032081	-0,0009272727273	-30,90909091	1
2	0,009	0,009121212121	0,00062302765	-0,0001212121212	-1,346801347	1
3	0,013	0,01431515152	0,0005239796153	-0,001315151515-	10,11655012	1
4	0,021	0,01950909091	0,0004458872357	0,001490909091	7,0995671	1
5	0,026	0,0247030303	0,0004011805352	0,001296969697	4,988344988	1
6	0,031	,0298969697	0,0004011805352	0,001103030303	3,558162268	1
7	0,034	0,03509090909	0,0004458872357	-0,001090909091	-3,20855615	1
8	0,041	0,04028484848	0,0005239796153	0,0007151515151	1,744271988	1
9	0,046	0,04547878788	0,00062302765	0,0005212121212	1,133069829	1
10	0,049	0,05067272727	0,0007346032081	-0,001672727273	-3,413729128	1

Reziduální součet čtverců : 1,24969697E-005

Průměr absolutních reziduí : 0,001025454545

Reziduální směr. odchylka : 0,001249848476

Reziduální rozptyl : 1,562121212E-006

Šikmost reziduí : 0,01493549723

Špičatost reziduí : 1,444696654



### Testování regresního tripletu

#### Fisher-Snedecorův test významnosti modelu

Hodnota kritéria F : 1424,731329  
Kvantil F (1-alfa, m-1, n-m) : 5,317655072  
Pravděpodobnost : 2,663982069E-010  
**Závěr : Model je významný**

#### Scottovo kritérium multikolinearity

Hodnota kritéria SC : 0,3326471122  
**Závěr : Model vykazuje multikolinearitu!**

#### Cook-Weisbergův test heteroskedasticity

Hodnota kritéria CW : 0,05707838479  
Kvantil  $\chi^2(1-\text{alfa}, 1)$  : 3,841458829  
Pravděpodobnost : 0,8111747274  
**Závěr : Rezidua vykazují homoskedasticitu.**

#### Jarque-Berrův test normality

Hodnota kritéria JB : 1,032796036  
Kvantil  $\chi^2(1-\text{alfa}, 2)$  : 5,991464547  
Pravděpodobnost : 0,5966658616  
**Závěr : Rezidua mají normální rozdělení.**

#### Waldův test autokorelace

Hodnota kritéria WA : 0,03307960698  
Kvantil  $\chi^2(1-\text{alfa}, 1)$  : 3,841458829  
Pravděpodobnost : 0,8556784698  
**Závěr : Autokorelace je nevýznamná**

#### Durbin-Watsonův test autokorelace

Hodnota kritéria DW : -1  
Kritické hodnoty DW 0 2  
**Závěr : Pozitivní autokorelace reziduí není prokázána.**

#### Znaménkový test reziduí

Hodnota kritéria Sg : 0,3354101966  
Kvantil  $N(1-\text{alfa}/2)$  : 1,959963999  
Pravděpodobnost : 0,7373156772  
**Závěr : V reziduích není trend.**

### Indikace vlivných dat

#### A. Analýza reziduí

Index	Standardní	Jackknife	Predikované	Diag(Hii)	Diag(H*ii)	Cook. vzdál.
1	-0,917022974	-0,9067779817	-0,001416667	0,3454545455	0,4142580019	-0,2419921737
2	-0,1118715914	-0,1047282424	-0,000161290	0,2484848485	0,2496605238	-0,01849490019
3	-1,159020033	-1,18853333	-0,001595588	0,1757575758	0,3141610087	-0,1235719888
4	1,276893053	1,338595524	0,001708333333	0,1272727273	0,3051406402	0,09310678509
5	1,095679607	1,111717777	0,001445945946	0,103030303	0,2376333657	0,06292754501
6	0,931839666	0,9231971317	0,00122972973	0,103030303	0,2003879728	0,05351781866
7	-0,9343119898	-0,9259460874	-0,00125	0,1272727273	0,2225024248	-0,06812691592
8	0,630250525	0,6047500351	0,0008676471	0,1757575758	0,2166828322	0,06719582804
9	0,4810478431	0,4566317529	0,0006935484	0,2484848485	0,2702230844	0,07952807083
10	-1,654237522	-1,9076981	-0,0025555556	0,3454545455	0,5693501455	-0,4365349016

## B. Analýza vlivu

Index	Atkinson.vzdál.	Andrews-Pregibon st.	Vliv na $Y^{\wedge}$	Vliv na par.LD(b)	Vliv na rozptyl LD(s)	Celk.vliv LD(b,s)
1	1,317517862	0,5857419981	-0,6587589311	0,5399383556	0,000162821996	0,55811567
2	0,1204411508	0,7503394762	-0,06022057542	0,005171297447	0,05205056116	0,05671323
3	1,097667753	0,6858389913	-0,5488338764	0,3517948401	0,03160194097	0,4188830
4	1,022369219	0,6948593598	0,5111846096	0,2928871894	0,07826044722	0,4142300
5	0,7535605407	0,7623666343	0,3767802703	0,1709023341	0,01669563415	0,19921997
6	0,6257747641	0,7996120272	0,312887382	0,123904272	0,0004560216455	0,1263255
7	0,7072030058	0,7774975752	-0,3536015029	0,1578766497	0,00052013596	0,1612782
8	0,5585157732	0,7833171678	0,2792578866	0,1053189585	0,01454984023	0,1148162
9	0,5251425266	0,7297769156	0,2625712633	0,09518754305	0,02816794794	0,11680981
10	<b>2,771821077</b>	0,4306498545	-1,385910539	1,65966119	0,5462638876	3,015797057

Intervalový odhad parametra úseku  $\beta_0$  a směrnice  $\beta_1$  bude

$$b_0 - t(1-\alpha/2, n-m)\sqrt{D(b_0)} \leq \beta_0 \leq b_0 + t(1-\alpha/2, n-m)\sqrt{D(b_0)}$$

po dosazení

$$-0,00127 - 2,306 \times 0,000854 \leq \beta_0 \leq -0,00127 + 2,306 \times 0,000854$$

$$\mathbf{- 0,00324 \leq \beta_0 \leq 0,0007028}$$

**Tento intervalový odhad spolehlivosti úseku regresní přímky obsahuje nulu, takže úsek  $\beta_0$  lze považovat za nulový.**

Analogickým dosazením do intervalu spolehlivosti směrnice obdržíme nerovnost

$$1,0388 - 2,306 \times 0,0275 \leq \beta_1 \leq 1,0388 + 2,306 \times 0,0275$$

$$\mathbf{0,975 \leq \beta_1 \leq 1,102}$$

**Tento intervalový odhad spolehlivosti směrnice regresní přímky obsahuje jednotku, takže směrnici  $\beta_1$  lze považovat za jednotkovou.**

Tabulka č.1

parametr	Odhady parametrů		95% intervaly spolehlivosti			
	$b_0$	$b_1$	úseku		směrnice	
hodnota	-0,00127	1,0388	-0,00323	0,0007	0,975	1,102

**Závěr:**  $r = 0,9972$ ,  $D = 99,44\%$ ,  $y = 1,0388(\pm 0,0275).x$ ,  
0 odlehlých bodů, 1 extrém (bod č.10).

Naměřené hodnoty  $P_b$  se statisticky významně neliší od atestovaných hodnot a tudíž metodu AAS lze považovat za správnou a přesnou.

## Příklad 4 – Vícerozměrný lineární regresní model

### Závislost obsahu cholesterolu na věku, výšce a váze pacienta

**Zadání:** Výběr 20 pacientů, nemocných s hyperliproteinemií byl vyšetřován na hladinu lipidů v plasmě totálního cholesterolu  $y$  s přihlédnutím k věku  $x_1$ , váze  $x_2$ , celkovému dennímu příjmu kalorií  $x_3$  a podílu tuků v nich  $x_4$ . Bude vyšetřována signifikance všech čtyř parametrů na obsah totálního cholesterolu s cílem zjistit míru jejich významnosti a korektnost navrhovaného vícerozměrného lineárního modelu s absolutním členem.

**Data:**

Cholesterol [mmol/kg]	Věk [rok]	Váha [kg]	Kalorie [KJ/den]	Tuky [%/den]
4500	15	58	210	5
4550	17	67	180	9
4350	19	64	100	9
4650	28	68	105	11
4700	22	70	145	13
4750	32	78	200	15
4800	33	80	228	18
4850	27	79	263	19
4900	34	82	280	22
4950	28	78	310	20
5000	36	90	345	18
5050	39	85	378	24
5080	18	64	410	23
5150	43	91	445	28
5200	48	90	490	30
5210	47	78	405	31
5290	50	90	410	30
5378	49	85	416	32
5410	51	91	495	34
5460	52	94	510	31

### 1. Řešení

Návrh modelu:  $y = b_0 + b_1 \cdot x_1 + b_2 \cdot x_2 + b_3 \cdot x_3 + b_4 \cdot x_4$   
Hladina významnosti : 0,05  
Kvantil  $t(1-\alpha/2, n-m)$  : 2,13144954555968  
Kvantil  $F(1-\alpha, m, n-m)$  : 3,05556827590656  
Absolutní člen : Ano  
Počet platných řádků : 20  
Počet parametrů : 5  
Metoda : Nejmenší čtverce  
Sloupce pro výpočet : Cholesterol [mmol/kg]  
Abs  
Věk [rok]  
Váha [kg]  
Kalorie [KJ/den]  
Tuky [%/den]  
Transformace : Bez transformace

## 2. Základní analýza

Charakteristiky proměnných

Proměnná	Průměr	Směr.Odch.	Kor.vs.Y	Významnost
Věk [rok]	34,4	12,49589406	0,879965	3,153607 E-007
Váha [kg]	79,1	10,75517012	0,804821	1,876176 E-005
Kalorie [KJ/den]	316,25	132,2600608	0,938938	8,971175 E-010
Tuky [%/den]	21,1	8,831164911	0,965902	5,229150 E-012

## 3. Analýza rozptylu

Průměr	Y :	4961,4		
Zdroj		Součet čtverců	Průměrný čtverec	Rozptyl
Celková variabilita		1917384,8	95869,24	100914,9895
Variabilita vysvětlená modelem		1830798,121	91539,90609	96357,79589
Reziduální variabilita		86586,67874	4329,333937	4557,193618
Hodnota kritéria F :		79,2904065		
Kvantil F (1-alfa, m-1, n-m) :		3,055568276		
Pravděpodobnost :		6,642374463E-010		
Závěr :		Model je významný		

## 4. Odhady parametrů

Proměnná	Odhad	Směr.Odch.	Závěr	Pravděpod.	Spodní mez	Horní mez
Abs	4159,1	202,2	Významný	2,1 E-012	3728,1	4590,2
Věk [rok]	2,8	4,5	Nevýznamný	0,54	-6,8	12,4
Váha [kg]	0,4	3,8	Nevýznamný	0,91	-7,7	8,5
Kalorie [KJ/den]	0,9	0,3	Významný	0,02	0,1	1,6
Tuky [%/den]	18,9	7,1	Významný	0,02	3,7	34,0

## 5. Statistické charakteristiky regrese

Vícenásobný korelační koeficient R :	0,977160
Koeficient determinace R <sup>2</sup> :	0,954841
Predikovaný korelační koeficient Rp :	0,823935
Střední kvadratická chyba predikce MEP :	8847,899
Akaikeho informační kritérium :	177,4634

## 6. Testování regresního tripletu

Fisher-Snedecorův test významnosti modelu

Hodnota kritéria F :	79,290407
Kvantil F (1-alfa, m-1, n-m) :	3,0555683
Pravděpodobnost :	6,6423745 E-010
Závěr :	Model je významný

Scottovo kritérium multikolinearity

Hodnota kritéria SC :	-0,04844046748
Závěr :	Model je korektní.

Cook-Weisbergův test heteroskedasticity

Hodnota kritéria CW :	0,01086830286
Kvantil Chi <sup>2</sup> (1-alfa,1) :	3,841458829

Pravděpodobnost : 0,916970041  
Závěr : Rezidua vykazují homoskedasticitu.

Jarque-Berrův test normality  
Hodnota kritéria JB : 1,100431721  
Kvantil  $\chi^2(1-\alpha,2)$  : 5,991464547  
Pravděpodobnost : 0,5768252832  
Závěr : Rezidua mají normální rozdělení.

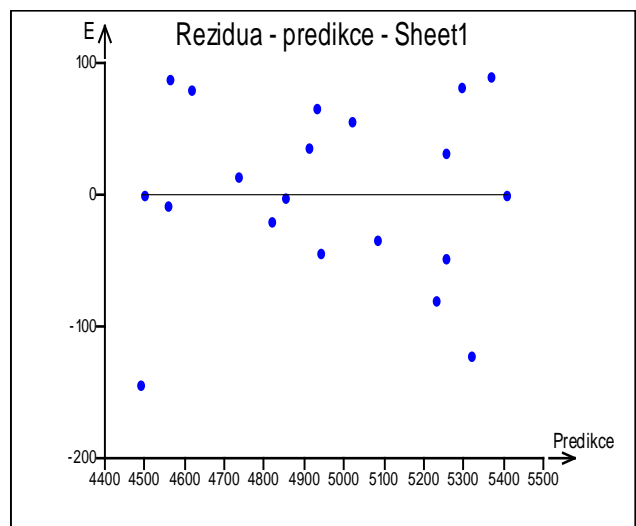
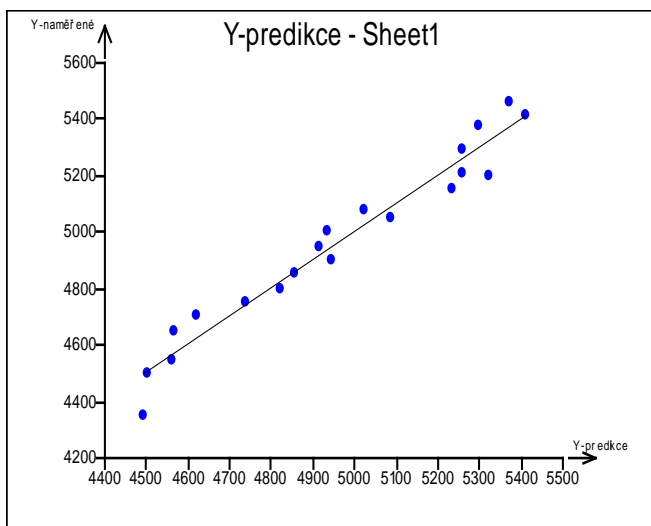
Waldův test autokorelace  
Hodnota kritéria WA : 0,05633998912  
Kvantil  $\chi^2(1-\alpha,1)$  : 3,841458829  
Pravděpodobnost : 0,8123771597  
Závěr : Autokorelace je nevýznamná.

Durbin-Watsonův test autokorelace  
Hodnota kritéria DW : -1  
Kritické hodnoty DW : 0,79 1,99  
Závěr : Pozitivní autokorelace reziduí není prokázána.

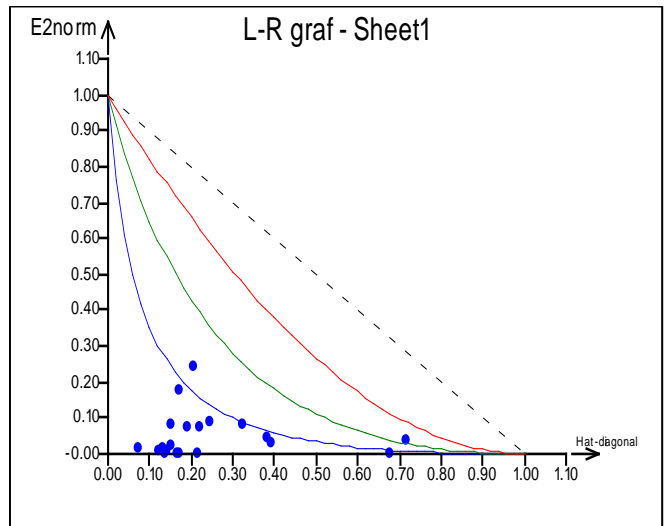
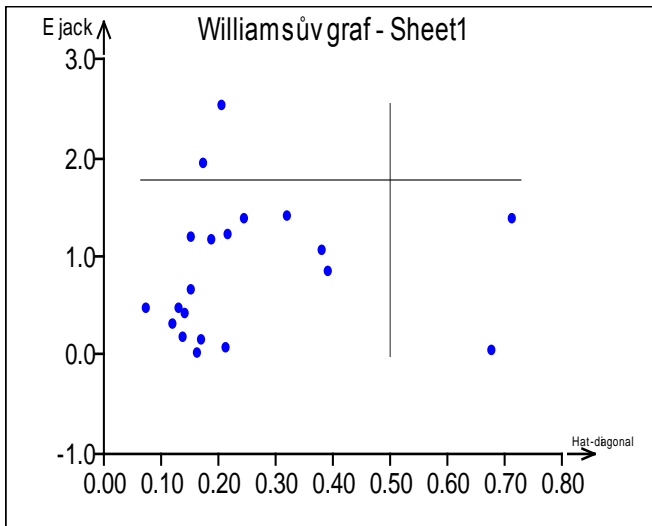
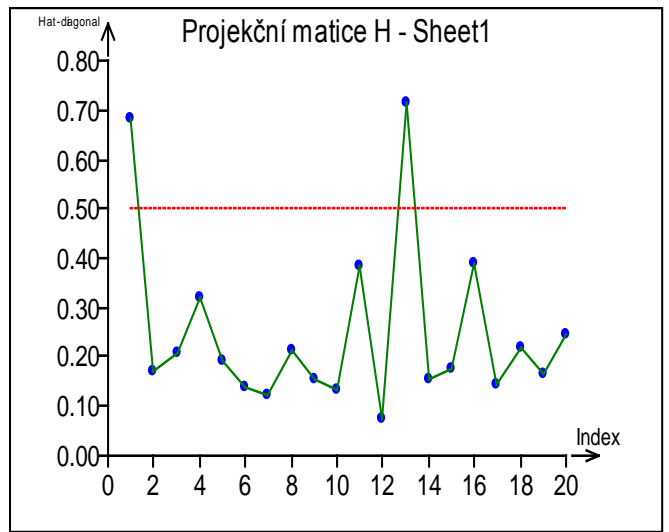
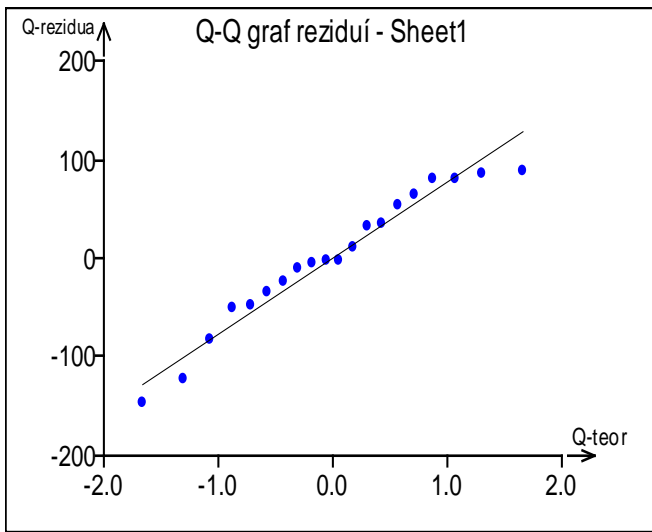
Znaménkový test reziduí  
Hodnota kritéria Sg : 0,1857480192  
Kvantil  $N(1-\alpha/2)$  : 1,959963999  
Pravděpodobnost : 0,8526423697  
Závěr : V reziduích není trend.

## 7. Grafická část:

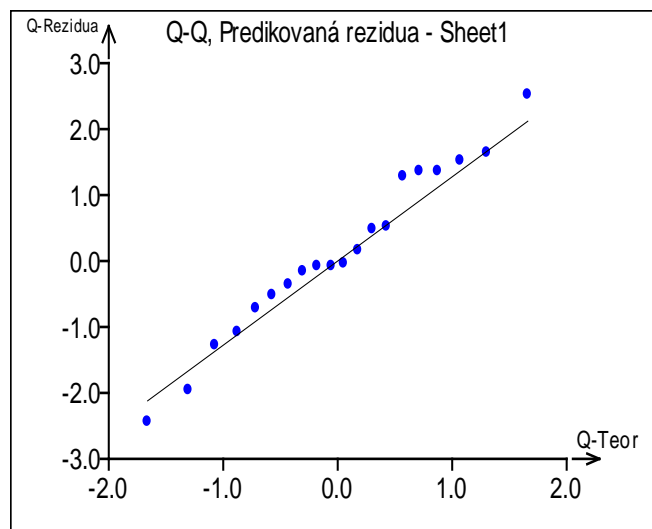
### Rezidua:



**Vlivná data:**



**Rankitový graf:**



## 8. Závěr:

Aplikací vícerozměrné lineární regrese na experimentální data mapující vliv čtyř parametrů (věk, váha, denní přísun kalorií a obsah tuků v nich) na hladinu totálního cholesterolu na vzorku 20 pacientů se zjistilo, že hlavní podíl na nárůst obsahu cholesterolu má denní přísun kalorií a obsah tuků v nich přítomných. Vliv věku a hmotnosti je nevýznamný.

Lineární model  $[y=(4200\pm 200)+(0,9\pm 0,3)*\text{Kalorie}+(19\pm 7)*\text{Tuky}]$  korektně popisuje zkoumaný jev. Pro snížení hladiny totálního cholesterolu se na základě výsledků doporučuje pacientům kontrolovat zejména denní přísun kalorií a podíl tuků v nich.