

Semestrální práce

2.1 Tvorba lineárních regresních modelů při analýze dat

2.3 Kalibrace a limity její přesnosti

Ing. Ján Lengyel, CSc.
Centrální analytická laboratoř
Ústav jaderného výzkumu Řež, a. s.
Husinec – Řež 130
250 68 Řež

V Řeži, leden 2009

2.1 Tvorba lineárních regresních modelů při analýze dat

Úloha 1. Porovnání tří regresních přímek u jednoduchého lineárního regresního modelu

Zadání:

Pro stanovení bakteriálních endotoxinů v léčivých látkách a přípravcích se používá metoda KQCL, založená na fotometrickém sledování kinetiky vzniku žlutého zbarvení lyzátu (*Limulus Amebocyte Lysat* – LAL) s gram - negativním endotoxinem za vzniku žlutého p-nitroanilinu. Pomocí kalibračních přímek se určuje obsah bakteriálních endotoxinů v léčivých látkách a přípravcích.

Úkolem je ověřit shodu nebo neshodu tří kalibračních přímek - lineárně regresních modelů. Předepsaná míra linearity kalibračních přímek je určena hodnotou korelačního koeficientu $\geq 0,980$.

Data:

Naměřená data ze tří setů, jedná se tudíž o tři kalibrační přímky, jsou uvedena v Tabulce 1.

Tabulka 1 Koncentrace lyzátu c v jednotkách IU/mg a reakčních časů v jednotkách s.

Koncentrace standardu endotoxinu, c , v EU/mg	Reakční čas, t , v s, 1. set	Reakční čas, t , v s, 2. set	Reakční čas, t , v s, 3. set
0,005	3838	3565	3837
0,05	2418	2432	2426
0,5	1233	1349	1307
5	886	729	764
50	390	454	466

Závislost $\log t$ na $\log c$ má být lineární, obecně $y = \beta_0 + \beta_1 x$ a údaje jsou uvedeny v Tabulce 2.

Tabulka 2. Naměřená data závislosti logaritmu koncentrace lyzátu na logaritmu reakčního času pro tři sety KQCL souprav.

$\log c$	$\log t$, 1. set	$\log t$, 2. set	$\log t$, 3. set
- 2,301029996	3,584105	3,55206	3,583992
- 1,301029996	3,383456	3,385964	3,384891
- 0,301029996	3,090963	3,130012	3,116276
0,698970004	2,947434	2,862728	2,883093
1,698970004	2,591065	2,657056	2,668386

Užitý program:

Pro odhad parametrů β_{0j} , β_{1j} a $\hat{\sigma}_j^2$ regresních přímek a určení vlivných a odlehlých bodů byl použit program ADSTAT, modul Lineární regrese. Pro porovnání shody úseků úseček, směrnic a určení totožnosti kalibračních přímek závislosti $\log c = f(\log t)$ byl použit program MS Excel Příloha 1 s názvem „porovnání kalibračních přímek pro BE“.

Výstup:

1. Odhady statistických parametrů, např. β_{0j} , β_{1j} , RSC a $\hat{\sigma}_j^2$, z výstupů programu ADSTAT pro:
první KQCL set

V Ý S T U P

(1) PŘEDBĚŽNÁ STATISTICKÁ ANALÝZA:

Proměnná	Průměr	Směrodatná odchylka	Párový korelační koeficient	Spočtená hladina výz.
y	3.1194E+00	3.8574E-01	1.0000	-----
x1	-3.0103E-01	1.5811E+00	-0.9928	0.001

Párové korelační koeficienty mezi dvojicemi vysvětlujících proměnných

Spočtená hladina významnosti

(2) INDIKACE MULTIKOLINEARITY:

Č [j]	Vlastní čísla korel. matice I[j]	Čísla podmí- něnosti K[j]	Variance inflation factor VIF[j]	Vícenás.korel. koef pro X[j]
1	1.0000E+00	1.0000E+00	1.0000E+00	0.0000
Maximální číslo podmíněnosti K (K[j], K > 1000 indikuje silnou multikolaritu)				: 1.0000E+00
(VIF[j] > 10 indikuje silnou multikolaritu)				

(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	3.0465E+00	2.4391E-02	1.2490E+02	Zamítnuta	0.000
BI 11	-2.4221E-01	1.6869E-02	-1.4358E+01	Zamítnuta	0.001

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R	: 9.9280E-01
Koeficient determinace, R ²	: 9.8566E-01
Predikovaný korelační koeficient, Rp ²	: 9.7723E-01
Střední kvadratická chyba predikce, MEP	: 5.3584E-03
Akaikeho informační kritérium, AIC	: -2.7864E+01

(5) ANALÝZA KLASICKÝCH REZIDUÍ:

Bod i	Meřená hodnota yexp[i]	Predikovaná hodnota yvyp[i]	Směrodatná odchylka s(yvyp[i])	Klasické reziduum e[i]	Relativní reziduum er[i]
1	3.5841E+00	3.6038E+00	4.1322E-02	-1.9720E-02	-5.5021E-01
2	3.3835E+00	3.3616E+00	2.9219E-02	2.1841E-02	6.4553E-01
3	3.0910E+00	3.1194E+00	2.3857E-02	-2.8442E-02	-9.2016E-01
4	2.9474E+00	2.8772E+00	2.9219E-02	7.0240E-02	2.3831E+00
5	2.5911E+00	2.6350E+00	4.1322E-02	-4.3919E-02	-1.6950E+00

Rezidualní součet čtverců, RSC	: 8.5373E-03
Průměr absolutních hodnot reziduí, Me	: 3.6832E-02
Průměr relativních reziduí, Mer	: 1.2388E+00
Odhad reziduálního rozptylu, s ² (e)	: 2.8458E-03
Odhad směrodatné odchylky reziduí, s(e)	: 5.3346E-02
Odhad šikmosti reziduí, g1(e)	: 6.8475E-01
Odhad špičatosti reziduí, g2(e)	: 1.9959E+00

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

Fisher-Snedocorův test významnosti regrese, F	: 2.0615E+02
Tabulkový kvantil, $F(1-\alpha, m-1, n-m)$: 1.0128E+01
Závěr: Navržený model je přijat jako významný.	
Spočtená hladina významnosti	: 0.001
Scottovo kritérium multikolinearity, M	: -4.6629E-15
Závěr: Navržený model je korektní.	
Cook-Weisbergův test heteroskedasticity, Sf	: 1.8899E+00
Tabulkový kvantil, $\chi^2(1-\alpha, 1)$: 3.8415E+00
Závěr: Rezidua vykazují homoskedasticitu.	
Spočtená hladina významnosti	: 0.169
Jarque-Berraův test normality reziduí, L(e)	: 6.0079E-01
Tabulkový kvantil, $\chi^2(1-\alpha, 2)$: 5.9915E+00
Závěr: Normalita je přijata.	
Spočtená hladina významnosti	: 0.741
Waldův test autokorelace, Wa	: 3.1156E+01
Tabulkový kvantil, $\chi^2(1-\alpha, 1)$: 3.8415E+00
Závěr: Rezidua jsou autokorelována.	
Spočtená hladina významnosti	: 0.000
Znamékový test, Dt	: 1.2002E+00
Tabulkový kvantil, $N(1-\alpha/2)$: 1.6449E+00
Závěr: Rezidua nevykazují trend.	
Spočtená hladina významnosti	: 0.115

(7) INDIKACE VLIVNÝCH BODŮ:

(* indikuje odlehlý nebo vlivný bod)

Bod	Standardizované reziduum	Jackknife reziduum	Predikované reziduum	Diagonální prvky
i	$eS[i]$	$eJ[i]$	$eP[i]$	$H[i, i]$
1	-5.8449E-01	-5.0697E-01	-4.9300E-02	6.0000E-01
2	4.8936E-01	4.1653E-01	3.1202E-02	3.0000E-01
3	-5.9609E-01	-5.1837E-01	-3.5552E-02	2.0000E-01
4	1.5737E+00	3.0765E+00	1.0034E-01	3.0000E-01
5	-1.3017E+00	-1.6112E+00	-1.0980E-01	6.0000E-01
Bod	Zobecněné diag. prvky	Cookova vzdálenost	Atkinsonova vzdálenost	Vliv na predikci
i	$Hm[i, i]$	$D[i]$	$A[i]$	$DF[i]$
1	6.4555E-01	2.5622E-01*	7.6046E-01	-6.2091E-01
2	3.5588E-01	5.1315E-02	3.3397E-01	2.7268E-01
3	2.9475E-01	4.4415E-02	3.1743E-01	-2.5918E-01
4	8.7789E-01	5.3071E-01*	2.4667E+00*	2.0140E+00*
5	8.2594E-01	1.2709E+00*	2.4168E+00*	-1.9733E+00*
Bod	V ě r o h o d n o s t n í v z d á l e n o s t i			
i	$LD(b)[i]$	$LD(s^2)[i]$	$LD(b, s^2)[i]$	
1	7.8849E-01	2.5267E-02	7.9632E-01	
2	1.6819E-01	4.6761E-02	1.9547E-01	
3	1.4590E-01	2.2818E-02	1.5717E-01	
4	1.5146E+00	1.0314E+01*	1.8427E+01*	
5	3.0685E+00	1.1475E+00	8.9356E+00*	

druhý KQCL set

V Ý S T U P

(1) PŘEDBĚŽNÁ STATISTICKÁ ANALÝZA:

Proměnná	Průměr	Směrodatná odchylna	Párový korelační koeficient	Spočtená hladina výz.
y	3.1176E+00	3.6674E-01	1.0000	-----
x1	-3.0103E-01	1.5811E+00	-0.9973	0.000

Párové korelační koeficienty mezi dvojicemi vysvětlujících proměnných

Spočtená hladina významnosti

(2) INDIKACE MULTIKOLINEARITY:

Č [j]	Vlastní čísla korel. matice l[j]	Čísla podmí- něnosti K[j]	Variance inflation factor VIF[j]	Vícenás.korel. koef pro X[j]
1	1.0000E+00	1.0000E+00	1.0000E+00	0.0000
Maximální číslo podmíněnosti K (K[j], K > 1000 indikuje silnou multikolaritu) (VIF[j] > 10 indikuje silnou multikolaritu)				: 1.0000E+00

(3) ODHADY PARAMETRŮ A TESTY VÝZNAMNOSTI:

Parametr	Odhad	Směrodatná odchylna	Test H0: B[j] = 0 vs. HA: B[j] <> 0 t-kriterium	hypoteza H0 je	Hlad. výz.
B[0]	3.0479E+00	1.4166E-02	2.1516E+02	Zamítnuta	0.000
B[1]	-2.3132E-01	9.7971E-03	-2.3611E+01	Zamítnuta	0.000

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R	: 9.9732E-01
Koeficient determinace, R ²	: 9.9465E-01
Predikovaný korelační koeficient, R _p ²	: 9.9145E-01
Střední kvadratická chyba predikce, MEP	: 1.8316E-03
Akaikeho informační kritérium, AIC	: -3.3298E+01

(5) ANALÝZA KLASICKÝCH REZIDUÍ:

Bod	Meřená hodnota	Predikovaná hodnota	Směrodatná odchylna	Klasické reziduum	Relativní reziduum
i	yexp[i]	yyp[i]	s(yyp[i])	e[i]	er[i]
1	3.5521E+00	3.5802E+00	2.3998E-02	-2.8153E-02	-7.9258E-01
2	3.3860E+00	3.3489E+00	1.6969E-02	3.7076E-02	1.0950E+00
3	3.1300E+00	3.1176E+00	1.3855E-02	1.2448E-02	3.9770E-01
4	2.8627E+00	2.8862E+00	1.6969E-02	-2.3512E-02	-8.2130E-01
5	2.6571E+00	2.6549E+00	2.3998E-02	2.1408E-03	8.0569E-02

Rezidualní součet čtverců, RSC	: 2.8795E-03
Průměr absolutních hodnot reziduí, Me	: 2.0666E-02
Průměr relativních reziduí, Mer	: 6.3742E-01
Odhad reziduálního rozptylu, s ² (e)	: 9.5984E-04
Odhad směrodatné odchylny reziduí, s(e)	: 3.0981E-02
Odhad šikmosti reziduí, g1(e)	: 2.5458E-01
Odhad špičatosti reziduí, g2(e)	: 1.7170E+00

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

Fisher-Snedocorův test významnosti regrese, F	: 5.5750E+02
Tabulkový kvantil, $F(1-\alpha, m-1, n-m)$: 1.0128E+01
Závěr: Navržený model je přijat jako významný.	
Spočtená hladina významnosti	: 0.000
Scottovo kritérium multikolinearity, M	: -1.1102E-15
Závěr: Navržený model je korektní.	
Cook-Weisbergův test heteroskedasticity, Sf	: 1.6787E+00
Tabulkový kvantil, $\chi^2(1-\alpha, 1)$: 3.8415E+00
Závěr: Rezidua vykazují homoskedasticitu.	
Spočtená hladina významnosti	: 0.195
Jarque-Berraův test normality reziduí, L(e)	: 3.9695E-01
Tabulkový kvantil, $\chi^2(1-\alpha, 2)$: 5.9915E+00
Závěr: Normalita je přijata.	
Spočtená hladina významnosti	: 0.820
Waldův test autokorelace, Wa	: 5.7775E-01
Tabulkový kvantil, $\chi^2(1-\alpha, 1)$: 3.8415E+00
Závěr: Rezidua nejsou autokorelována.	
Spočtená hladina významnosti	: 0.447
Znamkový test, Dt	: 1.0911E-01
Tabulkový kvantil, $N(1-\alpha/2)$: 1.6449E+00
Závěr: Rezidua nevykazují trend.	
Spočtená hladina významnosti	: 0.457

(7) INDIKACE VLIVNÝCH BODŮ:

(* indikuje odlehlý nebo vlivný bod)

Bod	Standardizované reziduum	Jackknife reziduum	Predikované reziduum	Diagonální prvky
i	$eS[i]$	$eJ[i]$	$eP[i]$	$H[i, i]$
1	-1.4368E+00	-2.1007E+00	-7.0382E-02	6.0000E-01
2	1.4303E+00	2.0709E+00	5.2965E-02	3.0000E-01
3	4.4922E-01	3.7978E-01	1.5560E-02	2.0000E-01
4	-9.0706E-01	-8.6935E-01	-3.3588E-02	3.0000E-01
5	1.0925E-01	8.9384E-02	5.3519E-03	6.0000E-01
Bod	Zobecněné diag. prvky	Cookova vzdálenost	Atkinsonova vzdálenost	Vliv na predikci
i	$Hm[i, i]$	$D[i]$	$A[i]$	$DF[i]$
1	8.7525E-01	1.5483E+00*	3.1510E+00*	-2.5728E+00*
2	7.7737E-01	4.3840E-01*	1.6604E+00*	1.3557E+00*
3	2.5381E-01	2.5225E-02	2.3257E-01	1.8989E-01
4	4.9197E-01	1.7630E-01*	6.9703E-01	-5.6912E-01
5	6.0159E-01	8.9524E-03	1.3408E-01	1.0947E-01
Bod	V ě r o h o d n o s t n í v z d á l e n o s t i			
i	$LD(b)[i]$	$LD(s^2)[i]$	$LD(b, s^2)[i]$	
1	3.5456E+00	3.1155E+00	1.6354E+01*	
2	1.2820E+00	2.9649E+00	6.6408E+00	
3	8.3383E-02	5.6010E-02	1.2813E-01	
4	5.5563E-01	2.4509E-02	6.7231E-01	
5	2.9753E-02	1.1176E-01	1.3573E-01	

třetí KQCL set

V Ý S T U P

(1) PŘEDBĚŽNÁ STATISTICKÁ ANALÝZA:

Proměnná	Průměr	Směrodatná odchylka	Párový korelační koeficient	Spočtená hladina výz.
y	3.1273E+00	3.6922E-01	1.0000	-----
x1	-3.0103E-01	1.5811E+00	-0.9991	0.000

Párové korelační koeficienty mezi dvojicemi
vysvětlujících proměnných

Spočtená
hladina významnosti

(2) INDIKACE MULTIKOLINEARITY:

Č [j]	Vlastní čísla korel. matice l[j]	Čísla podmí- něnosti K[j]	Variance inflation factor VIF[j]	Vícenás.korel. koef pro X[j]
1	1.0000E+00	1.0000E+00	1.0000E+00	0.0000
Maximální číslo podmíněnosti K (K[j], K > 1000 indikuje silnou multikolaritu) (VIF[j] > 10 indikuje silnou multikolaritu)				: 1.0000E+00

(3) ODHADY PARAMETRŮ A TESTY VÝZNAMNOSTI:

Parametr	Odhad	Směrodatná odchylka	Test H0: B[j] = 0 us. t-kriterium	HA: B[j] <> 0 hypoteza H0 je	Hlad. výz.
BI 01	3.0571E+00	8.2967E-03	3.6847E+02	Zamítnuta	0.000
BI 11	-2.3330E-01	5.7381E-03	-4.0658E+01	Zamítnuta	0.000

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R	: 9.9909E-01
Koeficient determinace, R ²	: 9.9819E-01
Predikovaný korelační koeficient, Rp ²	: 9.9759E-01
Střední kvadratická chyba predikce, MEP	: 5.2403E-04
Akaikého informační kritérium, AIC	: -3.8647E+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	3.5840E+00	3.5939E+00	1.4055E-02	-9.9376E-03	-2.7728E-01
2	3.3849E+00	3.3606E+00	9.9387E-03	2.4262E-02	7.1679E-01
3	3.1163E+00	3.1273E+00	8.1149E-03	-1.1052E-02	-3.5464E-01
4	2.8831E+00	2.8940E+00	9.9387E-03	-1.0934E-02	-3.7923E-01
5	2.6684E+00	2.6607E+00	1.4055E-02	7.6603E-03	2.8708E-01

Rezidualní součet čtverců, RSC	: 9.8778E-04
Průměr absolutních hodnot reziduí, Me	: 1.2769E-02
Průměr relativních reziduí, Mer	: 4.0300E-01
Odhad reziduálního rozptylu, s ² (e)	: 3.2926E-04
Odhad směrodatné odchylky reziduí, s(e)	: 1.8146E-02
Odhad šiklosti reziduí, g1(e)	: 7.9904E-01
Odhad špičatosti reziduí, g2(e)	: 1.9931E+00

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

Fisher-Snedocorův test významnosti regrese, F	: 1.6531E+03
Tabulkový kvantil, $F(1-\alpha, m-1, n-m)$: 1.0128E+01
Závěr: Navržený model je přijat jako významný.	
Spočtená hladina významnosti	: 0.000
Scottovo kritérium multikolinearity, M	: 6.5503E-15
Závěr: Navržený model je korektní.	
Cook-Weisbergův test heteroskedasticity, Sf	: 1.1772E+00
Tabulkový kvantil, $\chi^2(1-\alpha, 1)$: 3.8415E+00
Závěr: Rezidua vykazují homoskedasticitu.	
Spočtená hladina významnosti	: 0.278
Jarque-Berraův test normality reziduí, L(e)	: 7.4329E-01
Tabulkový kvantil, $\chi^2(1-\alpha, 2)$: 5.9915E+00
Závěr: Normalita je přijata.	
Spočtená hladina významnosti	: 0.690
Waldův test autokorelace, Wa	: 1.7410E+00
Tabulkový kvantil, $\chi^2(1-\alpha, 1)$: 3.8415E+00
Závěr: Rezidua nejsou autokorelována.	
Spočtená hladina významnosti	: 0.187
Znaménkový test, Dt	: 1.0911E-01
Tabulkový kvantil, $N(1-\alpha/2)$: 1.6449E+00
Závěr: Rezidua nevykazují trend.	
Spočtená hladina významnosti	: 0.457

(7) INDIKACE VLIVNÝCH BODŮ:

(* indikuje odlehlý nebo vlivný bod)

Bod	Standardizované reziduum	Jackknife reziduum	Predikované reziduum	Diagonální prvky
i	$eS[i]$	$eJ[i]$	$eP[i]$	$H[i, i]$
1	-8.6593E-01	-8.1638E-01	-2.4844E-02	6.0000E-01
2	1.5981E+00	3.3845E+00	3.4661E-02	3.0000E-01
3	-6.8094E-01	-6.0468E-01	-1.3815E-02	2.0000E-01
4	-7.2018E-01	-6.4657E-01	-1.5619E-02	3.0000E-01
5	6.6750E-01	5.9063E-01	1.9151E-02	6.0000E-01
Bod	Zobecněné diag. prvky	Cookova vzdálenost	Atkinsonova vzdálenost	Vliv na predikci
i	$H_m[i, i]$	$D[i]$	$A[i]$	$DF[i]$
1	6.9998E-01	5.6238E-01*	1.2246E+00	-9.9985E-01
2	8.9595E-01	5.4730E-01*	2.7136E+00*	2.2156E+00*
3	3.2365E-01	5.7960E-02	3.7029E-01	-3.0234E-01
4	4.2102E-01	1.1114E-01*	5.1841E-01	-4.2328E-01
5	6.5941E-01	3.3416E-01*	8.8594E-01	7.2337E-01
Bod	V ě r o h o d n o s t n í v z d á l e n o s t i			
i	$LD(b)[i]$	$LD(s^2)[i]$	$LD(b, s^2)[i]$	
1	1.5920E+00	1.0616E-02	2.0100E+00	
2	1.5553E+00	1.3494E+01*	2.3312E+01*	
3	1.8956E-01	7.4912E-03	1.9031E-01	
4	3.5739E-01	2.7467E-03	3.6107E-01	
5	1.0056E+00	9.5254E-03	1.0561E+00	

první, druhý a třetí KQCL set dohromady

V Ý S T U P

(1) PŘEDBĚŽNÁ STATISTICKÁ ANALÝZA:

Proměnná	Průměr	Směrodatná odchylna	Párový korelační koeficient	Spočtená hladina výz.
y	3.1214E+00	3.4628E-01	1.0000	-----
x1	-3.0103E-01	1.4639E+00	-0.9960	0.000

Párové korelační koeficienty mezi dvojicemi vysvětlujících proměnných

Spočtená hladina významnosti

(2) INDIKACE MULTIKOLINEARITY:

Č [j]	Vlastní čísla korel. matice I[j]	Čísla podmí- něnosti K[j]	Variance inflation factor VIF[j]	Vícenás. korel. koef pro X[j]
1	1.0000E+00	1.0000E+00	1.0000E+00	0.0000

Maximální číslo podmíněnosti K : 1.0000E+00
 (K[j], K > 1000 indikuje silnou multikolaritu)
 (VIF[j] > 10 indikuje silnou multikolaritu)

(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	3.0505E+00	8.4584E-03	3.6065E+02	Zamítnuta		0.000
BI 11	-2.3561E-01	5.8499E-03	-4.0276E+01	Zamítnuta		0.000

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R	: 9.9602E-01
Koeficient determinace, R ²	: 9.9205E-01
Predikovaný korelační koeficient, Rp ²	: 9.9455E-01
Střední kvadratická chyba predikce, MEP	: 1.2165E-03
Akaikeho informační kritérium, AIC	: -1.0137E+02

(5) ANALÝZA KLASICKÝCH REZIDUÍ:

Bod i	Meřená hodnota y _{exp[i]}	Predikovaná hodnota y _{yyp[i]}	Směrodatná odchylna s(y _{yyp[i]})	Klasické reziduum e _[i]	Relativní reziduum er _[i]
1	3.5841E+00	3.5927E+00	1.4329E-02	-8.5508E-03	-2.3857E-01
2	3.5521E+00	3.5927E+00	1.4329E-02	-4.0596E-02	-1.1429E+00
3	3.5840E+00	3.5927E+00	1.4329E-02	-8.6638E-03	-2.4174E-01
4	3.3835E+00	3.3570E+00	1.0132E-02	2.6412E-02	7.8062E-01
5	3.3860E+00	3.3570E+00	1.0132E-02	2.8920E-02	8.5411E-01
6	3.3849E+00	3.3570E+00	1.0132E-02	2.7847E-02	8.2269E-01
7	3.0910E+00	3.1214E+00	8.2730E-03	-3.0469E-02	-9.8575E-01
8	3.1300E+00	3.1214E+00	8.2730E-03	8.5800E-03	2.7412E-01
9	3.1163E+00	3.1214E+00	8.2730E-03	-5.1560E-03	-1.6546E-01
10	2.9474E+00	2.8858E+00	1.0132E-02	6.1614E-02	2.0904E+00
11	2.8627E+00	2.8858E+00	1.0132E-02	-2.3092E-02	-8.0665E-01
12	2.8831E+00	2.8858E+00	1.0132E-02	-2.7271E-03	-9.4589E-02
13	2.5911E+00	2.6502E+00	1.4329E-02	-5.9143E-02	-2.2826E+00
14	2.6571E+00	2.6502E+00	1.4329E-02	6.8477E-03	2.5772E-01
15	2.6684E+00	2.6502E+00	1.4329E-02	1.8178E-02	6.8122E-01

Reziduální součet čtverců, RSC	: 1.3346E-02
Průměr absolutních hodnot reziduí, Me	: 2.3786E-02
Průměr relativních reziduí, Mer	: 7.8128E-01
Odhad reziduálního rozptylu, $s^2(e)$: 1.0266E-03
Odhad směrodatné odchylky reziduí, $s(e)$: 3.2041E-02
Odhad šikmosti reziduí, $g1(e)$: -2.7001E-02
Odhad špičatosti reziduí, $g2(e)$: 2.7305E+00
(6) TESTOVÁNÍ REGRESNÍHO TRIPLETU (DATA + MODEL + METODA):	
Fisher-Snedocorův test významnosti regrese, F	: 1.6222E+03
Tabulkový kvantil, $F(1-\alpha, m-1, n-m)$: 4.6672E+00
Závěr: Navržený model je přijat jako významný.	
Spočtená hladina významnosti	: 0.000
Scottovo kritérium multikolinearity, M	: 1.7764E-15
Závěr: Navržený model je korektní.	
Cook-Weisbergův test heteroskedasticity, Sf	: 4.2250E+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)	: 4.7203E-02
Tabulkový kvantil, $\chi^2(1-\alpha, 2)$: 5.9915E+00
Závěr: Normalita je přijata.	
Spočtená hladina významnosti	: 0.977
Waldův test autokorelace, Wa	: 7.3955E-02
Tabulkový kvantil, $\chi^2(1-\alpha, 1)$: 3.8415E+00
Závěr: Rezidua nejsou autokorelována.	
Spočtená hladina významnosti	: 0.786
Znamkový test, Dt	: 1.7949E-02
Tabulkový kvantil, $N(1-\alpha/2)$: 1.6449E+00
Závěr: Rezidua nevykazují trend.	
Spočtená hladina významnosti	: 0.493

(7) INDIKACE VLIVNÝCH BODŮ:

(* indikuje odlehlý nebo vlivný bod)

Bod	Standardizované reziduum eSIiI	Jackknife reziduum eJIiI	Predikované reziduum ePIiI	Diagonální prvky HIi, iI
1	-2.9837E-01	-2.8765E-01	-1.0688E-02	2.0000E-01
2	-1.4165E+00	-1.4800E+00	-5.0745E-02	2.0000E-01
3	-3.0231E-01	-2.9148E-01	-1.0830E-02	2.0000E-01
4	8.6890E-01	8.6017E-01	2.9347E-02	1.0000E-01
5	9.5141E-01	9.4767E-01	3.2133E-02	1.0000E-01
6	9.1611E-01	9.1004E-01	3.0941E-02	1.0000E-01
7	-9.8431E-01	-9.8304E-01	-3.2646E-02	6.6667E-02
8	2.7718E-01	2.6709E-01	9.1928E-03	6.6667E-02
9	-1.6657E-01	-1.6020E-01	-5.5243E-03	6.6667E-02
10	2.0270E+00	2.3548E+00*	6.8460E-02	1.0000E-01
11	-7.5968E-01	-7.4664E-01	-2.5658E-02	1.0000E-01
12	-8.9716E-02	-8.6223E-02	-3.0301E-03	1.0000E-01
13	-2.0637E+00	-2.4180E+00*	-7.3929E-02	2.0000E-01
14	2.3894E-01	2.3007E-01	8.5597E-03	2.0000E-01
15	6.3428E-01	6.1905E-01	2.2722E-02	2.0000E-01

Bod	Zobecněné diag. prvky	Cookova vzdálenost	Atkinsonova vzdálenost	Vliv na predikci
i	Hm[i, i]	DI[i]	AI[i]	DF[i]
1	2.0548E-01	1.1128E-02	3.6668E-01	-1.4382E-01
2	3.2348E-01	2.5082E-01*	1.8866E+00*	-7.3998E-01*
3	2.0562E-01	1.1424E-02	3.7156E-01	-1.4574E-01
4	1.5227E-01	4.1944E-02	7.3100E-01	2.8672E-01
5	1.6267E-01	5.0288E-02	8.0537E-01	3.1589E-01
6	1.5810E-01	4.6626E-02	7.7338E-01	3.0335E-01
7	1.3623E-01	3.4602E-02	6.6983E-01	-2.6273E-01
8	7.2182E-02	2.7438E-03	1.8199E-01	7.1384E-02
9	6.8659E-02	9.9087E-04	1.0916E-01	-4.2816E-02
10	3.8444E-01	2.2825E-01*	2.0012E+00*	7.8493E-01*
11	1.3995E-01	3.2062E-02	6.3452E-01	-2.4888E-01
12	1.0056E-01	4.4716E-04	7.3275E-02	-2.8741E-02
13	4.6209E-01*	5.3237E-01*	3.0824E+00*	-1.2090E+00*
14	2.0351E-01	7.1366E-03	2.9329E-01	1.1504E-01
15	2.2476E-01	5.0289E-02	7.8914E-01	3.0953E-01
Bod	V ě r o h o d n o s t n í v z d á l e n o s t i			
i	LD(b)[i]	LD(s ²)[i]	LD(b, s ²)[i]	
1	2.5658E-02	2.8353E-02	5.2486E-02	
2	5.6793E-01	7.5462E-02	7.1430E-01	
3	2.6340E-02	2.8187E-02	5.2967E-02	
4	9.6483E-02	6.2764E-04	9.6539E-02	
5	1.1560E-01	7.5874E-05	1.1649E-01	
6	1.0721E-01	3.8147E-05	1.0739E-01	
7	7.9640E-02	5.3815E-04	8.1068E-02	
8	6.3306E-03	2.9212E-02	3.5157E-02	
9	2.2865E-03	3.2789E-02	3.4927E-02	
10	5.1770E-01	8.0613E-01	1.5249E+00	
11	7.3807E-02	4.1388E-03	7.6404E-02	
12	1.0319E-03	3.4276E-02	3.5240E-02	
13	1.1808E+00	9.0240E-01	2.6077E+00	
14	1.6460E-02	3.0627E-02	4.6066E-02	
15	1.1560E-01	1.0448E-02	1.2222E-01	

Řešení:

Řešení je založeno na porovnání tří (M=3) navržených regresních modelů:

$$y_{ij} = \beta_{0j} + \beta_{1j} \cdot x_{ij} + \varepsilon_{ij}, \quad j=1, \dots, M, \quad i=1, \dots, n_j, \quad (\text{rov.1}).$$

Otestuje se zda:

- regresní přímky mají společný úsek úseček,
- regresní přímky mají společnou směrnici,
- regresní přímky jsou totožné.

Prvním krokem statistické analýzy je odhad parametrů β_{0j} , β_{1j} a $\hat{\sigma}_j^2$ pro všechna data zvlášť metodou nejmenších čtverců. Odhady získané z programu ADSTAT jsou v Tab.3.

Tab.3 Odhady parametrů β_{0j} , β_{1j} a $\hat{\sigma}_j^2$ z programu ADSTAT pro jednotlivé KQCL sety

Set	β_{0j}	β_{1j}	RSC	$\hat{\sigma}_j^2$
1	3,0465	- 0,24221	0,0085373	0,0028458
2	3,0479	- 0,23132	0,0028795	0,00095984
3	3,0571	- 0,23330	0,00098778	0,00032926

Druhým krokem je ověření shody reziduálních rozptylů Bartlettovým testem s testačním kritériem podle vztahu:

$$B = \frac{V \cdot \ln \hat{\sigma}_c^2 - \sum_{j=1}^{M=3} v_j \cdot \ln \hat{\sigma}_j^2}{L} \quad (\text{rov.2}),$$

kde:

$$V = \sum_{j=1}^{M=3} v_j \quad (\text{rov.3}) \text{ a } v_j = (n_j - m) = 5 - 2 = 3 \text{ představuje stupně volnosti}$$

modelů regresních přímek,

$$\hat{\sigma}_c^2 = \frac{\sum_{j=1}^{M=3} v_j \cdot \hat{\sigma}_j^2}{V} \quad \text{představuje sdružený rozptyl} \quad (\text{rov.4})$$

$$L = 1 + \frac{\sum_{j=1}^{M=3} v_j^{-1} - V^{-1}}{3 \cdot M - 3} \quad (\text{rov.5})$$

V Tab.4 je uvedena hodnota testačního kritéria B, včetně pomocných výpočtů – MS Excel Příloha 1 s názvem „porovnání kalibračních přímek pro BE“, list testační kritérium B.

Tab.4 Výpočet testačního kritéria B

v_j	V	$\hat{\sigma}_j^2$		L	$V \ln \hat{\sigma}_c^2$		$v_j \ln \hat{\sigma}_j^2$	B
3	9	0,0028458	0,000949	1,148148148	-59,2821		-17,5857	2,79216
3	9	0,00095984	0,00032				-20,8462	
3	9	0,00032926	0,00011				-24,056	
		$\hat{\sigma}_c^2 =$	0,001378			$\Sigma =$	-62,488	

Kvantil rozdělení χ^2 pro $M - 1 = 3 - 1 = 2$ stupňů volnosti a pro hladinu významnosti $\alpha = 0,05$ má hodnotu 5,991. Platí, že $B < \chi^2_{(1-\alpha)}$ a proto platí nulová hypotéza o shodě rozptylů kalibračních přímek u testovaných KQCL setů a přijatelným odhadem rozptylu je sdružený rozptyl.

Nyní lze provést testování:

a) homogenity úseků

Ověřuje se nulová hypotéza $H_0: \beta_{01} = \beta_{02} = \beta_{03}$. Když platí H_0 , tak lze získat sdružený odhad úseku b_{0c} , váhová kombinace odhadů jednotlivých úseků b_0 , podle vztahu:

$$b_{0c} = \frac{\sum_{j=1}^{M=3} w_{Bj} \cdot b_{0j}}{\sum_{j=1}^{M=3} w_{Bj}} \quad (\text{rov.6}),$$

kde váhový koeficient w_{Bj} , odpovídající odhadu úseku j -té přímky, je dán vztahem:

$$w_{Bj} = \frac{n_j \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2}{\sum_{i=1}^{n_j} x_{ij}^2} \quad (\text{rov.7}).$$

Testační statistika F_I se počítá podle vztahu:

$$F_I = \frac{\frac{1}{M-1} \sum_{j=1}^{M=3} w_{Bj} \cdot (b_{0j} - b_{0c})^2}{\frac{1}{n-2M} \sum_{j=1}^{M=3} RSC_j} \quad (\text{rov.8}),$$

kde $n = \sum_{j=1}^{M=3} n_j$ (rov.9).

Platí-li nulová hypotéza $H_0: \beta_{01} = \beta_{02} = \beta_{03}$, potom musí platit, že $F_I < F_{(1-\alpha)}(M-1, n-2M)$ na hladině významnosti $\alpha = 0,05$, $F_{0,95}(2,9) = 4,26$. V Tab.5 jsou shrnuty výsledky výpočtů podle rovnic (rov.6) až (rov.10), v MS Excel Příloha 1 s názvem „porovnání kalibračních přímek pro BE“, list homogenita úseků. Na hladině významnosti $\alpha = 0,05$ mají všechny kalibrační přímky stejný úsek a jeho odhad je daný hodnotou b_{0c} a rozptyl tohoto úseku se vypočte podle vztahu:

$$D(b_{0c}) = \frac{\frac{1}{n-2M} \sum_{j=1}^{M=3} RSC_j}{\sum_{j=1}^{M=3} w_{Bj}} \quad (\text{rov.10}).$$

Tab.5 Výpočet testačního kritéria F_I

b_{0c}	$D(b_{0c})$	F_I
3,0505	9,5E-5	0,11582

Porovnáním vypočtené hodnoty F_I s kritickou hodnotou F-rozdělení vyplývá, že nulová hypotéza o homogenitě úseků platí a kalibrační přímky mají společný průsečík s osou y .

b) homogenity směrnic

Ověřuje se nulová hypotéza $H_0: \beta_{11} = \beta_{12} = \beta_{13}$. Když platí H_0 , tak lze získat sdružený odhad úseku b_{1c} , váhová kombinace odhadů jednotlivých směrnic b_{1j} , podle vztahu:

$$b_{1c} = \frac{\sum_{j=1}^{M=3} w_{Sj} b_{1j}}{\sum_{j=1}^{M=3} w_{Sj}} \quad (\text{rov.11}),$$

kde

$$w_{sj} = \sum_{i=1}^{n_j=5} (x_{ij} - \bar{x}_j)^2 \quad (\text{rov.12}).$$

Testační statistika F_S se počítá podle vztahu:

$$F_I = \frac{\frac{1}{M-1} \sum_{j=1}^{M=3} w_{sj} (b_{1j} - b_{1c})^2}{\frac{1}{n-2M} \sum_{j=1}^{M=3} RSC_j} \quad (\text{rov.13})$$

Platí-li nulová hypotéza $H_0: \beta_{11} = \beta_{12} = \beta_{13}$, potom musí platit, že

$F_S < F_{(1-\alpha)}(M-1, n-2M)$ na hladině významnosti $\alpha = 0,05$, $F_{0,95}(2,9) = 4,26$.

V Tab.6 jsou shrnuty výsledky výpočtů podle rovnic (rov.11) až (rov.14), v MS Excel Příloha 1 s názvem „porovnání kalibračních přímek pro BE“, list homogenita směrnic. Nejlepším odhadem celkové směrnice je b_{1c} a její rozptyl se vypočte podle vztahu:

$$D(b_{1c}) = \frac{\frac{1}{n-2M} \sum_{j=1}^{M=3} RSC_j}{\sum_{j=1}^{M=3} w_{sj}} \quad (\text{rov.14}).$$

Tab.6 Výpočet testačního kritéria F_S

b_{1c}	$D(b_{1c})$	F_S
-0,23561	4,56E-5	0,245719

Porovnáním vypočtené hodnoty F_S s kritickou hodnotou F-rozdělení vyplývá, že nulová hypotéza o homogenitě směrnic platí a kalibrační přímky jsou rovnoběžné.

c) shoda regresních přímek

Ověřuje se nulová hypotéza $H_0: \beta_{01} = \beta_{0c}$, $\beta_{1j} = \beta_{1c}$. To se provádí porovnáním reziduálního součtu čtverců RSC_K , získaného po proložení všech M skupin dat jedinou společnou přímkou s odhady parametrů b_{0K} a b_{1K} , a reziduálního součtu čtverců, vytvořeného pro každou skupinu zvlášť – testační kritérium, podle vztahu:

$$F_A = \frac{\frac{RSC_K - RSC_c}{2M-2}}{\frac{RSC_c}{n-2M}} \quad (\text{rov.15}),$$

kde

$$RSC_c = \sum_{j=1}^{M=3} RSC_j \quad (\text{rov.16}).$$

Platí-li nulová hypotéza $H_0: \beta_{01} = \beta_{0c}$, $\beta_{1j} = \beta_{1c}$, potom musí platit, že

$F_A < F_{(1-\alpha)}(2M-2, n-2M)$ na hladině významnosti $\alpha = 0,05$, $F_{0,95}(4,9) = 3,63$.

V Tab.7 jsou shrnuty výsledky výpočtů podle rovnic (rov.15) a (rov.16), v MS Excel Příloha 1 s názvem „porovnání kalibračních přímek pro BE“, list totožnost přímek.

Tab.7 Výpočet testačního kritéria F_A

RSC_c	RSC_K	F_A	b_{0K}	b_{1K}
0,01232508	0,013346	0,18637	3,0505 (0,0085)	-0,23561 (0,0058)

Porovnáním vypočtené hodnoty F_A s kritickou hodnotou F-rozdělení vyplývá, že nulová hypotéza o shodě regresních přímek platí a kalibrační přímky jsou totožné se společným odhadem úseku b_{0K} a směrnici b_{1K} na hladině významnosti $\alpha = 0,05$.

Závěr:

Statistickým testováním kalibračních přímek závislosti log reakčního času na log koncentrace standardu endotoxinu byla ověřena totožnost kalibračních přímek u tří KQCL setů. To dokazuje, že všechny tři KQCL sety by měly poskytovat shodné výsledky obsahu bakteriálních endotoxinů v léčivých látkách a přípravcích. Porovnáním hodnot odhadů úseků a směrníc z tabulek Tab. 5, Tab.6 a Tab.7 je vidět shodu hodnot těchto odhadů a kalibrační přímku lze vyjádřit vztahem:

$$\log t = 3,0305(0,0085) - 0,23561(0,0058) \log c \quad (\text{rov.17}).$$

U KQCL metody předepsaná podmínka linearity – korelační koeficient $\geq 0,98$, je ve všech případech splněna. Rovněž podmínka, že rezidua nevykazují trendy je ve všech případech splněna. Proto žádný indikovaný odlehlý bod nebyl vyloučen z testování.

Úloha 2. Určení stupně polynomu**Zadání:**

Při výrobě kyseliny sírové byl studován vliv teploty na stupeň přeměny - x, vztažený na oxid siřičitý. Úkolem je určit stupeň polynomu této závislosti jednak pomocí metody nejmenších čtverců (MŇČ) a jednak metodou racionálních hodnot (RH).

Data:

Data, uvedená v tabulce Tabulka 1, jsou převzatá z publikace E. Hála, A. Reiser: Fyzikální chemie, díl 1., Academia, Praha, 1971, str.

Tabulka 1 Vliv teploty T v K na stupeň přeměny x, vztažený na SO_2 , v reakci $SO_2(g) + \frac{1}{2} O_2(g) = SO_3(g)$, takzvaná stechiometrická směs.

T (K)	1232	1130	1066	1023	979	940	899	852	788	701
X	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	0,975

Užitý program:

Pro výpočty byl použitý program ADSTAT, modul lineární regrese.

Výstup:
Výstupy č. 1

V Ý S T U P

(1) ZVOLENÁ STRATEGIE REGRESNÍ ANALÝZY:

Omezení, P : 1.0000E-34
Transformace : Polynom
Stupeň polynomu : 3
Váhy : Ne
Absolutní člen zahrnut : Ano

(2) PODMÍNKY A KVANTILY PRO STATISTICKÉ TESTY:

Hladina významnosti, alfa : 0.050
Počet bodů, n : 10
Počet parametrů, m : 4
Kvantil Studentova rozdělení $t(1-\alpha/2, n-m)$: 2.447
Kvantil rozd. Chí-kvadrát $\text{Chi-square}(1-\alpha, m)$: 9.488
Jméno výstupního souboru : RESULTS.TXT

V Ý S T U P

(1) PŘEDBĚŽNÁ STATISTICKÁ ANALÝZA:

Proměnná	Průměr	Směrodatná odchylka	Párový korelační koeficient	Spočtená hladina výz.
y	5.4750E-01	2.9871E-01	1.0000	-----
x1	9.6100E+02	1.6017E+02	-0.9898	0.000
x2	9.4661E+05	3.1031E+05	-0.9856	0.000
x3	9.5425E+08	4.6249E+08	-0.9733	0.000

Párové korelační koeficienty mezi dvojicemi vysvětlujících proměnných		Spočtená hladina významnosti
x1 versus x2 :	9.9617E-01	0.000
x1 versus x3 :	9.8539E-01	0.000
x2 versus x3 :	9.9649E-01	0.000

(2) INDIKACE MULTIKOLINEARITY:

Č [j]	Vlastní čísla korel. matice I[j]	Čísla podmí- něnosti K[j]	Variance inflation factor VIF[j]	Vícenás. korel. koef pro X[j]
1	1.2336E-05	2.4200E+05	1.3033E+04	1.0000
2	1.4614E-02	2.0428E+02	5.3902E+04	1.0000
3	2.9854E+00	1.0000E+00	1.4197E+04	1.0000

Maximální číslo podmíněnosti K : 2.4200E+05
(K[j], $K > 1000$ indikuje silnou multikolaritu)
(VIF[j] > 10 indikuje silnou multikolaritu)

(3) ODHADY PARAMETRŮ A TESTY VÝZNAMNOSTI:

Parametr	Odhad	Směrodatná odchylka	Test $H_0: B[j] = 0$ vs. $H_A: B[j] \neq 0$	t-kriterium	hypoteza H_0 je	Hlad. výz
BI 01	-6.9268E+00	4.7966E-01	-1.4441E+01	Zamítnuta	0.000	
BI 11	2.8021E-02	1.5359E-03	1.8244E+01	Zamítnuta	0.000	
BI 21	-3.1526E-05	1.6123E-06	-1.9553E+01	Zamítnuta	0.000	
BI 31	1.0887E-08	5.5518E-10	1.9609E+01	Zamítnuta	0.000	

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:	
Vícenásobný korelační koeficient, R	: 9.9984E-01
Koeficient determinace, R^2	: 9.9969E-01
Predikovaný korelační koeficient, R_p^2	: 9.9870E-01
Střední kvadratická chyba predikce, MEP	: 2.0884E-04
Akaikeho informační kritérium, AIC	: -9.7936E+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	1.0000E-01	1.0231E-01	6.2701E-03	-2.3054E-03	-2.3054E+00
2	2.0000E-01	1.9016E-01	3.8555E-03	9.8442E-03	4.9221E+00
3	3.0000E-01	3.0683E-01	3.5735E-03	-6.8331E-03	-2.2777E+00
4	4.0000E-01	4.0142E-01	3.1259E-03	-1.4151E-03	-3.5377E-01
5	5.0000E-01	5.0537E-01	2.7757E-03	-5.3733E-03	-1.0747E+00
6	6.0000E-01	5.9915E-01	2.8218E-03	8.5362E-04	1.4227E-01
7	7.0000E-01	6.9495E-01	3.1776E-03	5.0494E-03	7.2135E-01
8	8.0000E-01	7.9555E-01	3.5840E-03	4.4465E-03	5.5582E-01
9	9.0000E-01	9.0495E-01	3.7034E-03	-4.9506E-03	-5.5007E-01
10	9.7500E-01	9.7432E-01	6.1749E-03	6.8371E-04	7.0124E-02

Rezidualní součet čtverců, RSC	: 2.5076E-04
Průměr absolutních hodnot reziduí, Me	: 4.1755E-03
Průměr relativních reziduí, Mer	: 1.2973E+00
Odhad reziduálního rozptylu, $s^2(e)$: 4.1794E-05
Odhad směrodatné odchylky reziduí, s(e)	: 6.4648E-03
Odhad šikmosti reziduí, g1(e)	: 4.4674E-01
Odhad špičatosti reziduí, g2(e)	: 2.2391E+00

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

Fisher-Snedocorův test významnosti regrese, F	: 6.4030E+03
Tabulkový kvantil, $F(1-\alpha, m-1, n-m)$: 4.7571E+00

Závěr: Navržený model je přijat jako významný.

Spočtená hladina významnosti	: 0.000
------------------------------	---------

Scottovo kritérium multikolinearity, M	: 8.9170E-01
Závěr: Navržený model není korektní.	

Cook-Weisbergův test heteroskedasticity, Sf	: 4.5932E+00
Tabulkový kvantil, $\chi^2(1-\alpha, 1)$: 3.8415E+00

Závěr: Rezidua vykazují heteroskedasticitu.

Spočtená hladina významnosti	: 0.032
------------------------------	---------

Jarque-Berraův test normality reziduí, L(e)	: 5.7389E-01
Tabulkový kvantil, $\chi^2(1-\alpha, 2)$: 5.9915E+00

Závěr: Normalita je přijata.

Spočtená hladina významnosti	: 0.751
------------------------------	---------

Waldův test autokorelace, Wa	: 1.0130E+00
Tabulkový kvantil, $\chi^2(1-\alpha, 1)$: 3.8415E+00

Závěr: Rezidua nejsou autokorelována.

Spočtená hladina významnosti	: 0.314
------------------------------	---------

Znamkový test, Dt	: 3.3541E-01
Tabulkový kvantil, N(1-alpha/2)	: 1.6449E+00
Závěr: Rezidua nevykazují trend.	
Spočtená hladina významnosti	: 0.369

(7) INDIKACE VLIVNÝCH BODŮ:

(* indikuje odlehlý nebo vlivný bod)

Bod	Standardizované reziduum	Jackknife reziduum	Predikované reziduum	Diagonální prvky
i	eS[i]	eJ[i]	eP[i]	H[i, i]
1	-1.4640E+00	-1.6669E+00	-3.8853E-02	9.4066E-01*
2	1.8970E+00	2.7373E+00*	1.5278E-02	3.5567E-01
3	-1.2684E+00	-1.3534E+00	-9.8396E-03	3.0555E-01
4	-2.5006E-01	-2.2948E-01	-1.8469E-03	2.3379E-01
5	-9.2031E-01	-9.0654E-01	-6.5877E-03	1.8435E-01
6	1.4676E-01	1.3421E-01	1.0545E-03	1.9052E-01
7	8.9689E-01	8.7984E-01	6.6580E-03	2.4160E-01
8	8.2643E-01	8.0141E-01	6.4195E-03	3.0734E-01
9	-9.3428E-01	-9.2262E-01	-7.3689E-03	3.2817E-01
10	3.5719E-01	3.2960E-01	7.7992E-03	9.1234E-01*
Bod	Zobecněné diag. prvky	Cookova vzdálenost	Atkinsonova vzdálenost	Vliv na predikci
i	Hm[i, i]	D[i]	A[i]	DF[i]
1	9.6186E-01	8.4941E+00*	8.1284E+00*	-6.6368E+00*
2	7.4212E-01	4.9661E-01*	2.4908E+00*	2.0337E+00*
3	4.9175E-01	1.7696E-01	1.0995E+00	-8.9776E-01
4	2.4178E-01	4.7702E-03	1.5525E-01	-1.2676E-01
5	2.9949E-01	4.7858E-02	5.2784E-01	-4.3098E-01
6	1.9343E-01	1.2673E-03	7.9746E-02	6.5112E-02
7	3.4328E-01	6.4063E-02	6.0820E-01	4.9660E-01
8	3.8619E-01	7.5762E-02	6.5381E-01	5.3383E-01
9	4.2591E-01	1.0659E-01	7.8975E-01	-6.4483E-01
10	9.1420E-01	3.3196E-01*	1.3022E+00	1.0633E+00
Bod	V ě r o h o d n o s t n í v z d á l e n o s t i			
i	LD(b)[i]	LD(s ²)[i]	LD(b, s ²)[i]	
1	1.8965E+01*	6.3565E-01	7.9921E+01*	
2	2.8599E+00	4.3837E+00*	1.1829E+01*	
3	1.1152E+00	2.2933E-01	1.6801E+00	
4	3.1751E-02	4.3624E-02	7.2547E-02	
5	3.1407E-01	1.1130E-02	3.4547E-01	
6	8.4452E-03	5.0067E-02	5.7699E-02	
7	4.1822E-01	7.5421E-03	4.5143E-01	
8	4.9274E-01	1.2054E-03	5.1417E-01	
9	6.8652E-01	1.3679E-02	7.6213E-01	
10	1.9992E+00	3.4205E-02	2.0692E+00	

Výstupy č. 2

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

Fisher-Snedecorův test významnosti regrese, F : 1.1106E+04
 Tabulkový kvantil, $F(1-\alpha, m-1, n-m)$: 5.4095E+00

Závěr: Navržený model je přijat jako významný.
 Spočtená hladina významnosti : 0.000

Scottovo kritérium multikolinearity, M : 8.8255E-01
 Závěr: Navržený model není korektní.

Cook-Weisbergův test heteroskedasticity, S_f : 8.0546E-01
 Tabulkový kvantil, $\chi^2(1-\alpha, 1)$: 3.8415E+00

Závěr: Rezidua vykazují homoskedasticitu.
 Spočtená hladina významnosti : 0.369

Jarque-Berraův test normality reziduí, $L(e)$: 1.0627E+00
 Tabulkový kvantil, $\chi^2(1-\alpha, 2)$: 5.9915E+00

Závěr: Normalita je přijata.
 Spočtená hladina významnosti : 0.588

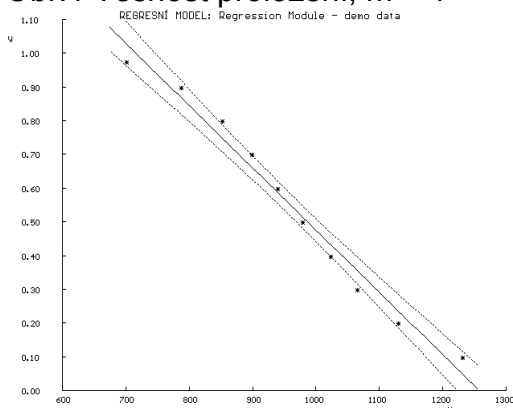
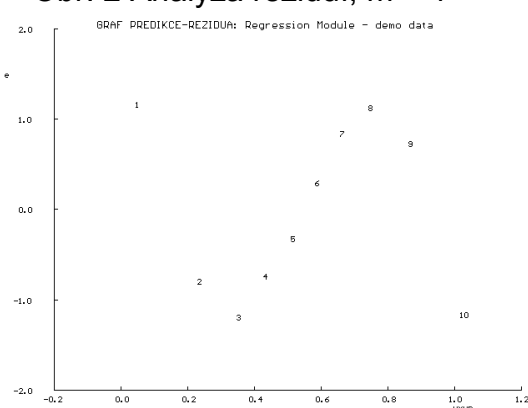
Waldův test autokorelace, W_a : 1.2281E+00
 Tabulkový kvantil, $\chi^2(1-\alpha, 1)$: 3.8415E+00

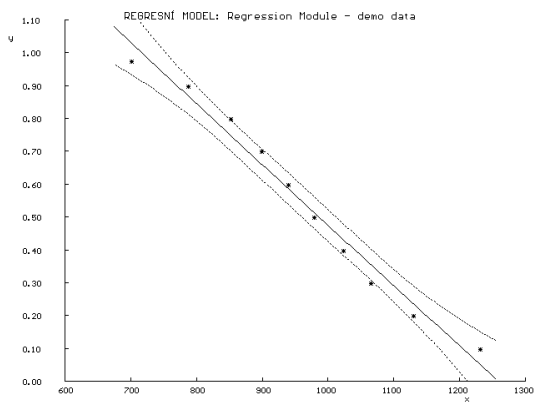
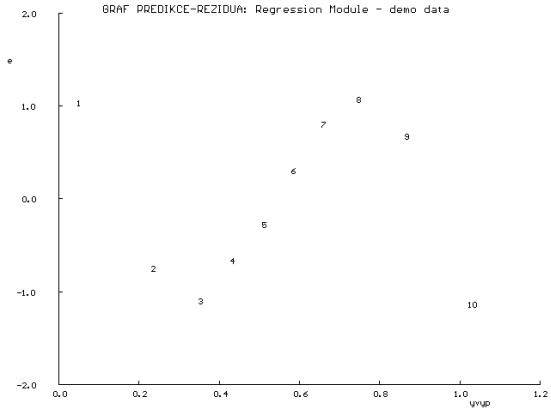
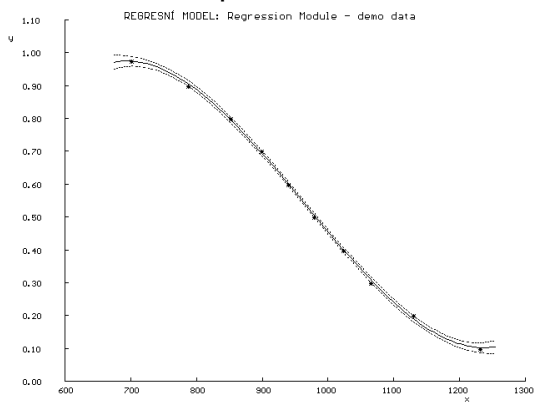
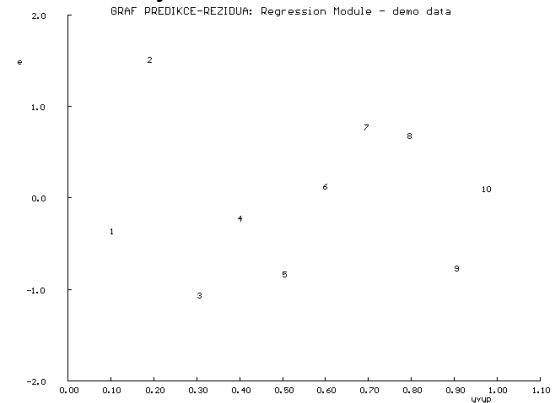
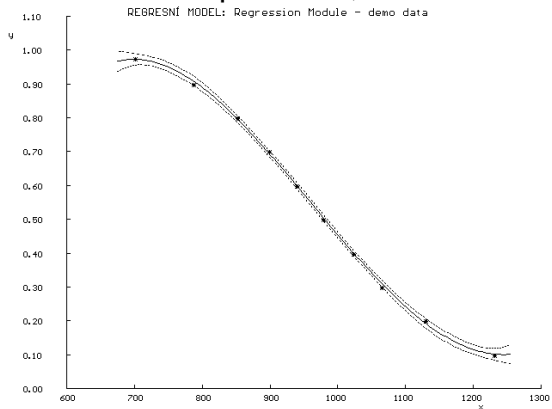
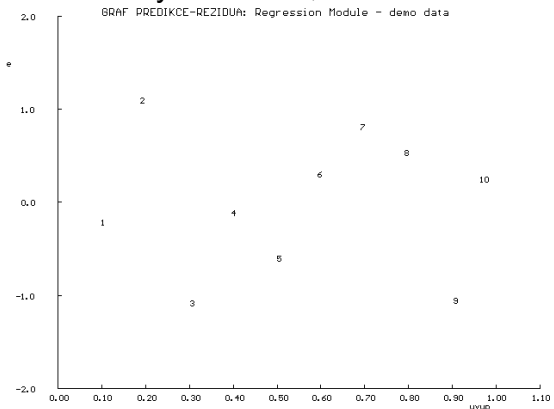
Závěr: Rezidua nejsou autokorelována.
 Spočtená hladina významnosti : 0.268

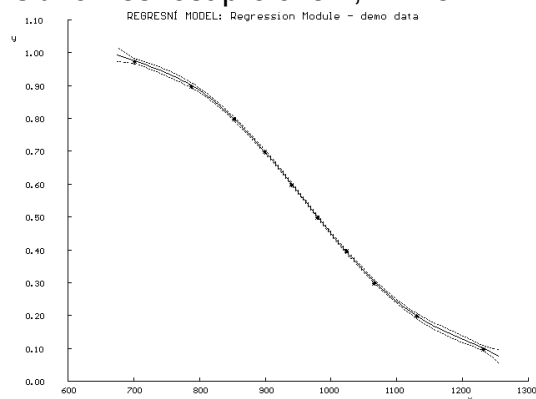
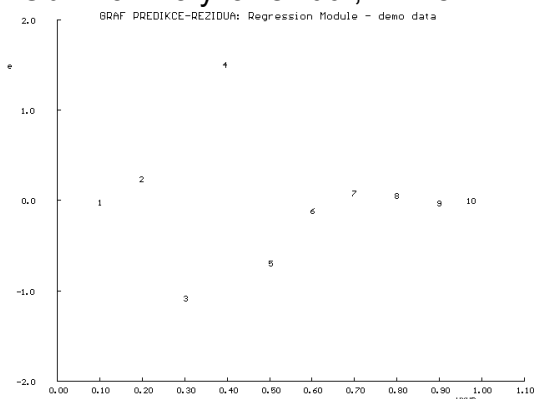
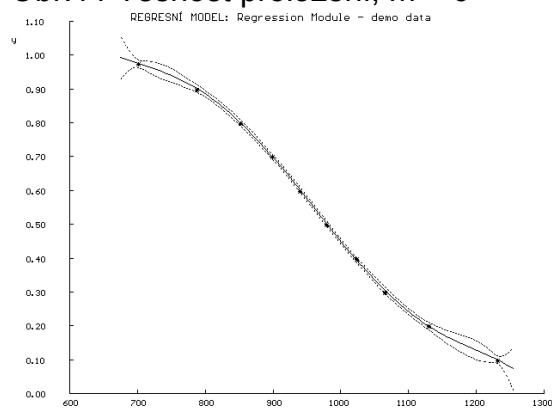
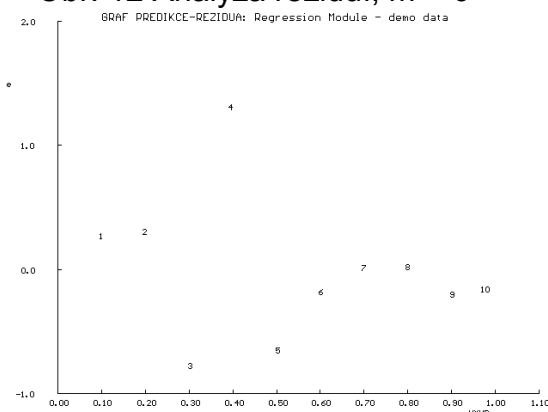
Znamkový test, D_t : 1.2247E+00
 Tabulkový kvantil, $N(1-\alpha/2)$: 1.6449E+00

Závěr: Rezidua nevykazují trend.
 Spočtená hladina významnosti : 0.110

Obrázky:

Obr.1 Těsnost proložení, $m = 1$ Obr. 2 Analýza reziduí, $m = 1$ 

Obr.3 Těsnost proložení, $m = 2$ Obr. 4 Analýza reziduí, $m = 2$ Obr.5 Těsnost proložení, $m = 3$ Obr. 6 Analýza reziduí, $m = 3$ Obr.7 Těsnost proložení, $m = 4$ Obr. 8 Analýza reziduí, $m = 4$ 

Obr.9 Těsnost proložení, $m = 5$ Obr. 10 Analýza reziduí, $m = 5$ Obr.11 Těsnost proložení, $m = 6$ Obr. 12 Analýza reziduí, $m = 6$ **Řešení:**

Nejprve se provede návrh modelu křivkové závislosti – určení stupně polynomu, pomocí MNČ.

Pro nalezení optimálního stupně polynomu byly použity tyto statistické charakteristiky: R – vícenásobný korelační koeficient, R^2 – koeficient determinace, R_p^2 – predikovaný korelační koeficient, MEP – střední kvadratická chyba predikce, AIC – Aikeho informační kritérium. Optimální stupeň polynomu je dán minimem statistických charakteristik MEP a AIC a maximem R_p^2 . V tabulce Tabulka 2 jsou hodnoty těchto charakteristik z programu ADSTAT pro různé stupně polynomu – m .

Tabulka 2 hodnoty statistických charakteristik R , R^2 , R_p^2 , MEP a AIC.

m	R	R^2	R_p^2	MEP	AIC
1	0,98977	0,97965	0,98052	0,0030983	-60,167
2	0,98979	0,97968	0,93432	0,01202	-58,178
3	0,99984	0,99969	0,9987	0,00020884	-97,936
4	0,99985	0,99971	0,98184	0,0028898	-96,63
5	0,99998	0,99995	0,99877	0,00019725	-113,1
6	0,99997	0,99995	0	2901,1	-109,95
7	0,71608	0,51278	0	41982000	-16,409

Jak je vidět z Tabulky 2, tak R_p^2 má maxima pro stupeň polynomu $m = 3$ a $m = 5$, MEP má minima pro $m = 3$ a $m = 5$ a AIC má minima pro $m = 3$ a $m = 5$.

Který stupeň polynomu nejlépe odpovídá závislosti se určí z těsnosti proložení pomocí grafů uvedených na obrázcích obr. 1 až obr.12. Z obrázků těsnosti proložení a anýzy reziduí lze určit optimální stupeň polynomu $m = 3$, protože konce Working Hottelingových pásů pro $m \geq 4$ se zřetelně rozevírají.

Kritika dat

Analýzou dat pro polynom třetího stupně – viz Výstupy č. 1 z programu ADSTAT, lze konstatovat, že:

1. Data vykazují silnou multikolinearitu, protože čísla podmíněnosti K ($K_{\max} = 2,42 \cdot 10^5$) i VIF jsou vyšší než limitní hodnoty pro $K = 1000$ a pro $VIF = 10$.
2. Navržený model je nekorektní v důsledku multikolinearity – Scottovo kritérium $M = 0,8917$
3. Analýza reziduí indikuje, že rezidua vykazují heteroskedasticitu – Cook-Weisbergův test.
4. Analýza reziduí indikuje podezřelý bod 1 a odlehlý bod 2, Jackknife reziduum, Cookova vzdálenost a kritérium DF_i .
5. Analýza reziduí indikuje vlivné body 1 a 10, diagonální prvky H_{ii} .

Vyloučením odlehlého bodu 2 se odstranila heteroskedasticita – viz Výstup č.2.

Vzhledem k tomu, že všechny nulové hypotézy jsou zamítnuté – všechny odhady parametrů β jsou nenulové, není třeba optimalizovat omezení filtru $P = 1 \cdot 10^{-34}$ a tím se nemusí použít metoda RH a polynomický model se určí pomocí MNČ.

Závěr:

Závislost stupně přeměny, vztažený na oxid SO_2 , na teplotě vyjádřená polynomicky nejlépe popisuje polynom třetího stupně, ale vzhledem k silné multikolinearitě, hodnota M Scottova kritéria je větší než 0,8, by bylo vhodné zvolit jiný model závislosti.

Úloha 3. Validizace nové analytické metody

Zadání:

Byla vypracována extrakčně – radiometrická metoda na stanovení malých množství neionogenních tenzidů založená na tvorbě trojných komplexů neionogenního tenzidu a bengálské červeně značené ^{131}I v přítomnosti Ba^{2+} iontů v alkalickém prostředí, $pH = 10$, a následnou extrakci do chloroformu. Metoda byla ověřena stanovením neionogenního tenzidu PEG – 1000, jehož koncentrace se měnila v rozsahu $10,0 \text{ mg/dm}^3$, $30,0 \text{ mg/dm}^3$ a $50,0 \text{ mg/dm}^3$ a každé stanovení se provedlo pětkrát.

Data:

Získané výsledky jsou uvedené v Tabulce 1.

Tabulka 1 Výsledky stanovení PEG – 1000 extrakčně radiometrickou metodou pomocí bengálské červeni značené ^{131}I v přítomnosti Ba^{2+} iontů

Daný obsah PEG – 1000, v mg/dm^3	10,0	30,0	50,0
Stanovený obsah PEG – 1000, v mg/dm^3	9,85	29,50	48,89
	11,02	29,87	50,43
	10,53	31,15	51,25
	10,02	30,04	49,52
	9,97	29,25	49,75

Užitý program:

Pro zpracování získaných výsledků byl použit program ADSTAT modul Lineární regrese a kalibrace.

Výstup:

Podmínky a kvantily pro statistické testy a výstupy statistického zpracování dat pomocí nástrojů lineární regrese z programu ADSTAT jsou uvedeny níže.

PODMÍNKY A KVANTILY PRO STATISTICKÉ TESTY:

Hladina významnosti, alfa : 0.050
 Počet bodů, n : 15
 Počet parametrů, m : 2
 Kvantil Studentova rozdělení $t(1-\alpha/2, n-m)$: 2.160
 Kvantil rozd. Chí-kuadrát $\text{Chi-square}(1-\alpha, m)$: 5.991
 Jméno výstupního souboru : RESULTS.TXT

V Ý S T U P

(1) PŘEDBĚŽNÁ STATISTICKÁ ANALÝZA:

Proměnná	Průměr	Směrodatná odchylka	Párový korelační koeficient	Spočtená hladina úz.
y	3.0069E+01	1.6786E+01	1.0000	-----
x1	3.0000E+01	1.6903E+01	0.9992	0.000

Párové korelační koeficienty mezi dvojicemi vysvětlujících proměnných	Spočtená hladina významnosti
---	------------------------------

(2) INDIKACE MULTIKOLINEARITY:

Č [j]	Vlastní čísla korel. matice I[j]	Čísla podmíněnosti K[j]	Variance inflation factor VIF[j]	Vícenás.korel. koef pro X[j]
1	1.0000E+00	1.0000E+00	1.0000E+00	0.0000

Maximální číslo podmíněnosti K : 1.0000E+00
 (K[j], $K > 1000$ indikuje silnou multikolaritu)
 (VIF[j] > 10 indikuje silnou multikolaritu)

(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.
$B_{[0]}$	3.0183E-01	3.8066E-01	7.9291E-01	Akceptována	0.442
$B_{[1]}$	9.9225E-01	1.1145E-02	8.9033E+01	Zamítnuta	0.000

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R	: 9.9918E-01
Koeficient determinace, R^2	: 9.9836E-01
Predikovaný korelační koeficient, R_p^2	: 9.9890E-01
Střední kvadratická chyba predikce, MEP	: 5.7712E-01
Akaikeho informační kritérium, AIC	: -8.6395E+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	9.8500E+00	1.0224E+01	2.8776E-01	-3.7433E-01	-3.8003E+00
2	1.1020E+01	1.0224E+01	2.8776E-01	7.9567E-01	7.2202E+00
3	1.0530E+01	1.0224E+01	2.8776E-01	3.0567E-01	2.9028E+00
4	1.0020E+01	1.0224E+01	2.8776E-01	-2.0433E-01	-2.0393E+00
5	9.9700E+00	1.0224E+01	2.8776E-01	-2.5433E-01	-2.5510E+00
6	2.9500E+01	3.0069E+01	1.8199E-01	-5.6933E-01	-1.9299E+00
7	2.9870E+01	3.0069E+01	1.8199E-01	-1.9933E-01	-6.6733E-01
8	3.1150E+01	3.0069E+01	1.8199E-01	1.0807E+00	3.4692E+00
9	3.0040E+01	3.0069E+01	1.8199E-01	-2.9333E-02	-9.7645E-02
10	2.9250E+01	3.0069E+01	1.8199E-01	-8.1933E-01	-2.8011E+00
11	4.8890E+01	4.9914E+01	2.8776E-01	-1.0243E+00	-2.0952E+00
12	5.0430E+01	4.9914E+01	2.8776E-01	5.1567E-01	1.0225E+00
13	5.1250E+01	4.9914E+01	2.8776E-01	1.3357E+00	2.6062E+00
14	4.9520E+01	4.9914E+01	2.8776E-01	-3.9433E-01	-7.9631E-01
15	4.9750E+01	4.9914E+01	2.8776E-01	-1.6433E-01	-3.3032E-01

Rezidualní součet čtverců, RSC	: 6.4586E+00
Průměr absolutních hodnot reziduí, M_e	: 5.3778E-01
Průměr relativních reziduí, M_{er}	: 2.2886E+00
Odhad reziduálního rozptylu, $s^2(e)$: 4.9682E-01
Odhad směrodatné odchylky reziduí, $s(e)$: 7.0485E-01
Odhad šikmosti reziduí, $g_1(e)$: 5.5539E-01
Odhad špičatosti reziduí, $g_2(e)$: 2.4220E+00

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

Fisher-Snedocorův test významnosti regrese, F	: 7.9269E+03
Tabulkový kvantil, $F(1-\alpha, m-1, n-m)$: 4.6672E+00
Závěr: Navržený model je přijat jako významný.	
Spočtená hladina významnosti	: 0.000

Scottovo kritérium multikolinearity, M	: 1.8097E-14
Závěr: Navržený model je korektní.	

Cook-Weisbergův test heteroskedasticity, Sf : 4.2250E+01
 Tabulkový kvantil, $\text{Chi}^2(1-\alpha,1)$: 3.8415E+00
 Závěr: Rezidua vykazují heteroskedasticitu.
 Spočtená hladina významnosti : 0.000

Jarque-Berraův test normality reziduí, L(e) : 9.7995E-01
 Tabulkový kvantil, $\text{Chi}^2(1-\alpha,2)$: 5.9915E+00
 Závěr: Normalita je přijata.
 Spočtená hladina významnosti : 0.613

Waldův test autokorelace, Wa : 9.4167E-02
 Tabulkový kvantil, $\text{Chi}^2(1-\alpha,1)$: 3.8415E+00
 Závěr: Rezidua nejsou autokorelována.
 Spočtená hladina významnosti : 0.759

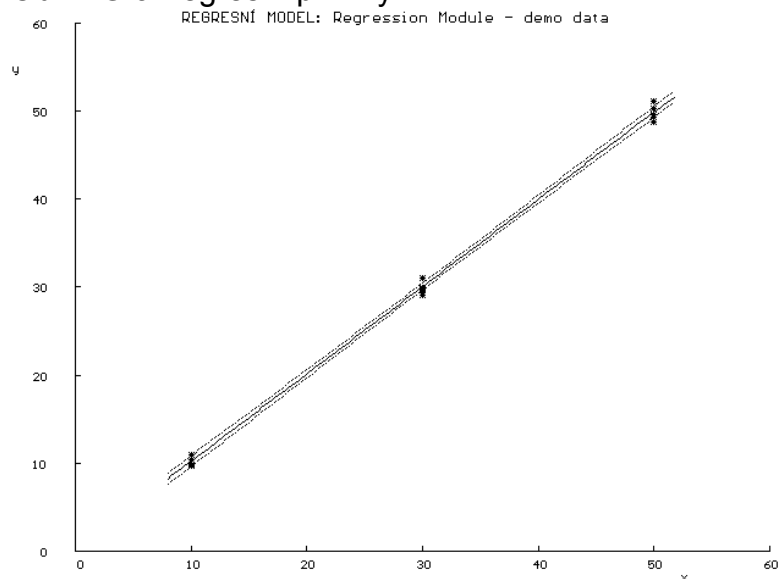
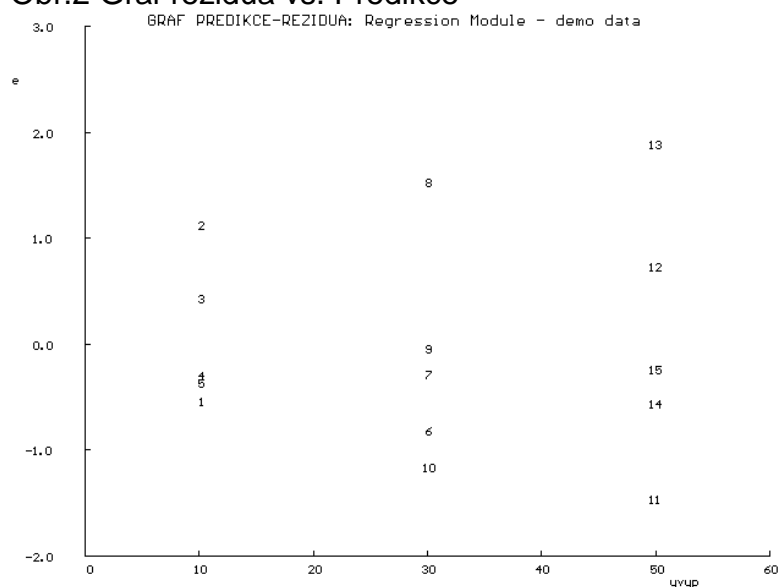
Znamkový test, Dt : -1.0146E-01
 Tabulkový kvantil, $N(1-\alpha/2)$: 1.6449E+00
 Závěr: Rezidua nevykazují trend.
 Spočtená hladina významnosti : 0.460

(7) INDIKACE VLIVNÝCH BODŮ:

(* indikuje odlehlý nebo vlivný bod)

Bod	Standardizované reziduum	Jackknife reziduum	Predikované reziduum	Diagonální prvky
i	eS[i]	eJ[i]	eP[i]	H[i, i]
1	-5.8177E-01	-5.6637E-01	-4.4920E-01	1.6667E-01
2	1.2366E+00	1.2648E+00	9.5480E-01	1.6667E-01
3	4.7505E-01	4.6043E-01	3.6680E-01	1.6667E-01
4	-3.1756E-01	-3.0630E-01	-2.4520E-01	1.6667E-01
5	-3.9527E-01	-3.8207E-01	-3.0520E-01	1.6667E-01
6	-8.3608E-01	-8.2579E-01	-6.1000E-01	6.6667E-02
7	-2.9273E-01	-2.8217E-01	-2.1357E-01	6.6667E-02
8	1.5870E+00	1.6981E+00	1.1579E+00	6.6667E-02
9	-4.3076E-02	-4.1389E-02	-3.1428E-02	6.6667E-02
10	-1.2032E+00	-1.2263E+00	-8.7786E-01	6.6667E-02
11	-1.5920E+00	-1.7047E+00	-1.2292E+00	1.6667E-01
12	8.0142E-01	7.8974E-01	6.1880E-01	1.6667E-01
13	2.0758E+00	2.4392E+00*	1.6028E+00	1.6667E-01
14	-6.1285E-01	-5.9750E-01	-4.7320E-01	1.6667E-01
15	-2.5540E-01	-2.4600E-01	-1.9720E-01	1.6667E-01

Bod	V ě r o h o d n o s t n í u z d á l e n o s t i		
i	LD(b)[i]	LD(s ²)[i]	LD(b, s ²)[i]
1	7.7902E-02	1.3428E-02	8.8274E-02
2	3.4879E-01	2.4092E-02	3.9735E-01
3	5.1988E-02	1.9540E-02	6.9005E-02
4	2.3254E-02	2.7531E-02	4.9422E-02
5	3.6012E-02	2.3829E-02	5.7890E-02
6	5.7502E-02	1.4056E-03	5.8233E-02
7	7.0606E-03	2.8587E-02	3.5222E-02
8	2.0615E-01	1.6876E-01	4.0905E-01
9	7.0668E-06	3.4750E-02	3.4893E-02
10	1.1885E-01	1.8361E-02	1.4368E-01
11	5.7374E-01	1.7235E-01	8.5040E-01
12	1.4749E-01	2.5032E-03	1.4803E-01
13	9.6282E-01	9.3620E-01	2.3245E+00
14	8.6424E-02	1.1652E-02	9.4954E-02
15	1.5045E-02	3.0040E-02	4.4160E-02

Obrázky:**Obr.1 Graf regresní přímky****Obr.2 Graf rezidua vs. Predikce****Řešení:**

Nejdříve se provede vyšetření regresního tripletu.

V případě, že metoda poskytuje spolehlivé výsledky, bude závislost stanoveného množství PEG – 1000 v mg/dm^3 (y) na přidaném množství PEG – 1000 v mg/dm^3 (x) lineární $y = \beta_0 + \beta_1 x$, s nulovým úsekem regresní přímky, β_0 , a jednotkovou směrnici regresní přímky, β_1 . Bodové odhady parametrů regresní přímky vypočítané metodou nejmenších čtverců z programu ADSTAT jsou $b_0 = 0,30183 (\pm 0,38066)$, $b_1 = 0,99225 (\pm 0,011145)$ a Studentův kvantil $t_{(1-\alpha/2)}(13) = 2,160$.

Podle nerovnice (1) se určí intervalové odhady parametrů regresní přímky b_0 a b_1 :

$$b_j - t_{(1-\alpha/2)}(n-m)\sqrt{D(b_j)} \leq \beta_j \leq b_j + t_{(1-\alpha/2)}(n-m)\sqrt{D(b_j)} \quad (1),$$

kde $D(b_j)$ je rozptyl bodových odhadů parametrů regresní přímky.

Po dosazení jednotlivých hodnot pro b_0 a b_1 , vychází intervalové odhady parametrů regresní přímky: $-0,5204 \leq \beta_0 \leq 1,1241$, nula leží v tomto intervalu spolehlivosti, takže parametr úseku regresní přímky $\beta_0 = 0$, a $0,9682 \leq \beta_1 \leq 1,0163$, jednička leží v tomto intervalu spolehlivosti, takže $\beta_1 = 1$, směrnice regresní přímky je jednotková.

Závěr:

Navržená metoda extrakčně – radiometrického stanovení malých množství neionogenních tenzidů poskytuje spolehlivé výsledky.

Úloha 4. Vícerozměrný lineární regresní model**Zadání:**

Z databáze pro radiační ochranu pro mimořádné události byl studován vliv počtu obyvatel lokalit, vzdálenost lokalit od jaderného zařízení, naměřené dávky před opatřeními, dávky po opatřeních na hodnoty odvrácených dávek, kterými mohou být zasaženi obyvatelé lokalit.

Data:

počet obyvatel, x_1	vzdálenost, km, x_2	dávka bez opatření, mSv, x_3	dávka po opatřeních, mSv, x_4	odvrácená dávka, mSv, y
650	3	91,1	16,6	74,5
306	7	46,8	14,3	32,5
309	5	33,9	10,2	23,7
742	11	28,7	8,7	20
393	4	27,7	8,5	19,2
274	12	18	5,4	12,6
400	32	15,8	4,7	11,1
171	16	14,7	4,4	10,3
338	17	11,2	3,3	7,9
391	12	10,5	3	7,5
482	6	8,7	2,4	6,3
3036	20	6,8	2,1	4,7
151	25	6,8	2	4,8

Užitý program:

ADSTAT, modul lineární regrese.

Výstup:

Výstupu č.1.

V Ý S T U P

(1) PŘEDBĚŽNÁ STATISTICKÁ ANALÝZA:

Proměnná	Průměr	Směrodatná odchylka	Párový korelační koeficient	Spočtená hladina výz.
y	1.8085E+01	1.8878E+01	1.0000	-----
x1	5.8792E+02	7.5427E+02	-0.0986	0.749
x2	1.3077E+01	8.7318E+00	-0.5363	0.059
x3	2.4669E+01	2.3295E+01	0.9963	0.000
x4	6.5846E+00	4.7702E+00	0.9078	0.000

Párové korelační koeficienty mezi dvojicemi vysvětlujících proměnných

	Spočtená hladina významnosti
x1 versus x2 :	0.618
x1 versus x3 :	0.702
x1 versus x4 :	0.546
x2 versus x3 :	0.047
x2 versus x4 :	0.027
x3 versus x4 :	0.000

(2) INDIKACE MULTIKOLINEARITY:

Č [j]	Vlastní čísla korel. matice l[j]	Čísla podmíněnosti K[j]	Variance inflation factor VIF[j]	Vícenás.korel. koef pro X[j]
1	5.5256E-02	4.4668E+01	1.0673E+00	0.2512
2	5.1960E-01	4.7502E+00	1.5986E+00	0.6119
3	9.5694E-01	2.5793E+00	8.9266E+00	0.9423
4	2.4682E+00	1.0000E+00	9.8796E+00	0.9480

Maximální číslo podmíněnosti K : 4.4668E+01

(K[j], $K > 1000$ indikuje silnou multikolaritu)

(VIF[j] > 10 indikuje silnou multikolaritu)

(3) ODHADY PARAMETRŮ A TESTY VÝZNAMNOSTI:

Parametr	Odhad	Směrodatná odchylka	Test $H_0: B[j] = 0$ vs. $H_A: B[j] \neq 0$	t-kriterium	hypoteza H_0 je	Hlad. výz.
BI 01	2.5252E-07	6.7109E-07	3.7628E-01	3.7628E-01	Akceptována	0.716
BI 11	-2.6396E-10	2.5120E-10	-1.0508E+00	-1.0508E+00	Akceptována	0.324
BI 21	-1.5218E-09	2.6556E-08	-5.7304E-02	-5.7304E-02	Akceptována	0.956
BI 31	1.0000E+00	2.3522E-08	4.2513E+07	4.2513E+07	Zamítnuta	0.000
BI 41	-1.0000E+00	1.2085E-07	-8.2750E+06	-8.2750E+06	Zamítnuta	0.000

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R	: 1.0000E+00
Koeficient determinace, R^2	: 1.0000E+00
Predikovaný korelační koeficient, R_p^2	: 1.0000E+00
Střední kvadratická chyba predikce, MEP	: 1.0633E-09
Akaikeho informační kritérium, AIC	:-3.6731E+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	7.4500E+01	7.4500E+01	6.3520E-07	4.2081E-08	5.6484E-08
2	3.2500E+01	3.2500E+01	4.7132E-07	9.6974E-07	2.9838E-06
3	2.3700E+01	2.3700E+01	2.9891E-07	-1.0188E-06	-4.2986E-06
4	2.0000E+01	2.0000E+01	2.4697E-07	-8.9814E-07	-4.4907E-06
5	1.9200E+01	1.9200E+01	2.9282E-07	-3.0980E-08	-1.6136E-07
6	1.2600E+01	1.2600E+01	2.1023E-07	3.6405E-07	2.8893E-06
7	1.1100E+01	1.1100E+01	4.9197E-07	-6.2692E-08	-5.6479E-07
8	1.0300E+01	1.0300E+01	2.3032E-07	3.3204E-07	3.2237E-06
9	7.9000E+00	7.9000E+00	2.3908E-07	1.2018E-07	1.5212E-06
10	7.5000E+00	7.5000E+00	2.7923E-07	-1.3086E-07	-1.7448E-06
11	6.3000E+00	6.3000E+00	4.0661E-07	3.4456E-07	5.4691E-06
12	4.7000E+00	4.7000E+00	6.2533E-07	1.3243E-07	2.8177E-06
13	4.8000E+00	4.8000E+00	3.5442E-07	-1.6364E-07	-3.4093E-06
Rezidualní součet čtverců, RSC					: 3.2290E-12
Průměr absolutních hodnot reziduí, Me					: 3.5463E-07
Průměr relativních reziduí, Mer					: 2.5870E-06
Odhad reziduálního rozptylu, s ² (e)					: 4.0362E-13
Odhad směrodatné odchylky reziduí, s(e)					: 6.3531E-07
Odhad šikmosti reziduí, g1(e)					: -4.6413E-01
Odhad špičatosti reziduí, g2(e)					: 3.3136E+00
(6) TESTOVÁNÍ REGRESNÍHO TRIPLETU (DATA + MODEL + METODA):					
Fisher-Snedocorův test významnosti regrese, F					: 2.2518E+15
Tabulkový kvantil, F(1-alpha, m-1, n-m)					: 3.8379E+00
Závěr: Navržený model je přijat jako významný.					
Spočtená hladina významnosti					: 0.000
Scottovo kritérium multikolinearity, M					: 6.5528E-01
Závěr: Navržený model není korektní.					
Cook-Weisbergův test heteroskedasticity, Sf					: 5.3144E-01
Tabulkový kvantil, Chi ² (1-alpha, 1)					: 3.8415E+00
Závěr: Rezidua vykazují homoskedasticitu.					
Spočtená hladina významnosti					: 0.466
Jarque-Berraův test normality reziduí, L(e)					: 5.2001E-01
Tabulkový kvantil, Chi ² (1-alpha, 2)					: 5.9915E+00
Závěr: Normalita je přijata.					
Spočtená hladina významnosti					: 0.771
Waldův test autokorelace, Wa					: 4.0029E-03
Tabulkový kvantil, Chi ² (1-alpha, 1)					: 3.8415E+00
Závěr: Rezidua nejsou autokorelována.					
Spočtená hladina významnosti					: 0.950
Znamékový test, Dt					: 2.2428E-02
Tabulkový kvantil, N(1-alpha/2)					: 1.6449E+00
Závěr: Rezidua nevykazují trend.					
Spočtená hladina významnosti					: 0.491

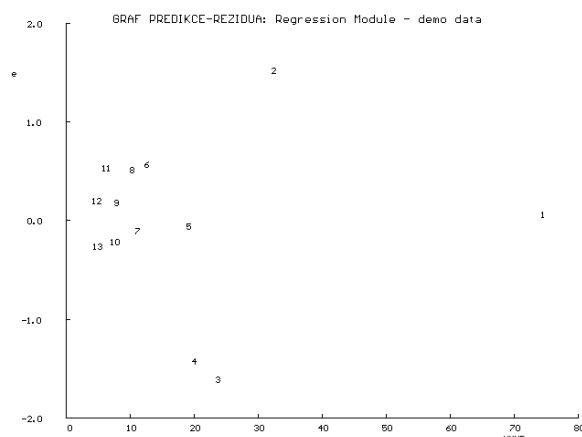
(7) INDIKACE VLIVNÝCH BODŮ:

(* indikuje odlehlý nebo vlivný bod)

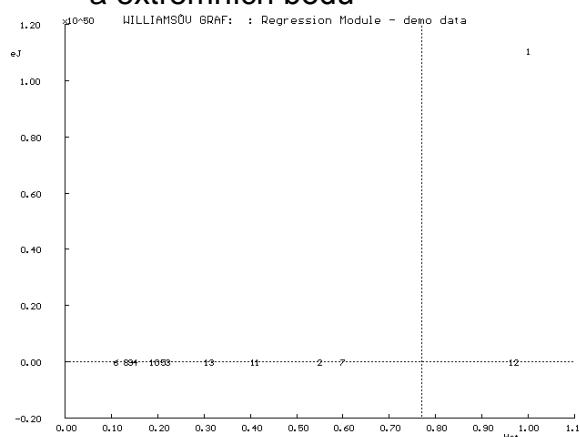
Bod	Standardizované reziduum $eS[i]$	Jackknife reziduum $eJ[i]$	Predikované reziduum $eP[i]$	Diagonální prvky $H[i, i]$
1	3.4994E+00	1.1066E+50*	1.1746E-04	9.9964E-01*
2	2.2764E+00	3.5877E+00*	2.1568E-06	5.5038E-01
3	-1.8173E+00	-2.2184E+00	-1.3084E-06	2.2136E-01
4	-1.5344E+00	-1.7085E+00	-1.0580E-06	1.5111E-01
5	-5.4949E-02	-5.1410E-02	-3.9337E-08	2.1244E-01
6	6.0724E-01	5.8158E-01	4.0882E-07	1.0950E-01
7	-1.5596E-01	-1.4611E-01	-1.5660E-07	5.9966E-01
8	5.6080E-01	5.3520E-01	3.8229E-07	1.3143E-01
9	2.0417E-01	1.9148E-01	1.4000E-07	1.4162E-01
10	-2.2932E-01	-2.1522E-01	-1.6219E-07	1.9317E-01
11	7.0584E-01	6.8183E-01	5.8362E-07	4.0963E-01
12	1.1807E+00	1.2153E+00	4.2484E-06	9.6883E-01*
13	-3.1037E-01	-2.9208E-01	-2.3759E-07	3.1122E-01
Bod	Zobecněné diag. prvky $Hm[i, i]$	Cookova vzdálenost $D[i]$	Atkinsonova vzdálenost $A[i]$	Vliv na predikci $DF[i]$
1	1.0002E+00*	6.8339E+03*	7.3940E+51*	5.8455E+51*
2	8.4162E-01	1.2686E+00*	5.0210E+00*	3.9694E+00*
3	5.4279E-01	1.8778E-01	1.4962E+00	-1.1828E+00
4	4.0093E-01	8.3820E-02	9.1182E-01	-7.2086E-01
5	2.1273E-01	1.6289E-04	3.3773E-02	-2.6700E-02
6	1.5055E-01	9.0686E-03	2.5797E-01	2.0394E-01
7	6.0088E-01	7.2867E-03	2.2619E-01	-1.7882E-01
8	1.6558E-01	9.5180E-03	2.6335E-01	2.0819E-01
9	1.4609E-01	1.3755E-03	9.8380E-02	7.7776E-02
10	1.9848E-01	2.5181E-03	1.3320E-01	-1.0531E-01
11	4.4639E-01	6.9137E-02	7.1840E-01	5.6795E-01
12	9.7426E-01*	8.6647E+00*	8.5704E+00*	6.7755E+00*
13	3.1952E-01	8.7050E-03	2.4835E-01	-1.9634E-01
Bod	V ě r o h o d n o s t n í v z d á l e n o s t i			
i	$LD(b)[i]$	$LD(s^2)[i]$	$LD(b, s^2)[i]$	
1	1.0868E+02*	1.0000E+34*	1.0000E+34*	
2	7.5899E+00	8.5423E+00*	3.5553E+01*	
3	1.4426E+00	1.5555E+00	3.9539E+00	
4	6.6380E-01	5.1354E-01	1.4043E+00	
5	1.3234E-03	4.0179E-02	4.1401E-02	
6	7.3474E-02	6.9397E-03	7.8241E-02	
7	5.9070E-02	3.7566E-02	9.2383E-02	
8	7.7104E-02	1.0231E-02	8.4537E-02	
9	1.1171E-02	3.5495E-02	4.5865E-02	
10	2.0444E-02	3.4222E-02	5.3233E-02	
11	5.4994E-01	1.6023E-03	5.5457E-01	
12	2.4163E+01*	8.3765E-02	7.8781E+01*	
13	7.0536E-02	2.9325E-02	9.5409E-02	

Obrázky:

Obr. 1 Analýza klasických reziduí



Obr. 2 Viliamsův graf odlehlých a extrémních bodů

**Řešení:**

Nejjednodušší model lineární regresní závislosti má tvar podle vztahu:

$$y = \beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \beta_3 \cdot x_3 + \beta_4 \cdot x_4 \quad (\text{rov. 1}),$$

kde proměnné jsou v prvních mocninách.

Výsledky testování tohoto navrženého modelu jsou ve výstupu č.1. Lze konstatovat že:

1. Pearsonův párový korelační koeficient závislosti y na nezávisle proměnných x_i na silnou závislost y na x_3 a x_4 a i na silnou korelaci mezi x_3 a x_4 .
2. Studentův test ukazuje, že odhady parametrů b_0 , b_1 a b_2 jsou nulové, nulové hypotézy $b_i = 0$ jsou akceptované. To odpovídá realitě, že počet obyvatel lokality a vzdálenost lokality nemají v daném případě vliv na odvrácenou dávku.
3. Vícenásobný korelační koeficient ukazuje, že navržený model je statisticky významný.
4. Koeficient determinace a predikovaný koeficient determinace ukazují, že všechny body odpovídají navrženému modelu.
5. Hodnota střední kvadratické chyby predikce má velice malou hodnotu, co svědčí o optimálním modelu.
6. Analýzy reziduí ukazuje, že data odpovídají normálnímu rozdělení, vykazují homoskedasticitu, nejsou autokorelovaná a nevykazují trend.
7. Navržený model závislosti vykazuje multikolinearitu, Scottovo kritérium v rozmezí $\langle 0,3; 0,8 \rangle$. Model je z tohoto hlediska málo vyhovující.
8. Jackknife rezidua indikují odlehlý bod 1 a 2, diagonální prvky H_{ii} a H_{mm} indikují jako extrémní body 1 a 12. Ostatní rezidua rovněž vykazují jak odlehlé, tak vlivné body. Ale z výběru není důvod je vyloučit.

Na obrázcích obr. 1 je uvedena analýza klasických reziduí a na obr. 2. je Viliamsův graf na indikaci odlehlých a extrémních bodů.

Závěr:

Odvrácené dávky nezávisí na počtu obyvatel v lokalitách, ani na vzdálenosti lokalit od jaderných zařízení a závisí jenom na dávkách před opatřeními a po přijatých opatřeních podle vztahu:

$$y = 1,00(2 \cdot 10^{-8}) \cdot x_3 - 1,00(1 \cdot 10^{-7}) \cdot x_4 \quad (\text{rov.2}).$$

2.3 Kalibrace a limity její přesnosti

Úloha 1. Lineární kalibrace

Zadání:

Mezioperační stanovení obsahu L-fenylacetylkarbinolu (L-FAK) v surovém ketolu, při výrobě L-efedrinium chloridu, se provádí polarimetrickou metodou pomocí kalibrační přímky závislosti specifické otáčivosti α v $^\circ$ na koncentraci ketolu (navážka ketolu v g / 50 ml ethanolu). Je třeba určit parametry kalibrační přímky a bodový a intervalový odhad pro tři neznámé vzorky, s naměřenými hodnotami α : 4,00, 5,50 a 8,00, i limity přesnosti.

Data:

Hodnoty kalibrační přímky závislosti specifické otáčivosti α na navážce ketolu v 50 ml ethanolu jsou uvedeny v Tabulce 1.

Tabulka 1 Hodnoty kalibrační přímky závislosti specifické otáčivosti α na navážce ketolu v 50 ml ethanolu.

m v g	α v $^\circ$
0,8171	3,59
0,8351	3,69
1,2215	5,41
1,2231	5,49
1,5820	6,96
1,5821	7,02
2,0135	8,93
2,0233	8,98
2,4025	10,69
2,4027	10,68

Užitý program:

ADSTAT modul lineární regrese a modul kalibrace.

Výstup:
Výstup č. 1

V Ý S T U P

(1) PŘEDBĚŽNÁ STATISTICKÁ ANALÝZA:

Proměnná	Průměr	Směrodatná odchylka	Párový korelační koeficient	Spočtená hladina významnosti
y	7.1440E+00	2.6245E+00	1.0000	-----
x1	1.6103E+00	5.8892E-01	0.9999	0.000

Párové korelační koeficienty mezi dvojicemi vysvětlujících proměnných

Spočtená hladina významnosti

(2) INDIKACE MULTIKOLINEARITY:

Č [j]	Vlastní čísla korel. matice	Čísla podmíněnosti	Variance inflation factor	Vícenásobný korelační koef. pro X[j]
1	1.0000E+00	1.0000E+00	1.0000E+00	0.0000
Maximální číslo podmíněnosti K (K [j], K > 1000 indikuje silnou multikolinearitu) (VIF [j] > 10 indikuje silnou multikolinearitu)				: 1.0000E+00

(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. významnosti
B[0]	-3.1637E-02	3.2792E-02	-9.6480E-01	-9.6480E-01	Akceptována	0.363
B[1]	4.4561E+00	1.9239E-02	2.3162E+02	2.3162E+02	Zamítnuta	0.000

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R	: 9.9993E-01
Koeficient determinace, R^2	: 9.9985E-01
Predikovaný korelační koeficient, R_p^2	: 9.9990E-01
Střední kvadratická chyba predikce, MEP	: 1.2695E-03
Akaikeho informační kritérium, AIC	:-6.5865E+01

(5) ANALÝZA KLASICKÝCH REZIDUÍ:

od i	Meřená hodnota $y_{exp}[i]$	Predikovaná hodnota $y_{ypp}[i]$	Směrodatná odchylka $s(y_{ypp}[i])$	Klasické reziduum $e[i]$	Relativní reziduum $er[i]$
1	3.5900E+00	3.6095E+00	1.8665E-02	-1.9454E-02	-5.4190E-01
2	3.6900E+00	3.6897E+00	1.8383E-02	3.3570E-04	9.0976E-03
3	5.4100E+00	5.4115E+00	1.3095E-02	-1.5075E-03	-2.7865E-02
4	5.4900E+00	5.4186E+00	1.3077E-02	7.1363E-02	1.2999E+00
5	6.9600E+00	7.0179E+00	1.0762E-02	-5.7937E-02	-8.3242E-01
6	7.0200E+00	7.0184E+00	1.0762E-02	1.6177E-03	2.3044E-02
7	8.9300E+00	8.9407E+00	1.3255E-02	-1.0750E-02	-1.2038E-01
8	8.9800E+00	8.9844E+00	1.3367E-02	-4.4202E-03	-4.9222E-02
9	1.0690E+01	1.0674E+01	1.8650E-02	1.5821E-02	1.4800E-01
10	1.0680E+01	1.0675E+01	1.8653E-02	4.9306E-03	4.6166E-02

Rezidualní součet čtverců, RSC	: 9.2425E-03
Průměr absolutních hodnot reziduí, Me	: 1.8814E-02
Průměr relativních reziduí, Mer	: 3.0980E-01
Odhad reziduálního rozptylu, $s^2(e)$: 1.1553E-03
Odhad směrodatné odchylky reziduí, $s(e)$: 3.3990E-02
Odhad šikmosti reziduí, $g_1(e)$: 5.8489E-01
Odhad špičatosti reziduí, $g_2(e)$: 4.3808E+00

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

Fisher-Snedocorův test významnosti regrese, F	: 5.3649E+04
Tabulkový kvantil, $F(1-\alpha, m-1, n-m)$: 5.3177E+00
Závěr: Navržený model je přijat jako významný.	
Spočtená hladina významnosti	: 0.000
Scottovo kritérium multikolinearity, M	: 1.1335E-13
Závěr: Navržený model je korektní.	
Cook-Weisbergův test heteroskedasticity, Sf	: 1.6367E+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)	: 1.3646E+00
Tabulkový kvantil, $\chi^2(1-\alpha, 2)$: 5.9915E+00
Závěr: Normalita je přijata.	
Spočtená hladina významnosti	: 0.505
Waldův test autokorelace, Wa	: 2.7888E+00
Tabulkový kvantil, $\chi^2(1-\alpha, 1)$: 3.8415E+00
Závěr: Rezidua nejsou autokorelována.	
Spočtená hladina významnosti	: 0.095
Znamkový test, Dt	: 1.0062E+00
Tabulkový kvantil, $N(1-\alpha/2)$: 1.6449E+00
Závěr: Rezidua nevykazují trend.	
Spočtená hladina významnosti	: 0.157

(7) INDIKACE VLIVNÝCH BODŮ:

(* indikuje odlehlý nebo vlivný bod)

Bod	Standardizované reziduum eS <i>l</i> i	Jackknife reziduum eJ <i>l</i> i	Predikované reziduum eP <i>l</i> i	Diagonální prvky H <i>l</i> , <i>l</i>
1	-6.8486E-01	-6.6028E-01	-2.7854E-02	3.0156E-01
2	1.1742E-02	1.0984E-02	4.7450E-04	2.9252E-01
3	-4.8060E-02	-4.4963E-02	-1.7702E-03	1.4843E-01
4	2.2746E+00	3.5799E+00*	8.3762E-02	1.4803E-01
5	-1.7970E+00	-2.1767E+00	-6.4392E-02	1.0026E-01
6	5.0174E-02	4.6941E-02	1.7979E-03	1.0025E-01
7	-3.4345E-01	-3.2367E-01	-1.2678E-02	1.5208E-01
8	-1.4144E-01	-1.3247E-01	-5.2288E-03	1.5465E-01
9	5.5676E-01	5.3120E-01	2.2636E-02	3.0106E-01
10	1.7352E-01	1.6262E-01	7.0554E-03	3.0116E-01
Bod	Zobecněné diag. prvky Hm <i>l</i> , <i>l</i>	Cookova vzdálenost D <i>l</i> i	Atkinsonova vzdálenost A <i>l</i> i	Vliv na predikci DF <i>l</i> i
1	3.4251E-01	1.0126E-01	8.6772E-01	-4.3386E-01
2	2.9253E-01	2.8503E-05	1.4125E-02	7.0627E-03
3	1.4867E-01	2.0130E-04	3.7543E-02	-1.8771E-02
4	6.9903E-01*	4.4948E-01*	2.9844E+00*	1.4922E+00*
5	4.6343E-01	1.7991E-01*	1.4532E+00	-7.2659E-01
6	1.0054E-01	1.4026E-04	3.1338E-02	1.5669E-02
7	1.6459E-01	1.0579E-02	2.7415E-01	-1.3708E-01
8	1.5676E-01	1.8298E-03	1.1332E-01	-5.6659E-02
9	3.2814E-01	6.6762E-02	6.9726E-01	3.4863E-01
10	3.0379E-01	6.4881E-03	2.1351E-01	1.0676E-01

Bod	V ě r o h o d n o s t n í v z d á l e n o s t i		
i	LD(b)[i]	LD(s ²)[i]	LD(b,s ²)[i]
1	2.4999E-01	9.9495E-03	2.5196E-01
2	4.9392E-06	5.3588E-02	5.3652E-02
3	5.0323E-04	5.3317E-02	5.3770E-02
4	1.0649E+00	6.1252E+00*	8.9881E+00*
5	4.3995E-01	9.7607E-01	1.6548E+00
6	3.5063E-04	5.3291E-02	5.3607E-02
7	2.6412E-02	3.9748E-02	6.3907E-02
8	4.5736E-03	5.1130E-02	5.5257E-02
9	1.6553E-01	2.1208E-02	1.7748E-01
10	1.6207E-02	4.9898E-02	6.4552E-02

Výstup č. 2

V Ý S T U P

(1) STATISTICKÉ CHARAKTERISTIKY PROMĚNNÝCH:

Proměnná	Průměr	Směrodatná odchylka	Var.koef.
x	1.6533E+00	6.0774E-01	36.759
y	7.3278E+00	2.7146E+00	37.045

Součty čtverců :

Sxx	: 2.9548E+00	Syy	: 5.8951E+01	Sxy	: 1.3198E+01
SumX ²	: 2.7556E+01	SumY ²	: 5.4222E+02	SumXY	: 1.2223E+02

Korelační koeficient : 0.99997

(2) PARAMETRY KALIBRACE:

Parametr	Odhad	Směrodatná odchylka	Test H ₀ : B _i j = 0 vs. H _A : B _i j <> 0	t-kriterium	hypotéza H ₀ je	Hlad.výz.
Úsek	-5.6745E-02	2.1984E-02	2.5812E+00	Zamítnuta		0.036
Směrnice	4.4665E+00	1.2564E-02	3.5550E+02	Zamítnuta		0.000

(3) ANALÝZA REZIDUÍ:

Bod	Měřená hodnota	Predikovaná hodnota	Absolutní reziduum	Relativní reziduum
i	y _{exp} [i]	y _{yyp} [i]	e[i]	er[i]
1	3.5900E+00	3.5928E+00	2.8367E-03	7.9018E-02
2	3.6900E+00	3.6732E+00	-1.6766E-02	-4.5437E-01
3	5.4100E+00	5.3991E+00	-1.0908E-02	-2.0163E-01
4	6.9600E+00	7.0093E+00	4.9266E-02	7.0785E-01
5	7.0200E+00	7.0097E+00	-1.0287E-02	-1.4654E-01
6	8.9300E+00	8.9366E+00	6.5629E-03	7.3492E-02
7	8.9800E+00	8.9803E+00	3.3518E-04	3.7326E-03
8	1.0690E+01	1.0674E+01	-1.5966E-02	-1.4936E-01
9	1.0680E+01	1.0675E+01	-5.0735E-03	-4.7505E-02

Reziduální součet čtverců, RSC	: 3.2650E-03
Průměr absolutních hodnot reziduí, Me	: 1.3111E-02
Průměr relativních reziduí, Mer[z]	: 0.207
Odhad reziduálního rozptylu, s ² (e)	: 4.6643E-04
Odhad směrodatné odchylky reziduí, s(e)	: 2.1597E-02

(4) KALIBRAČNÍ MEZE:

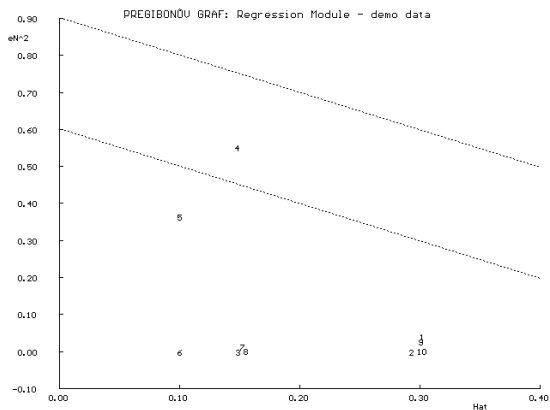
Kritická úroveň,	yc:	-4.7602E-03	xc:	1.1639E-02
Limita detekce,	yd:	4.6575E-02	xd:	2.3132E-02
Mez stanovitelnosti,	ys:	3.0818E-01	xs:	8.1702E-02

(5) KALIBRAČNÍ TABULKA:

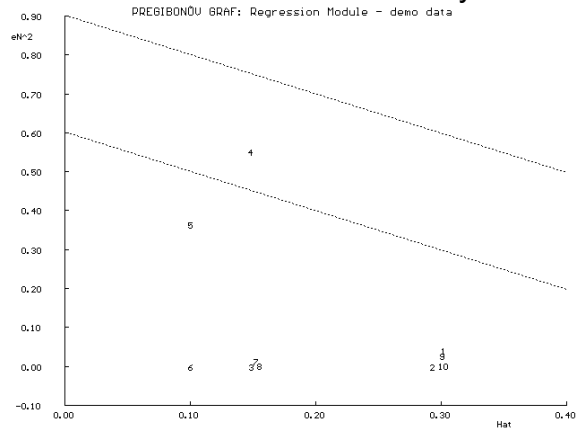
Měřená hodnota $y_{exp[i]}$	Přímý odhad $x_{vyp[i]}$	Naszodiho odhad $x_{vyp[i]}$	Konfidenční interval dolní mez $Ll_{x_{vyp[i]}}$	Konfidenční interval horní mez $Lu_{x_{vyp[i]}}$
4.0000E+00	9.0826E-01	9.0827E-01	8.9523E-01	9.2129E-01
5.5000E+00	1.2441E+00	1.2441E+00	1.2317E+00	1.2564E+00
8.0000E+00	1.8038E+00	1.8038E+00	1.7917E+00	1.8159E+00

Obrázky:

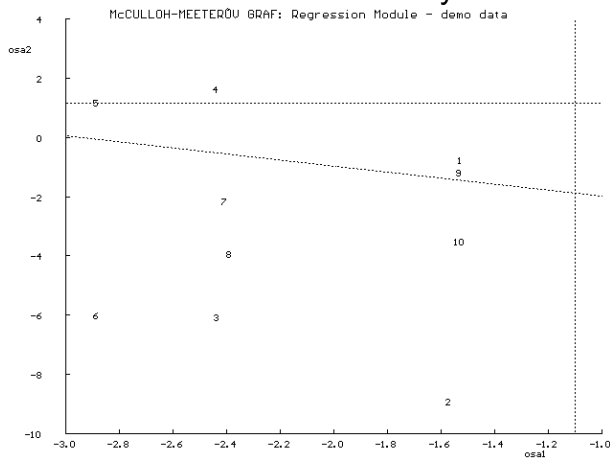
Obr. 1 Grafická indikace odlehlých bodů



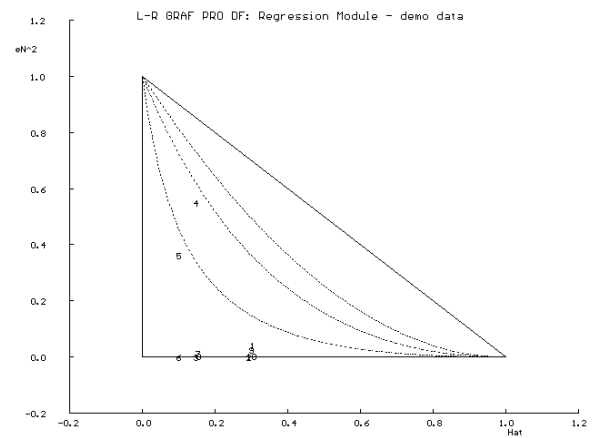
Obr. 2 Grafická indikace odlehlých bodů



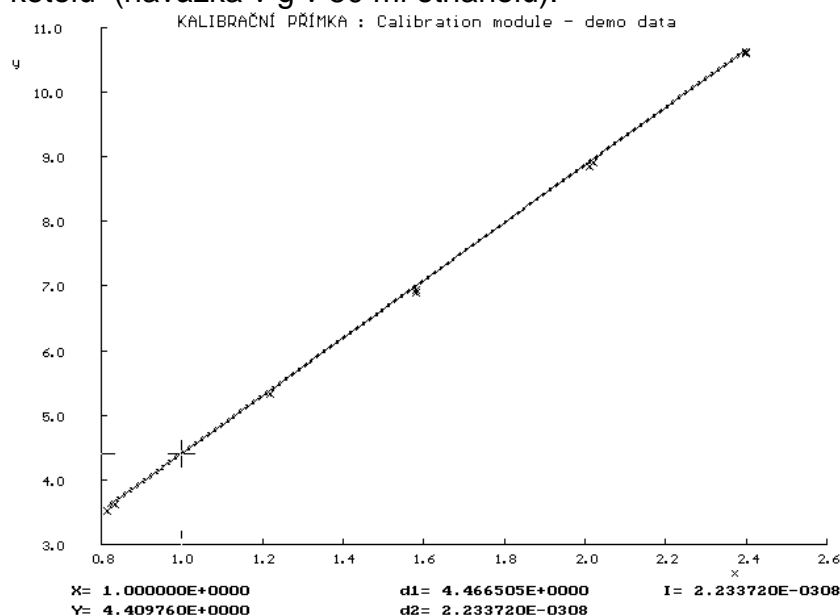
Obr. 3 Grafická indikace odlehlých bodů



Obr. 4 Grafická indikace odlehlých bodů



Obr. 5 kalibrační závislost specifické otáčivosti L-FAK na koncentraci surového ketolu (navážka v g v 50 ml ethanolu).



Řešení:

Vyšetření regresního tripletu.

Výsledky z programu ADSTAT jsou uvedeny jako Výstup č. 1, z kterých vyplývá, že:

1. daná závislost je přímková,
2. rezidua vykazují heteroskedasticitu
3. ve výběru dat je indikován odlehlý bod 4.

Na obrázcích obr.1 až obr. 4 je graficky znázorněna indikace bodu 4 jako odlehlého bodu.

Po odstranění odlehlého bodu 4, bylo dosaženo homoskedasticity reziduí a vylepšily se statistické charakteristiky regrese, uvedené v Tabulce 2.

Tabulka 2 Srovnání původních a vylepšených statistických charakteristik regrese

	Původní výběr	Výběr bez odlehlého bodu 4
R_p^2	0.99990	0,99996
MEP	0,0012695	0,00052084
AIC	-65,865	-67,296

Přímková závislost bez odlehlého bodu 4 je vhodná jako model uvedené kalibrační závislosti.

Navržení kalibrační závislosti.

Parametry kalibrační závislosti jsou uvedené ve Výstupu č. 2 z programu ADSTAT i s kalibrační tabulkou pro stanovení L-FAK ve třech vzorcích surového ketolu s naměřenou specifickou otáčivostí α : 4,00°, 5,50° a 8,00°.

Závislost kalibrační přímky lze popsat rovnicí (rov.1):

$$\alpha = -0,06(0,02) + 4,47(0,01).c \quad (\text{rov.1}),$$

kde c je koncentrace ketolu v ethanolu (g/50 ml).

Kalibrační přímka je uvedena na obr. 5.

Závěr:

Metoda lineární kalibrační závislosti specifické otáčivosti na koncentraci ketolu je vhodná na mezioperační stanovení L-FAK v surovém ketolu při výrobě L-efedrinium chloridu.

Úloha 2. Nelineární kalibrace**Zadání:**

Železité ionty s tironem (dvojsodná sůl 1,2 dihydroxybenzen – 3,5 – disulfonové kyseliny) tvoří komplex, který při pH = 9,6 vykazuje absorpční maximum při vlnové délce $\lambda = 480$ nm. Absorbance roste s rostoucím nadbytkem tironu vůči Fe^{3+} iontům. Úkolem je určit kalibrační závislost absorbance na koncentračním poměru tiron/ Fe^{3+} ($[\text{Fe}^{3+}] = 8 \cdot 10^{-5}$ mol/l) pro naměřené hodnoty absorbance 0,200, 0,250 a 0,280 určit koncentrační poměry tiron/ Fe^{3+} .

Data:

Hodnoty uvedené závislosti jsou v Tabulce 1.

Tabulka 1 Hodnoty absorbance v závislosti na koncentračním poměru tiron/ Fe^{3+}

tiron/ Fe^{3+}	1,0	1,5	2,0	2,5	3,0	4,0	5,0
A	0,105	0,172	0,235	0,290	0,331	0,380	0,405

Užitý program:

ADSTAT modul lineární regrese a modul kalibrace.

Výstup:

Výstup č. 1

U Ý S T U P**(1) PŘEDBĚŽNÁ STATISTICKÁ ANALÝZA:**

Proměnná	Průměr	Směrodatná odchylna	Párový korelační koeficient	Spočtená hladina významnosti
y	2.7400E-01	1.0984E-01	1.0000	-----
x1	2.7143E+00	1.4100E+00	0.9578	0.001

Párové korelační koeficienty mezi dvojicemi vysvětlujících proměnných

Spočtená hladina významnosti

(2) INDIKACE MULTIKOLINEARITY:

Č [j]	Vlastní čísla korel. matice	Čísla podmíněnosti	Variance inflation factor	Vícenás.korel. koef pro X[j]
1	1.0000E+00	1.0000E+00	1.0000E+00	0.0000
Maximální číslo podmíněnosti K (K [j], K > 1000 indikuje silnou multikolaritu) (VIF [j] > 10 indikuje silnou multikolaritu)				: 1.0000E+00

(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. v
BI 01	7.1485E-02	3.0170E-02		2.3694E+00	Akceptována	0.06
BI 11	7.4611E-02	1.0017E-02		7.4485E+00	Zamítnuta	0.00

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R	: 9.5777E-01
Koeficient determinace, R^2	: 9.1733E-01
Predikovaný korelační koeficient, R_p^2	: 8.7206E-01
Střední kvadratická chyba predikce, MEP	: 2.4768E-03
Akaikeho informační kritérium, AIC	: -4.5452E+01

(5) ANALÝZA KLASICKÝCH REZIDUÍ:

Bod	Meřená hodnota	Predikovaná hodnota	Směrodatná odchylka	Klasické reziduum	Relativní reziduum
i	y _{exp[i]}	y _{vyp[i]}	s(y _{vyp[i]})	e _i	e _{r[i]}
1	1.0500E-01	1.4610E-01	2.1584E-02	-4.1096E-02	-3.9139E+01
2	1.7200E-01	1.8340E-01	1.7859E-02	-1.1401E-02	-6.6286E+00
3	2.3500E-01	2.2071E-01	1.4906E-02	1.4293E-02	6.0823E+00
4	2.9000E-01	2.5801E-01	1.3251E-02	3.1988E-02	1.1030E+01
5	3.3100E-01	2.9532E-01	1.3386E-02	3.5683E-02	1.0780E+01
6	3.8000E-01	3.6993E-01	1.8353E-02	1.0072E-02	2.6505E+00
7	4.0500E-01	4.4454E-01	2.6367E-02	-3.9539E-02	-9.7627E+00

Rezidualní součet čtverců, RSC	: 5.9844E-03
Průměr absolutních hodnot reziduí, Me	: 2.6296E-02
Průměr relativních reziduí, Mer	: 1.2296E+01
Odhad reziduálního rozptylu, $s^2(e)$: 1.1969E-03
Odhad směrodatné odchylky reziduí, s(e)	: 3.4596E-02
Odhad šikmosti reziduí, g ₁ (e)	: -2.8914E-01
Odhad špičatosti reziduí, g ₂ (e)	: 1.5702E+00

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

Fisher-Snedocorův test významnosti regrese, F	: 5.5481E+01
Tabulkový kvantil, $F(1-\alpha, m-1, n-m)$: 6.6079E+00
Závěr: Navržený model je přijat jako významný.	
Spočtená hladina významnosti	: 0.001
Scottovo kritérium multikolinearity, M	: 3.3307E-16
Závěr: Navržený model je korektní.	
Cook-Weisbergův test heteroskedasticity, Sf	: 7.9803E+00
Tabulkový kvantil, $\chi^2(1-\alpha, 1)$: 3.8415E+00
Závěr: Rezidua vykazují heteroskedasticitu.	
Spočtená hladina významnosti	: 0.005
Jarque-Berraův test normality reziduí, L(e)	: 6.9381E-01
Tabulkový kvantil, $\chi^2(1-\alpha, 2)$: 5.9915E+00
Závěr: Normalita je přijata.	
Spočtená hladina významnosti	: 0.707

Waldův test autokorelace, W_a : 1.5160E+00
 Tabulkový kvantil, $\chi^2(1-\alpha,1)$: 3.8415E+00
 Závěr: Rezidua nejsou autokorelována.
 Spočtená hladina významnosti : 0.218

Znaménkový test, D_t : -7.8824E-01
 Tabulkový kvantil, $N(1-\alpha/2)$: 1.6449E+00
 Závěr: Rezidua nevykazují trend.
 Spočtená hladina významnosti : 0.215

(7) INDIKACE VLIVNÝCH BODŮ:

(* indikuje odlehlý nebo vlivný bod)

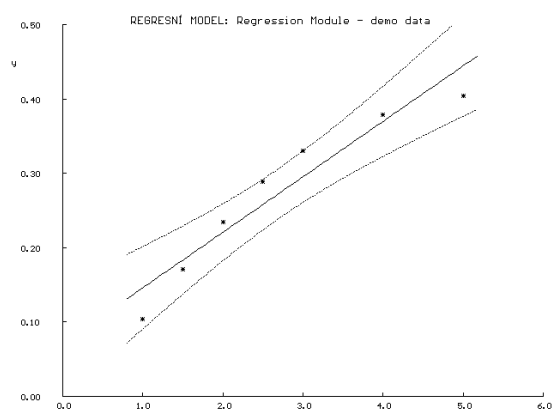
Bod	Standardizované reziduum $eS[i]$	Jackknife reziduum $eJ[i]$	Predikované reziduum $eP[i]$	Diagonální prvky $H[i, i]$
1	-1.5200E+00	-1.8536E+00	-6.7284E-02	3.8922E-01
2	-3.8478E-01	-3.4937E-01	-1.5543E-02	2.6647E-01
3	4.5782E-01	4.1835E-01	1.7551E-02	1.8563E-01
4	1.0010E+00	1.0012E+00	3.7488E-02	1.4671E-01
5	1.1185E+00	1.1554E+00	4.1965E-02	1.4970E-01
6	3.4344E-01	3.1087E-01	1.4017E-02	2.8144E-01
7	-1.7653E+00	-2.5723E+00	-9.4329E-02	5.8084E-01*

Bod	Zobecněné diag. prvky $H_m[i, i]$	Cookova vzdálenost $D[i]$	Atkinsonova vzdálenost $A[i]$	Vliv na predikci $DF[i]$
1	6.7143E-01	7.3611E-01*	2.3395E+00*	-1.4797E+00*
2	2.8819E-01	2.6892E-02	3.3294E-01	-2.1057E-01
3	2.1977E-01	2.3889E-02	3.1581E-01	1.9974E-01
4	3.1769E-01	8.6128E-02	6.5639E-01	4.1514E-01
5	3.6246E-01	1.1013E-01*	7.6651E-01	4.8478E-01
6	2.9839E-01	2.3099E-02	3.0762E-01	1.9455E-01
7	8.4207E-01	2.1590E+00*	4.7876E+00*	-3.0280E+00*

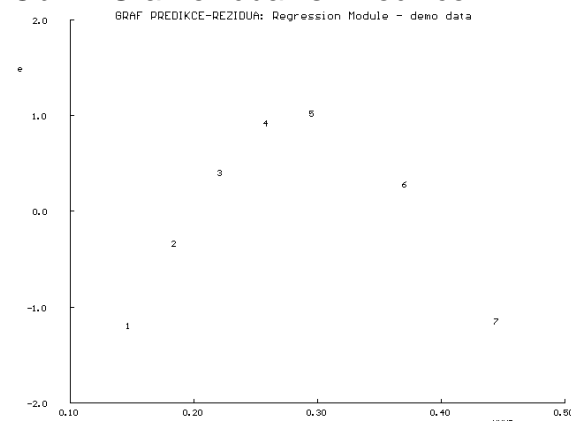
Bod	V ě r o h o d n o s t n í v z d á l e n o s t i		
i	$LD(b)[i]$	$LD(s^2)[i]$	$LD(b, s^2)[i]$
1	1.8066E+00	8.9259E-01	4.1767E+00
2	7.4896E-02	5.1734E-02	1.1824E-01
3	6.6571E-02	4.1811E-02	1.0165E-01
4	2.3710E-01	1.7289E-02	2.7580E-01
5	3.0177E-01	6.5573E-02	4.1810E-01
6	6.4380E-02	5.6905E-02	1.1368E-01
7	4.3576E+00	3.1710E+00	1.6924E+01*

Obrázky:

Obr. 1 Graf regresní křivky



Obr. 2 Graf rezidua vs. Predikce

**Řešení:****Vyšetření regresního tripletu.**

Výsledky vyšetření regresního tripletu a parametry lineární regrese jsou uvedeny ve výstupu č.1. Navržený model lineární regrese vykazuje heteroskedasticitu reziduí a i grafy regresní přímky – obr. 1, a analýzy reziduí – obr. 2, svědčí o tom, že lineární regresní model, vyjádřený (rov.1), není vhodný pro uvedenou kalibrační závislost.

$$A = 0,075(0,010) \cdot [\text{tiron}/\text{Fe}^{3+}] \quad (\text{rov.1}),$$

kde odhad parametru b_0 má nulovou hodnotu.

Metody lineárního, kvadratického a kubického spline lze v takovém případě použít pro nalezení optimálního tvaru kalibrační závislosti. Výsledky jsou uvedeny v Tabulce 2.

Tabulka 2

Metoda	Počet uzlových bodů	X_D	σ_e
Lineární spline	1	4,96E-1	9E-3
	2	2,35E-1	3E-3
	3	4,85E-1	4E-3
Kvadratický spline	1	3,19E-1	3E-3
	2	2,60E-1	1E-3
	3	2,72E-1	2E-4
Kubický spline	1	3,66E-1	9E-4
	2	1,141	6E-4
	3	Mnoho uzlů	2E-8

Z Tabulky 2 vyplývá, že optimální model kalibrační závislosti poskytuje metoda kvadratického spline se třemi uzlovými body. Za nejlepší kalibrační model se považuje takový, který má nejmenší limitu detekce x_D a odhad směrodatné odchylky reziduí σ_e při nejnižším počtu uzlových bodů.

Kalibrační rovnici lze vyjádřit podle (rov.2):

$$A = f_i \cdot [\text{tiron}/\text{Fe}^{3+}]^2 + g_i \cdot [\text{tiron}/\text{Fe}^{3+}] + h_i \quad \text{pro body } k_{i-1} < [\text{tiron}/\text{Fe}^{3+}] \leq k_i \quad (\text{rov.2}).$$

Koeficienty kalibrační rovnice jsou uvedeny v Tabulce 3.

Tabulka 3 Koeficienty kalibrační rovnice

bod k_i	f_i	g_i	h_i
2,000	-7E-3	1,51E-1	-3,9E-2
3,000	-2,7E-2	2,32E-1	-1,19E-1
4,000	-2,0E-2	1,87E-1	-5,3E-2
5,000	-4E-3	6,3E-2	1,96E-1

Inverzní hodnoty koncentračního poměru $[\text{tiron}/\text{Fe}^{3+}]$ pro naměřené hodnoty absorbance: 0,200, 0,250 a 0,280 jsou uvedeny v Tabulce 4.

Tabulka 4 Inverzní hodnoty koncentračního poměru $[\text{tiron}/\text{Fe}^{3+}]$ pro naměřené hodnoty absorbance: 0,200, 0,250 a 0,280

Naměřená A_{exp}	Inverzní odhad $[\text{tiron}/\text{Fe}^{3+}]$	Konfidenční interval	
		Dolní mez	Horní mez
0,200	1,72	1,70	1,74
0,250	2,12	2,10	2,14
0,280	2,40	2,37	2,43

Závěr:

Uvedená kalibrační závislost, kvadratický spline se třemi uzlovými body, je vhodná pro stanovení koncentračního poměru $[\text{tiron}/\text{Fe}^{3+}]$, za uvedených experimentálních podmínek, z naměřených hodnot absorbancí.

Úloha 3. Rozlišení mezi lineární a nelineární kalibrací

Zadání:

Roztok aromatického ketonu v tetrachlormethanu vykazuje při vlnočtu 1680 cm^{-1} absorpční maximum a absorbance roztoku roste s rostoucí koncentrací aromatického ketonu. Pro tři naměřené hodnoty absorbance 0,030, 0,060 a 0,180 je třeba určit odpovídající koncentrace.

Data:

Hodnoty kalibrační závislosti absorbance na koncentraci aromatického ketonu v tetrachlormethanu jsou uvedeny v Tabulce 1.

Tabulka 1 Hodnoty kalibrační závislosti

c [mol/l]	A
0,095	0,020
0,189	0,041
0,250	0,055
0,311	0,071
0,482	0,113
0,897	0,220

Užitý program:

ADSTAT, moduly lineární regrese a kalibrace.

Výstup:

Výstup č. 1

V Ý S T U P

(1) ZVOLENÁ STRATEGIE REGRESNÍ ANALÝZY:

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

(2) PODMÍNKY A KVANTILY PRO STATISTICKÉ TESTY:

Hladina významnosti, alfa	:	0.050
Počet bodů, n	:	6
Počet parametrů, m	:	2
Kvantil Studentova rozdělení $t(1-\alpha/2, n-m)$:	2.776
Kvantil rozd. Chí-kvadrát $\text{Chi-square}(1-\alpha, m)$:	5.991
Jméno výstupního souboru	:	RESULTS.TXT

(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.
BI 01	-6.3604E-03	1.2633E-03	-5.0349E+00	Zamítnuta		0.007
BI 11	2.5097E-01	2.7778E-03	9.0350E+01	Zamítnuta		0.000

(4) STATISTICKÉ CHARAKTERISTIKY REGRESE:

Vícenásobný korelační koeficient, R	:	9.9976E-01
Koeficient determinace, R^2	:	9.9951E-01
Predikovaný korelační koeficient, R_p^2	:	9.9854E-01
Střední kvadratická chyba predikce, MEP	:	1.2806E-05
Akaikeho informační kritérium, AIC	:	-7.4321E+01

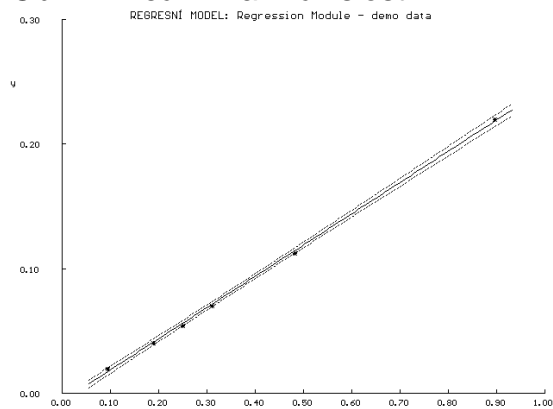
(6) TESTOVÁNÍ REGRESNÍHO TRIPLETU (DATA + MODEL + METODA):

Fisher-Snedocorův test významnosti regrese, F	:	8.1631E+03
Tabulkový kvantil, $F(1-\alpha, m-1, n-m)$:	7.7086E+00
Závěr: Navržený model je přijat jako významný.	:	
Spočtená hladina významnosti	:	0.000
Scottovo kritérium multikolinearity, M	:	-2.5868E-14
Závěr: Navržený model je korektní.	:	
Cook-Weisbergův test heteroskedasticity, Sf	:	5.5744E+00
Tabulkový kvantil, $\text{Chi}^2(1-\alpha, 1)$:	3.8415E+00
Závěr: Rezidua vykazují heteroskedasticitu.	:	
Spočtená hladina významnosti	:	0.018
Jarque-Berraův test normality reziduí, L(e)	:	6.1203E-01
Tabulkový kvantil, $\text{Chi}^2(1-\alpha, 2)$:	5.9915E+00
Závěr: Normalita je přijata.	:	
Spočtená hladina významnosti	:	0.736
Waldův test autokorelace, W_a	:	1.1856E-06
Tabulkový kvantil, $\text{Chi}^2(1-\alpha, 1)$:	3.8415E+00
Závěr: Rezidua nejsou autokorelována.	:	
Spočtená hladina významnosti	:	0.999

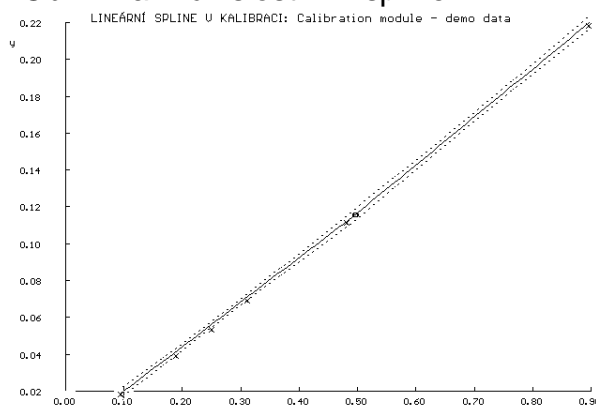
Znamkový test, Dt : -1.7678E-01
 Tabulkový kvantil, N(1-alpha/2) : 1.6449E+00
 Závěr: Rezidua nevykazují trend.
 Spočtená hladina významnosti : 0.430

Obrázky:

Obr.1 Lineární kal. závislost



Obr.2 Kal. závislost lin. spline

**Řešení:**

Vyšetření regresního tripletu.

Nejprve se provede vyšetření regresního tripletu. Výsledky jsou uvedeny ve výstupu č. 1 z programu ADSTAT. Výsledky ukazují, že model lineární kalibrační křivky je vhodný, ale analýza reziduí vykazuje heteroskedasticitu. Jako vlivné body se jeví bod 1 a bod 6. Vzhledem k malému počtu bodů není vhodné vyloučit indikovaný odlehlý bod 1.

Rovnice kalibrační přímky lze vyjádřit podle (rov.1):

$$A = -0,006(0,001) + 0,251(0,003).c \quad (\text{rov.1})$$

Kalibrační závislost je znázorněná na obr. 1

Vhodné je model kalibrační závislosti zpřesňovat pomocí metod spline.

Výsledky vyšetřování pomocí metody lineárního, kvadratického a kubického spline jsou uvedeny v Tabulce 2.

Tabulka 2

Metoda	Počet uzlových bodů	X_D	σ_e
Přímka	-	0,027	0,002
Lineární spline	1	0,0287	0,001
	2	0,047	0,001
	3	0,146	0,001
Kvadratický spline	1	0,05	0,001
	2	0,78	0,0005
	3	Mnoho uzlů	2E-8
Kubický spline	1	0,232	0,0006
	2	Mnoho uzlů	2E-8

Z Tabulky 2 vyplývá, že optimální model kalibrační závislosti poskytuje metoda lineárního spline s jedním uzlovým bodem.

Kalibrační rovnici lze vyjádřit podle (rov.2):

$$A = g_i \cdot c + h_i \quad \text{pro body } k_{i-1} < c \leq k_i \quad (\text{rov.2}).$$

Koeficienty kalibrační rovnice jsou uvedeny v Tabulce 3. a kalibrační závislost je znázorněná na obr.2.

Tabulka 3 Koeficienty kalibrační rovnice

bod k_i	g_i	h_i
0,496	0,242	-0,004
0,897	0,260	-0,013

Inverzní hodnoty koncentrace aromatického ketonu v tetrachlormethanu pro naměřené hodnoty absorbance 0,030, 0,060 a 0,180 jsou uvedeny v Tabulce 4.

Tabulka 4 Inverzní hodnoty koncentrace pro naměřené hodnoty absorbance 0,030, 0,060 a 0,180

Naměřená A_{exp}	Inverzní odhad c	Konfidenční interval	
		Dolní mez	Horní mez
0,030	0,141	0,132	0,150
0,060	0,265	0,259	0,272
0,180	0,743	0,734	0,753

Závěr:

Lineární spline model s jedním uzlovým bodem nejlépe vystihuje průběh kalibrační závislosti absorbance na koncentraci aromatického ketonu v tetrachlormethanu.