

**Univerzita Pardubice
Chemicko-technologická fakulta
Katedra analytické chemie**

**12. licenční studium PYTHAGORAS
Statistické zpracování dat**

Kalibrace a limity její přesnosti

Semestrální práce
2009

RNDr. Markéta Vaňková, Ph.D.
Endokrinologický ústav
Národní 8, 116 94 Praha 1

Úvod do problematiky detekce RET/PTC přeskupení u papilárního karcinomu štítné žlázy

Papilární karcinom štítné žlázy (PTC) je nejčastějším diferencovaným karcinomem štítné žlázy s maximem výskytu je kolem 40. roku věku, ale postihuje i mladší věkové skupiny včetně dětí. Nejčastější příčinou je aktivace proto-onkogenu *RET*. *RET* onkogen je aktivován *RET/PTC* přeskupením, které vzniká fúzí intracelulární tyrozinkinázové domény *RET* genu a jiného konstitutivně exprimovaného genu. *RET/PTC* je spojováno s agresivnější formou PTC a často tvorbou vzdálených metastáz.

Real Time PCR

Polymerázová řetězová reakce (PCR) je základní metodou v molekulární genetice a slouží k namnožení (amplifikaci) vybraného úseku molekuly DNA. Množství amplifikovaného produktu závisí v každé PCR na množství templátové DNA, která je přidána do reakce. Reakce tedy probíhá tzv. kvantitativně. Přesné určení tohoto množství je možné pomocí sondy označené fluorescenčním barvivem. Real Time PCR monitoruje množství PCR produktu během amplifikace v její exponenciální fázi. Speciální termocyklery (light cyklery), určené pro kvantitativní PCR v reálném čase, jsou schopny v průběhu PCR ozařovat vzorek excitačním zářením, které vybudí fluorescenci uvolněného barviva. Tuto fluorescenci přístroj po každém cyklu změří a výsledek předá řídicímu softwaru, který zobrazuje průběžně (v reálném čase) množství uvolněné fluorescence, které odpovídá množství vzniklého produktu. V průběhu PCR tak dochází k uvolňování dalších a dalších molekul fluorescenčního barviva, takže roste fluorescence reakční směsi. Tento nárůst je přímo úměrný množství produktu, který v reakci vzniká.

Při Real Time PCR se využívá sonda, která má sekvenci specifickou pro určitý produkt. Na sondu je napojen tzv. reporter - R (fluorescenční barvivo) a quencher - Q (zhášec). Pokud je reporter navázán na sondě v blízkosti Q, není schopen fluorescence. Sonda musí být navržena tak, aby hybridizovala s templátovou DNA za stejných podmínek jako primery. Během PCR nasedne na templátovou DNA uvnitř cílové sekvence. DNA-polymeráza, která provádí extenzi jednoho ze dvou primerů, narazí při syntéze komplementárního vlákna na nasedlou sondu. Protože DNA-polymeráza má kromě své schopnosti syntetizovat komplementární vlákno

také tzv. exonukleázovou aktivitu, štěpí nasedlou sondu na jednotlivé nukleotidy. Tím se reporter dostane z těsné blízkosti zhášedce a je schopen fluorescence.

Výsledek reakce tak známe často dřív, než proběhnou všechny cykly. Kromě tohoto urychlení a vypuštění zdlouhavé detekce produktu na gelu je tato technika hlavně výhodná tam, kde potřebujeme znát přesné množství vstupní templátové DNA - zejména u sledování exprese genů.

Absolutní kvantifikace slouží ke zjištění přesného množství DNA. Pro určení koncentrace neznámého vzorku je potřeba vytvořit standardní křivku (kalibrační křivku z pozitivní kontroly s několika řády přesné koncentrace), se kterou se porovnává počet cyklů, po kterých fluorescence vzorku vzroste nad pozadí fluorescence a na základě toho se stanoví přesné množství koncentrace vzorku přidané do reakce. Hlavní výhodou této metody, ze které těží také lékařská diagnostika, je její vysoká citlivost. Protože dokáže z pouhých 2 vláken vytvořit po 30 cyklech 2^{30} molekul produktu (107 milionů), proto stačí pro pomnožení určité sekvence velmi malé množství DNA.

Relativní kvantifikace slouží ke zjištění změn v expresi genů. Při této metodě se nezjišťuje přesná koncentrace genů, ale to, zda došlo ke snížení či zvýšení exprese určitého genu. Například během rakovinného bujení, kdy je změněná regulace některých genů (především těch odpovědných za proliferaci buněk, buněčný cyklus a apoptózu). Protože je potřeba odstranit rozdíly mezi jednotlivými vzorky (liší se kvalita RNA a množství buněk, z kterého byla RNA izolována), normalizuje se množství fluorescenčního signálu cílového genu k referenčnímu genu. Většinou se jedná o tzv. housekeeping geny, které se vyskytují a exprimují ve všech tkáních stejně. Je potřeba vytvořit standardní křivku jak pro cílový gen, tak pro referenční (nejedná se o typickou kalibrační křivku, není potřeba znát přesnou koncentraci, stačí vytvořit tzv. ředící řadu s několika řády). Dále se do reakce přidává pozitivní kontrola, která zároveň slouží jako kalibrátor k porovnání vzorků v jednom běhu i porovnání jednotlivých běhů mezi sebou. Pro přesnější analýzu se amplifikace genu provádí v dubletech nebo tripletech (jeden vzorek ve dvou až třech reakcích).

Detekce fúzních forem RET genu

Kromě detekce jednotlivých typů RET/PTC přeskupení se provádí ještě amplifikace několika kontrolních genů. Exprese *GAPDH* genu slouží ke kontrole proběhlé

reverzní transkripce a potvrzení přítomnosti cDNA. Původ folikulárních buněk je potvrzen expresí genu **PAX8**. Pro relativní kvantifikaci používáme housekeeping gen **β -actin**, kódující vysoce konzervovaný protein zahrnutý v motilitě, struktuře a integritě buněk.

Exprese extracelulární domény (EC) *RET* genu ukazuje, že se jedná o medulární karcinom štítné žlázy a nikoliv PTC. Je to proto, že *RET* proto-onkogen se za normálních okolností ve folikulárních buňkách štítné žlázy neexprimuje. Pokud dojde k přeskupení a vzniku *RET/PTC* formy, *RET* onkogen extracelulární doménu ztrácí a exprimuje se pouze tyrozinkinázová doména.

Při pozitivní EC doméně se provádí PCR amplifikace **kalcitoninu**, který se exprimuje u medulárního karcinomu štítné žlázy.

Ke zjištění přítomnosti fúzní formy *RET* genu se provádí exprese tyrozinkinázové domény (TK) *RET* genu.

V současné době je popsáno 11 forem *RET/PTC* přeskupení. V naší laboratoři provádíme detekci dvou nejčastějších fúzních forem - ***RET/PTC1*** a ***RET/PTC3***.

Úkol 1 Lineární kalibrace

V tomto úkolu byly použity certifikované standardy od výrobce (rozdíl od úlohy 2, kde byly použity vlastní standardy vyrobené ředěním pozitivní kontroly).

Zadání úlohy

Vytvořit kalibrační křivku koncentrace PCR produktu housekeepingového genu **β -actin**, která bude sloužit pro kvantifikaci genetického materiálu po izolaci z tkáně štítné žlázy pacientů. Pomocí kalibrační křivky určit koncentraci DNA ve vzorcích 5 pacientů.

**Data pro kalibrační křivku (3 měření pro každou koncentraci)
a data 5 pacientů (2 měření)**

concentration	CP	patient	CP measurement1	CP measurement2
1	21.05	A	21	21.1
0.1	24.86	B	21.9	22
0.01	27.95	C	20.9	20.7
0.001	31.42	D	23.4	23.9
0.0001	35	E	27.3	27.6
LOG concentration	LOG CP	patient	LOG CP measurement 1	LOG CP measurement2
0	1.32325	A	1.322219295	1.324282455
-1	1.3955	B	1.340444115	1.342422681
-2	1.44638	C	1.320146286	1.315970345
-3	1.49721	D	1.369215857	1.378397901
-4	1.54407	E	1.436162647	1.440909082

concentration - ředící řada standardu

CP - počet cyklů, po kterých je vzorek přístrojem detekován (fluorescence se dostává ze šumu)

Vzhledem k ředění standardu v rozsahu 5 řádů byla u dat pro vytvoření modelu použita logaritmická transformace.

Program: QC.Expert 3.1

KALIBRACE data.xls

kalibrace lineární.vts

Výsledek

Kalibrační model - lineární regrese

1. Kritika dat

Sleduje se především výskyt vlivných bodů, které mohou způsobovat zkreslení odhadů a zvyšovat rozptyl. Vlivné body lze dělit do tří skupin, 1) vlivné body vzniklé hrubou chybou při měření, 2) vlivné body s vysokým vlivem a 3) vlivné body jako důsledek nesprávně navrženého regresního modelu. Podle výskytu dělíme vlivné body na vybočující (liší se hodnotami na ose y) a na extrémny (liší se hodnotami na ose x). Některé body jsou vybočující i extrémny zároveň.

K odhalení vlivných bodů se používá analýza projekční matice a analýza reziduí, vhodné jsou grafické diagnostiky.

1.1.Regresní diagnostika

Vícenásobná lineární regrese

Hladina významnosti :	0.05
Kvantil $t(1-\alpha/2, n-m)$:	3.18244630528356
Kvantil $F(1-\alpha, m, n-m)$:	10.127964486013
Absolutní člen :	Ano
Počet platných řádků :	5
Počet parametrů :	2
Metoda :	Nejmenší čtverce
Sloupce pro výpočet :	LOG CP Abs LOG concentration
Transformace :	Bez transformace

Základní analýza

Charakteristiky proměnných

Proměnná	Průměr	Směr.Odch.	Kor.vs.Y	Významnost
LOG concentration	-2	1.58113883	-0.99622	0.000277859

Analýza rozptylu

Průměr Y :	1.441281852			
Zdroj	Součet čtverců	Průměrný čtverec		Rozptyl
Celková variabilita	0.02974543885	0.005949087769		0.007436359711
Variabilita vysvětlená modelem	0.02952150358	0.005904300716		0.007380375895
Reziduální variabilita	0.0002239352652	4.478705303E-005		5.598381629E-005
Hodnota kritéria F :	395.4915751			
Kvantil F (1- α , m-1, n-m) :	10.12796449			
Pravděpodobnost :	0.0002778598211			
Závěr :	Model je významný			

Odhady parametrů

Proměnná	Odhad	Směr.Odch.	Závěr	Pravděpodobnost	Spodní mez	Horní mez
Abs	1.332	0.0066923	Významný	2.792842E-007	1.311316	1.353912
LOG concentration	-0.054	0.0027321	Významný	0.0002778598	-0.06302	-0.045638

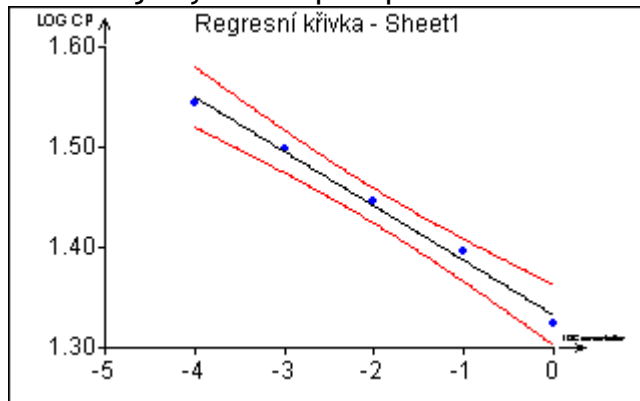
Statistické charakteristiky regrese

Vícenásobný korelační koeficient R :	0.9962286936
Koeficient determinace R^2 :	0.99247161
Predikovaný korelační koeficient R_p :	0.9365517795
Střední kvadratická chyba predikce MEP:	0.0001918220666
Akaikeho informační kritérium :	-46.06795728

1.2. Grafické diagnostiky

Průběh regresního modelu

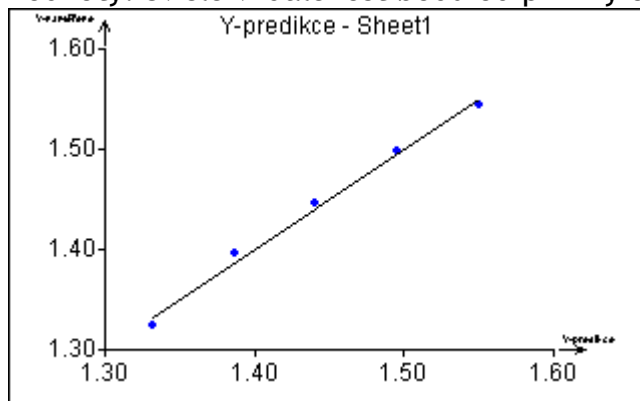
Červeně je vyznačen pás spolehlivosti modelu na zadané hladině významnosti.



Rezidua

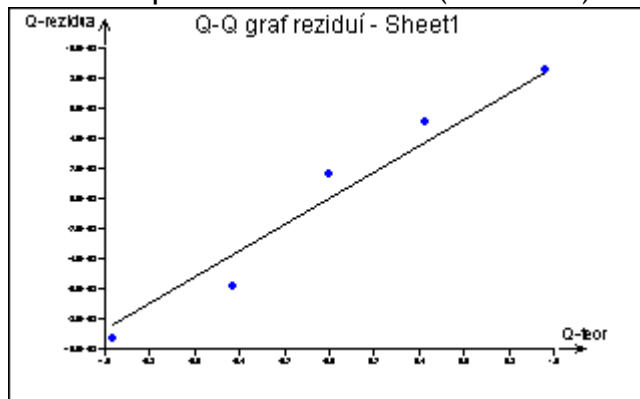
Graf vyjadřující těsnost proložení

Na ose X jsou vypočítané hodnoty závisle proměnné, na ose Y jsou naměřené hodnoty. Svislé vzdálenost bodu od přímky odpovídá reziduu.



Q-Q graf pro posouzení normality reziduí

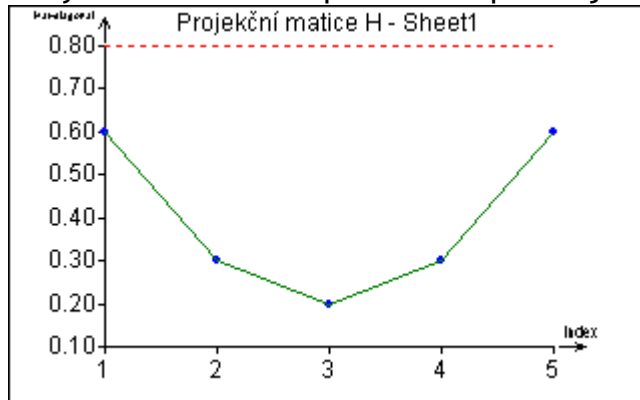
Přímka odpovídá normálnímu (Gausovu) rozdělení reziduí.



Vlivná data

Projekční matice

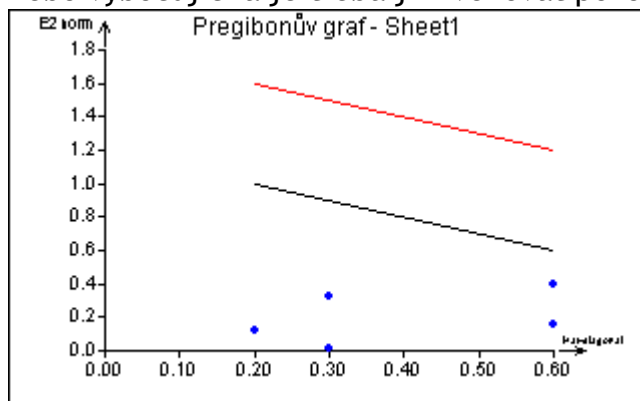
Diagonální prvky projekční matice vyjadřují míru vlivu jednotlivých dat na regresi. Body nad vodorovnou přímkou se považují za silně vlivné.



V datech nejsou silné vlivné body.

Pregibonův graf

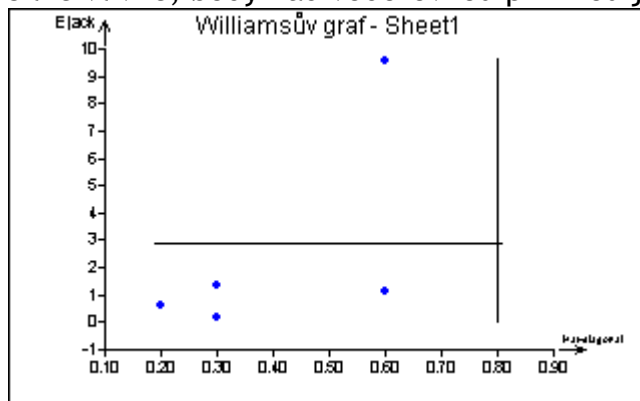
Graf pro společné posouzení vybočujících bodů a vlivných bodů. Body nad nižší (černou) přímkou se považují za vlivné, nad vyšší (červenou) přímkou za silně vlivné nebo vybočující a je třeba jim věnovat pozornost.



V datech nejsou vlivné nebo vybočující body.

Williamsův graf

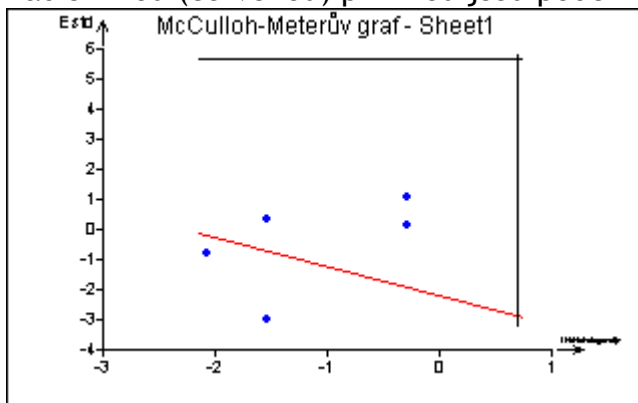
Slouží k indikaci vlivných i vybočujících bodů. Body vpravo od svislé přímky jsou silně vlivné, body nad vodorovnou přímkou jsou silně vybočující.



V datech je jeden silně vybočující bod (1).

McCulloh-Meterův graf

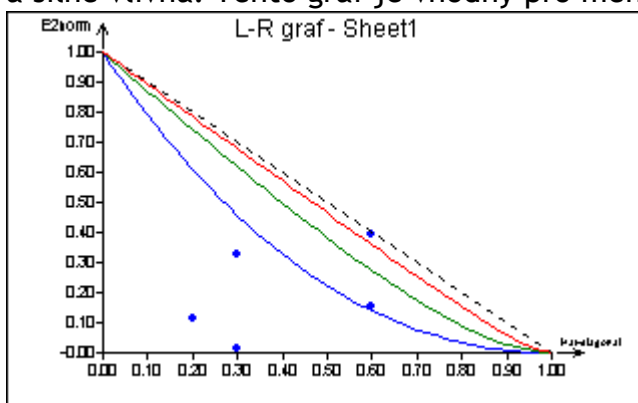
Další alternativa k indikaci vlivných a vybočujících bodů. Body vpravo od svislé přímky jsou silně vlivné, body nad vodorovnou přímkou jsou silně vybočující. Body nad šikmou (červenou) přímkou jsou podezřelé vybočující nebo vlivné.



V datech jsou tři podezřelé vybočující nebo vlivné body (1,2,5).

L-R graf

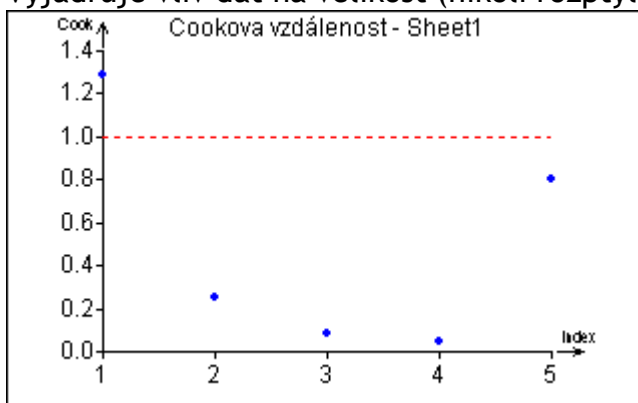
Další alternativa k indikaci vlivných bodů. Hyperbolické křivky jsou linie stejného vlivu. Podle polohy bodů vůči třem křivkám lze data rozdělit na slabě vlivná, vlivná a silně vlivná. Tento graf je vhodný pro menší rozsahy dat.



V datech je jeden vlivný bod (1).

Cookova vzdálenost

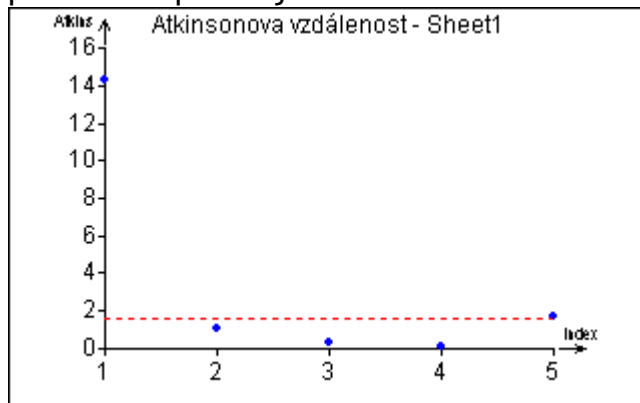
Vyjadřuje vliv dat na velikost (nikoli rozptyl) odhadovaných parametrů.



V datech je jeden vybočující bod (1).

Atkinsonova vzdálenost

Je další diagnostikou k posouzení vlivu jednotlivých dat. Data nad vodorovnou přímkou se považují za silně vlivná.



V datech je jeden silně vlivný bod (1).

Testování regresního tripletu

Fisher-Snedecorův test významnosti modelu

Hodnota kritéria F : 395.4915751

Kvantil F (1-alfa, m-1, n-m) : 10.12796449

Pravděpodobnost : 0.0002778598211

Závěr : **Model je významný**

Scottovo kritérium multikolinearity

Hodnota kritéria SC : -0.961262106

Závěr : **Model je korektní.**

Cook-Weisbergův test heteroskedasticity

Hodnota kritéria CW : 0.2803530358

Kvantil $\chi^2(1-\alpha, 1)$: 3.841458829

Pravděpodobnost : 0.5964699179

Závěr : **Rezidua vykazují homoskedasticitu.**

Jarque-Berrův test normality

Hodnota kritéria JB : 0.5023167893

Kvantil $\chi^2(1-\alpha, 2)$: 5.991464547

Pravděpodobnost : 0.7778991467

Závěr : **Rezidua mají normální rozdělení.**

Waldův test autokorelace

Hodnota kritéria WA : 0.08099963131
Kvantil $\chi^2(1-\alpha,1)$: 3.841458829
Pravděpodobnost : 0.7759472846
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 : 0.1091089451
Kvantil $N(1-\alpha/2)$: 1.959963999
Pravděpodobnost : 0.9131160801
Závěr : **V reziduích není trend.**

Závěr kritiky dat

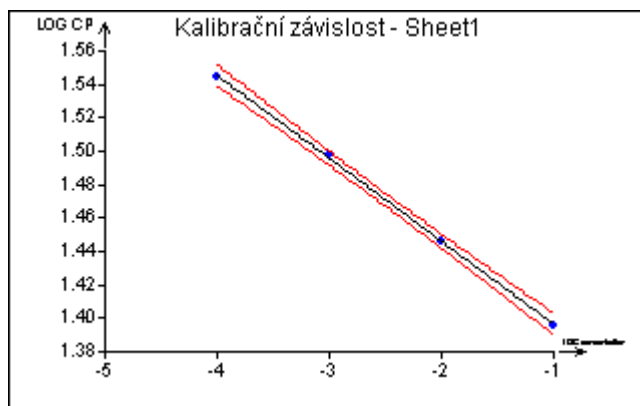
V datech byl identifikován jeden vlivný bod (1). Jeho eliminací Akaikovo informační kritérium stoupá (z -103 na -50) a střední kvadratická chyba predikce MEP klesá (z 0,00018 na 0,000007). Přesto bude bod eliminován.

3. Kalibrace

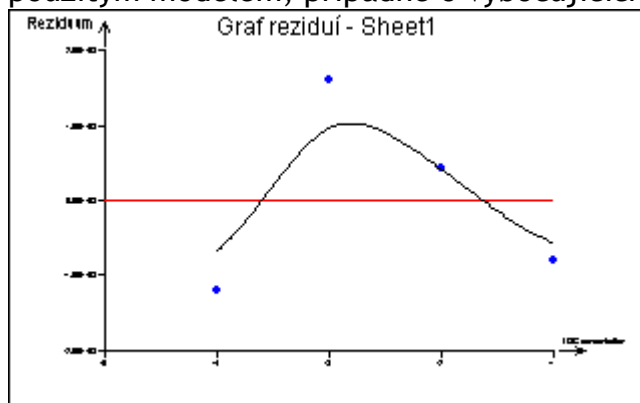
3.1. Tvorba modelu

Na základě grafických diagnostik byl vybrán lineární model.

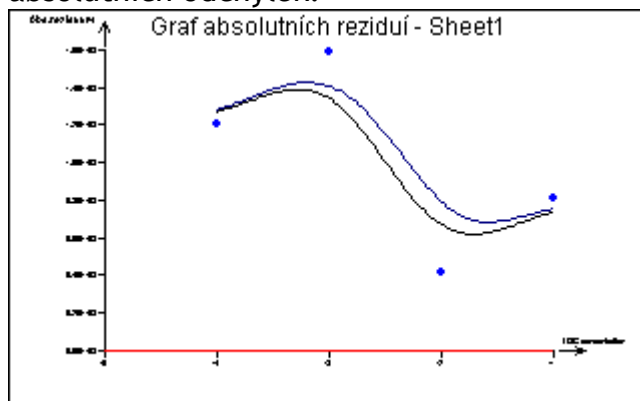
Graf kalibrační závislosti s vyznačeným pásem spolehlivosti (červené čáry).



Graf reziduí proložený neparametrickým jádrovým modelem. Výrazné zakřivení může svědčit o nelinearitě kalibrační závislosti, kterou se nepodařilo postihnout použitým modelem, případně o vybočujících bodech.



Graf absolutních reziduí s neparametrickým jádrovým proložením. Horní křivka je odmocninou z proložení čtverců odchylek, spodní křivka je přímým proložením absolutních odchylek.



3.2. Použití modelu

Počet dat: 4
Hladina významnosti : 0.05
Volba kalibračního modelu : Automatická
Použitý kalibrační model : Lineární
Vhodnost použitého modelu : Vyhovuje
Použita vážená regrese : Ne

Parametry kalibračního modelu

Parametr	Odhad	Sm. odchylka	Spodní mez	Horní mez
Abs.	1.3466	0.0018974957	1.3384937	1.3548222
X	-0.049	0.0006928674	-0.052633	-0.0466713

Významnost absolutního členu

Hodnota	Spodní mez	Horní mez	Závěr
1.346658008	1.338493743	1.354822273	Významný

Citlivost metody : -0.0496525129

Zvolený faktor K : 1.959963999

Vypočítaná sm.odch. slepého signálu : 0.001549298783

Kalibrační meze

c..kritická

d..detekce

q..kvantifikace

Metoda	Yc	Yd	Yq	Xc	Xd	Xq
Metoda podle ISO 11843-2						
	1.4295	1.50900	1.51235	-1.66855	-3.269	-3.337
Přímá metoda analytu						
	1.35608	1.35652	1.356967	-0.18986	-0.1987	-0.2076
Přímá metoda signálu, IUPAC						
	1.35608	1.36491	1.373207	-0.18986	-0.3677	-0.5347
Kombinovaná metoda Ebel,Kamm						
	1.34712	1.35652	1.356967	-0.00949	-0.19875	-0.2076
Metoda K*Sigma z regrese						
	1.35608	1.3655	1.374939	-0.1898	-0.37973	-0.5695
Metoda K*Sigma, ACS						
	1.34969	1.3527	1.355767	-0.0611	-0.12231	-0.18346

Kalibrační tabulka					
pacient	Zpětný odhad	Spodní mez	Horní mez	Naměřené hodnoty	
A	0.4714188997	0.85697	0.08828	1.3222	1.3242
B	0.105223476	0.47422	-0.2614	1.3404	1.3424
C	0.5759968797	1.34547	-0.18856	1.3201	1.3159
D	-0.5467773817	1.12783	-2.21227	1.3692	1.3783
E	-1.850417051	-0.9850	-2.71322	1.4361	1.4409

Analýza reziduí

Reziduální součet čtverců : 4.800653439E-006

Půměrné absolutní reziduum : 0.00100470625

Korelační koeficient : 0.9998053336

měření	Naměřené X	Naměřené Y	Vypočítané Y	Reziduum
1	-1	1.395501124	1.396310521	-0.00080
2	-2	1.446381812	1.445963034	0.000418
3	-3	1.497206181	1.495615547	0.001590
4	-4	1.544068044	1.54526806	-0.001200

Závěr

Kalibrace má dvě části: tvorbu modelu a použití tohoto modelu. Při tvorbě modelu se získané hodnoty y porovnávají s danými hodnotami x , které jsou např. certifikovanými standardy. Závislost y na x se popíše regresním modelem. Použití kalibračního modelu pak spočívá v odhadu neznáme hodnoty x na základě jednoho nebo více měření y . Tento zpětný odhad musí být charakterizovaný intervalem spolehlivosti na hladině významnosti α . Interval spolehlivosti odhadu x je dán šířkou pásu spolehlivosti kalibračního modelu. K určení minimálních měřitelných hodnot a úrovně šumu jsou vypočítávány kalibrační meze (kritická hodnota, meze detekce a meze stanovení).

V této úloze představují hodnoty x řadící řadu standardu (certifikované standardy od výrobce) v rozmezí 5 řádů. Každá hodnota y se měřila 3 krát. Měření vzorků od 5 pacientů se vždy 2 krát opakovalo.

Kalibrační model této úlohy je lineární (přímka). Kalibrační meze byly vypočítány pěti metodami, které nabízí modul kalibrace. Výsledek použití modelu je vidět v kalibrační tabulce.

Pod kritickou úrovní (c), tj. pod nejmenší hodnotou rozeznatelnou od šumu, je zpětný odhad pro vzorek E (všechny metody) a vzorek D (5 metod ze 6). Meze detekce (d), tj. hodnoty nad níž můžeme s pravděpodobností $1-\alpha$ prokázat přítomnost vzorku, dosáhl zpětný odhad pro vzorek A, B a C (všechny metody), vzorek E a D (1 metoda ze 6). Mez kvantifikace Y_q , tj. hodnota nad níž lze stanovit skutečnou hodnotu y s relativní chybou menší než α , byla dosažena u zpětného odhadu u vzorků A, B a C (všechny metody) a vzorek D a E (1 metoda z 6).

Pod mezí kvantifikace není doporučena kvantitativní analýza, o kterou v tomto případě jde především, vzorky D a E proto nebudou doporučeny pro další analýzy. Vzorky A, B a C mají dostatečnou koncentraci pro další analýzy.

Úkol 2 Nelineární kalibrace

V tomto úkolu byly použity standardy vyrobené ředěním z pozitivní kontroly. Tento postup předcházet použití certifikovaných standardů od výrobce (rozdíl od úlohy 1).

Zadání úlohy

Vytvořit kalibrační křivku koncentrace PCR produktu housekeepingového genu β -actin, která bude sloužit pro kvantifikaci genetického materiálu po izolaci z tkáně štítné žlázy pacientů. Pomocí kalibrační křivky určit koncentraci DNA ve vzorcích 5 pacientů.

Data pro kalibrační křivku (průměr ze 3 měření pro každou koncentraci) a data 5 pacientů (2 měření)

concentration	CP	patient	CP measurement1	CP measurement2
1	20,82	A	21	21,1
0,1	24,42	B	21,9	22
0,01	27,95	C	20,9	20,7
0,001	31,28	D	23,4	23,9
0,0001	34,65	E	27,3	27,6
LOG concentration	LOG CP	patient	LOG CP measurement 1	LOG CP measurement2
0	1,3184807	A	1.322219295	1.324282455
-1	1,3877457	B	1.340444115	1.342422681
-2	1,4463818	C	1.320146286	1.315970345
-3	1,4952667	D	1.369215857	1.378397901
-4	1,5397032	E	1.436162647	1.440909082

concentration - ředící řada pozitivního vzorku

CP - počet cyklů, po kterých je vzorek přístrojem detekován (fluorescence se dostává ze šumu)

Vzhledem k rozsahu ředění standardu přes 5 řádů, bude pro sestavení modelu použita logaritmická transformace.

Program: QC.Expert 3.1

KALIBRACE data.xls
kalibrace nelineární.vts

Výsledek

Kalibrační model - lineární regrese

1. Kritika dat

Sleduje se především výskyt vlivných bodů, které mohou způsobovat zkreslení odhadů a zvyšovat rozptyl. Vlivné body lze dělit do tří skupin, 1) vlivné body vzniklé hrubou chybou při měření, 2) vlivné body s vysokým vlivem a 3) vlivné body jako důsledek nesprávně navrženého regresního modelu. Podle výskytu dělíme vlivné body na vybočující (liší se hodnotami na ose y) a na extrémy (liší se hodnotami na ose x). Některé body jsou vybočující i extrémy zároveň.

K odhalení vlivných bodů se používá analýza projekční matice a analýza reziduí, vhodné jsou grafické diagnostiky.

1.1. Regresní diagnostika

Vícenásobná lineární regrese

Hladina významnosti : 0,05
Kvantil $t(1-\alpha/2, n-m)$: 3,18244630528356
Kvantil $F(1-\alpha, m, n-m)$: 10,127964486013
Absolutní člen : Ano
Počet platných řádků : 5
Počet parametrů : 2
Metoda : Nejmenší čtverce
Sloupce pro výpočet : LOG mean CP
Abs
LOG mean concentration
Transformace : Bez transformace

Základní analýza

Charakteristiky proměnných

Proměnná	Průměr	Směr. Odch.	Kor. vs. Y	Významnost
LOG mean concentration	-2	1,58113883	-0,99579	0,00032751

Analýza rozptylu

Průměr Y :	1,437515636		
Zdroj	Součet čtverců	Průměrný čtverec	Rozptyl
Celková variabilita	0,03050246631	0,006100493261	0,007625616577
Variabilita vysvětlená modelem	0,03024627243	0,006049254487	0,007561568109
Reziduální variabilita	0,000256193872	5,123877439E-005	6,404846799E-005
Hodnota kritéria F :	354,18028		
Kvantil F (1-alfa, m-1, n-m) :	10,12796449		
Pravděpodobnost :	0,0003275196833		
Závěr :	Model je významný		

Odhady parametrů

Proměnná	Odhad	Směr.Odch.	Závěr	Pravděpodobnost	Spodní mez	Horní mez
Abs	1,327	0,0071581	Významný	3,45698E-007	1,304742	1,350302
LOG mean concentration	-0,054	0,0029222	Významný	0,00032751	- 0,06429	-0,045696

Statistické charakteristiky regrese

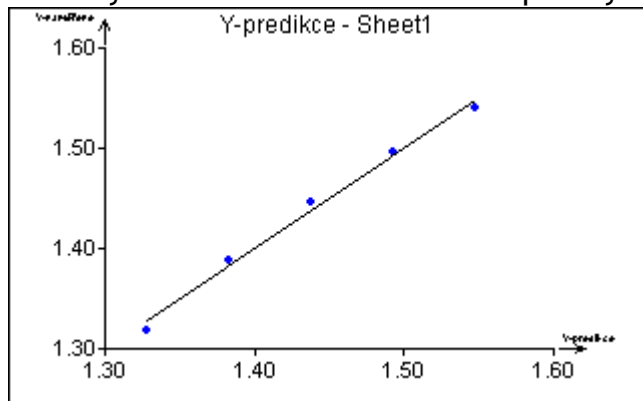
Vícenásobný korelační koeficient R :	0,9957915847
Koeficient determinace R ² :	0,9916008801
Predikovaný korelační koeficient Rp :	0,9300722651
Střední kvadratická chyba predikce MEP :	0,000217162046
Akaikeho informační kritérium :	-45,39507

1.2. Grafické diagnostiky

Rezidua

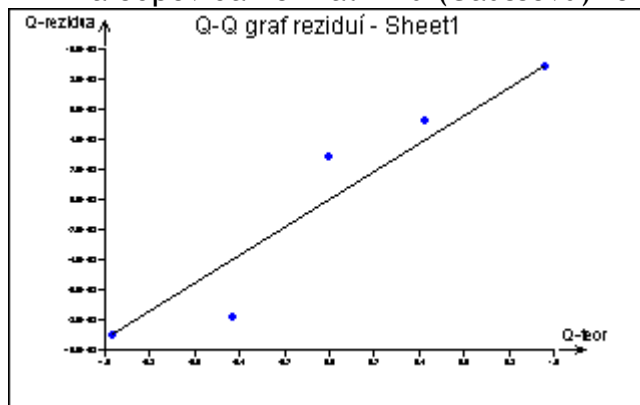
Graf vyjadřující těsnost proložení

Na ose X jsou vypočítané hodnoty závisle proměnné, na ose Y jsou naměřené hodnoty. Svislé vzdálenosti bodů od přímky odpovídá reziduu.



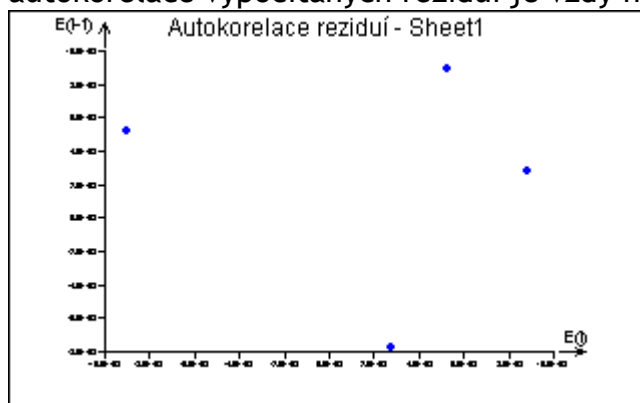
Q-Q graf pro posouzení normality reziduí

Přímka odpovídá normálnímu (Gaussovu) rozdělení reziduí.



Grafické posouzení autokorelace reziduí prvního řádu

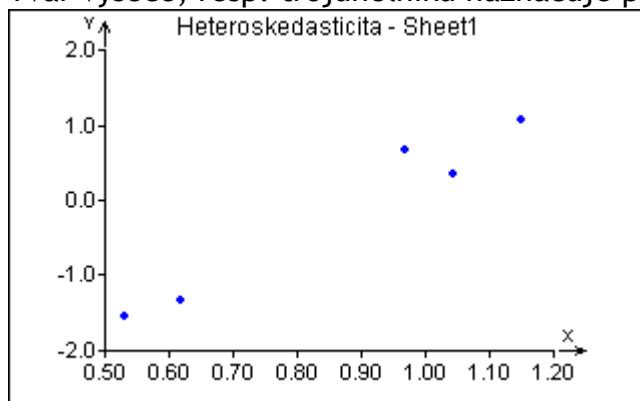
Na ose X je i -té reziduum, na ose Y je $(i-1)$ reziduum. "Mrak" bodů s kladnou směrnici, naznačuje pozitivní autokorelaci, klesající trend negativní autokorelaci. Autokorelace reziduí nemusí nutně dokazovat autokorelaci chyb, neboť autokorelace vypočítaných reziduí je vždy nenulová.



Rezidua nejsou autokorelována.

Grafické posouzení heteroskedasticity (nekonstantnosti rozptylu)

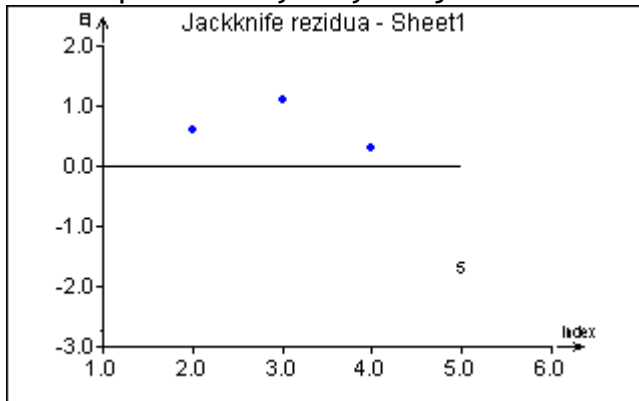
Tvar výseče, resp. trojúhelníku naznačuje přítomnost heteroskedasticity.



Data mají konstantní rozptyl.

Jackknife rezidua

Daleko citlivěji indikují vybočující měření, než klasická rezidua. V přítomnosti většího počtu blízkých vybočujících hodnot však mohou selhat.

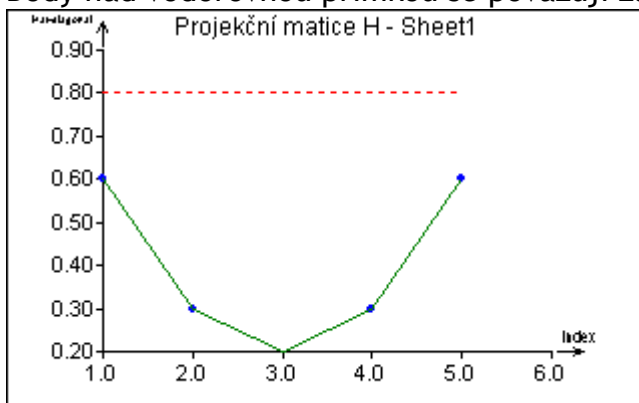


V datech jsou vybočující body (1,5).

Vlivná data

Projekční matice

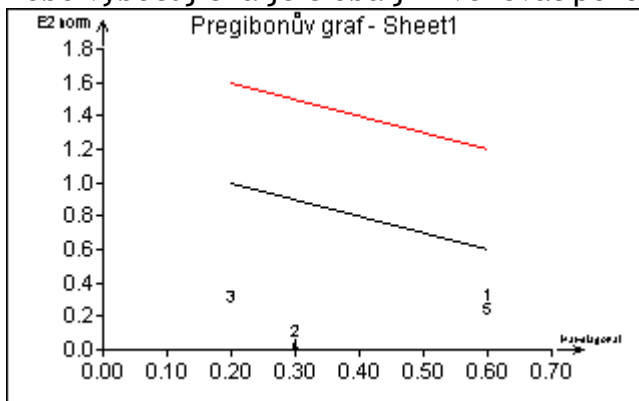
Diagonální prvky projekční matice vyjadřují míru vlivu jednotlivých dat na regresi. Body nad vodorovnou přímkou se považují za silně vlivné.



V datech nejsou silné vlivné body.

Pregibonův graf

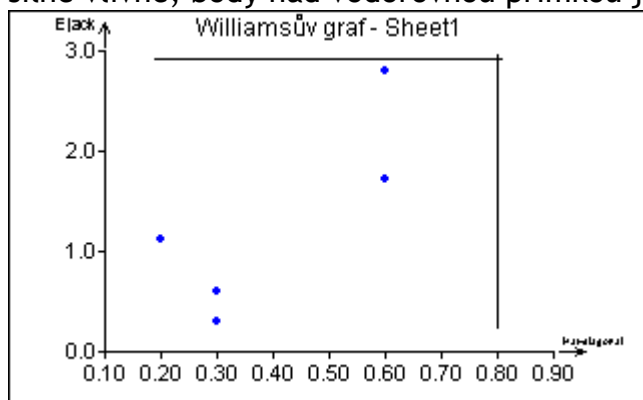
Graf pro společné posouzení vybočujících bodů a vlivných bodů. Body nad nižší (černou) přímkou se považují za vlivné, nad vyšší (červenou) přímkou za silně vlivné nebo vybočující a je třeba jim věnovat pozornost.



V datech nejsou vlivné nebo vybočující body.

Williamsův graf

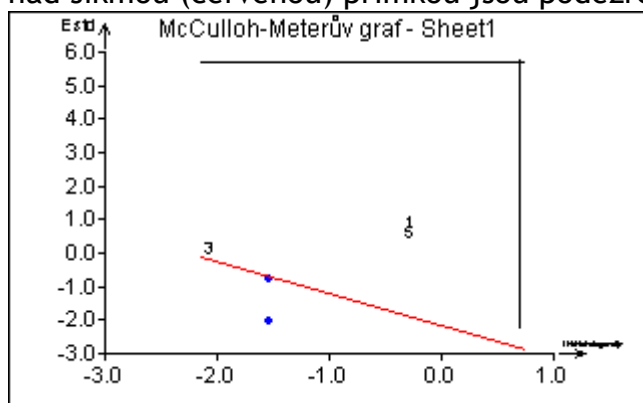
Slouží k indikaci vlivných i vybočujících bodů. Body vpravo od svislé přímky jsou silně vlivné, body nad vodorovnou přímkou jsou silně vybočující.



V datech nejsou vlivné nebo vybočující body.

McCulloh-Meterův graf

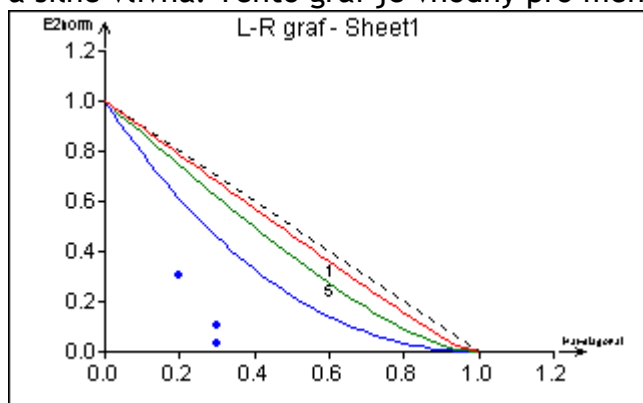
Další alternativa k indikaci vlivných a vybočujících bodů. Body vpravo od svislé přímky jsou silně vlivné, body nad vodorovnou přímkou jsou silně vybočující. Body nad šikmou (červenou) přímkou jsou podezřelé vybočující nebo vlivné.



V datech jsou dva podezřelé vybočující nebo vlivné body (1,5).

L-R graf

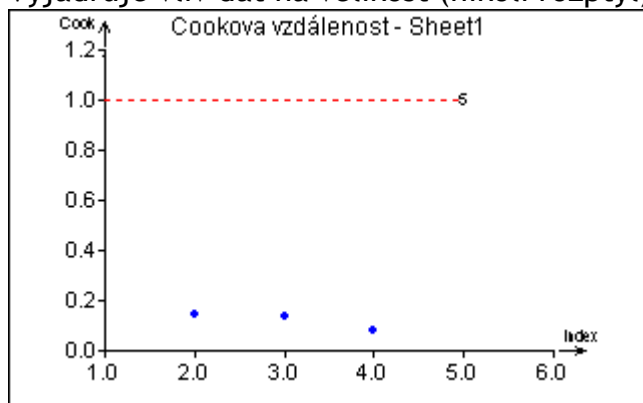
Další alternativa k indikaci vlivných bodů. Hyperbolické křivky jsou linie stejného vlivu. Podle polohy bodů vůči třem křivkám lze data rozdělit na slabě vlivná, vlivná a silně vlivná. Tento graf je vhodný pro menší rozsahy dat.



V datech jsou dva vlivné body (1,5).

Cookova vzdálenost

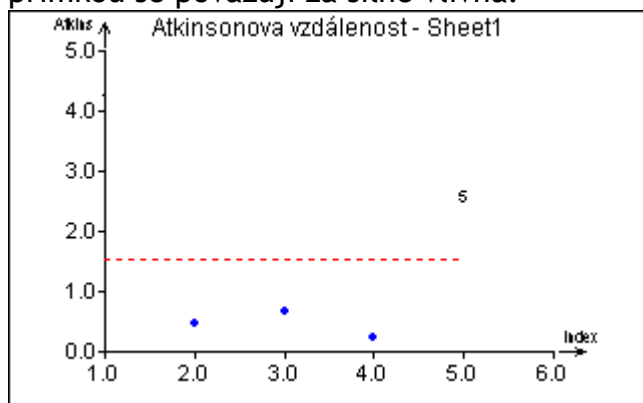
Vyjadřuje vliv dat na velikost (nikoli rozptyl) odhadovaných parametrů.



V datech jsou dva vlivné body (1,5).

Atkinsonova vzdálenost

Je další diagnostikou k posouzení vlivu jednotlivých dat. Data nad vodorovnou přímkou se považují za silně vlivná.



V datech jsou dva silně vlivné body (1,5).

Testování regresního tripletu

Fisher-Snedecorův test významnosti modelu

Hodnota kritéria F : 354,18028

Kvantil F (1-alfa, m-1, n-m) : 10,12796449

Pravděpodobnost : 0,0003275196833

Závěr : Model je významný

Scottovo kritérium multikolinearity

Hodnota kritéria SC : -0,9600436536
Závěr : Model je korektní.

Cook-Weisbergův test heteroskedasticity

Hodnota kritéria CW : 0,02582970621
Kvantil $\chi^2(1-\alpha,1)$: 3,841458829
Pravděpodobnost : 0,8723169068
Závěr : Rezidua vykazují homoskedasticitu.

Jarque-Berrův test normality

Hodnota kritéria JB : 0,6163510937
Kvantil $\chi^2(1-\alpha,2)$: 5,991464547
Pravděpodobnost : 0,7347863172
Závěr : Rezidua mají normální rozdělení.

Waldův test autokorelace

Hodnota kritéria WA : 0,0001720869163
Kvantil $\chi^2(1-\alpha,1)$: 3,841458829
Pravděpodobnost : 0,9895334987
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 : 0,1091089451
Kvantil $N(1-\alpha/2)$: 1,959963999
Pravděpodobnost : 0,9131160801
Závěr : V reziduích není trend.

Závěr kritiky dat

V datech byly identifikovány dva vlivné body (1, 5). Jejich eliminací Akaikovo informační kritérium AIC stoupne (z -45 na -32) a střední kvadratická chyba predikce MEP klesne (z 0,0002 na 0,00007). Body budou v modelu ponechány.

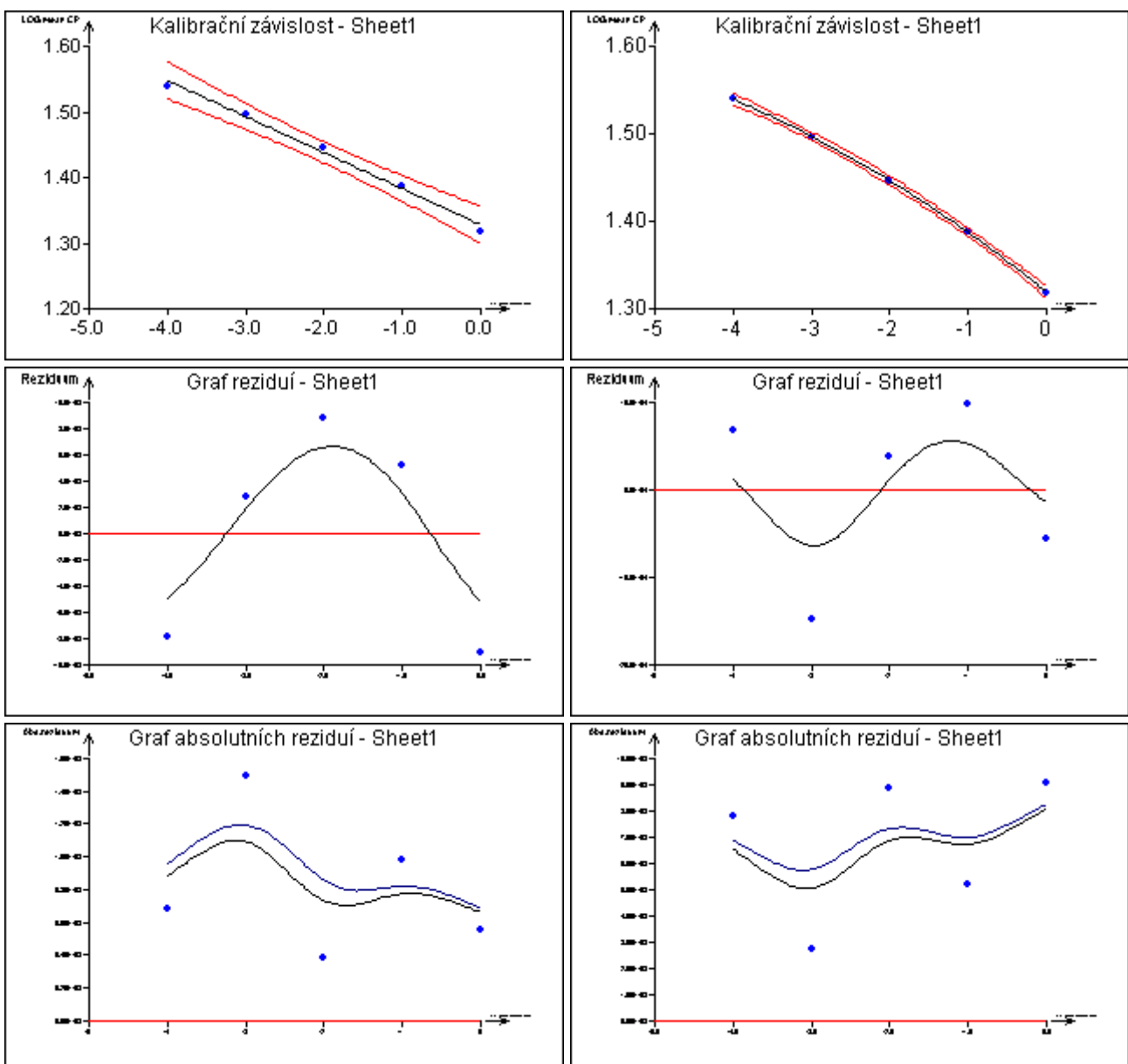
3. Kalibrace

Posouzení vhodnosti modelu podle grafických diagnostik

3.1. Tvorba modelu

a) lineární model - přímka

b) nelineární model (kvadratický) - parabola



Na základě grafických diagnostik byl vybrán kvadratický model.

3.2. Použití modelu

Počet dat: 5
Hladina významnosti : 0,05
Volba kalibračního modelu : Automatická
Použitý kalibrační model : Kvadratický
Vhodnost použitého modelu : Vyhovuje
Použita vážená regrese : Ne

Parametry kalibračního modelu

Parametr	Odhad	Sm. odchylka	Spodní mez	Horní mez
Abs.	1,3190	0,0013473889	1,31323819	1,3248328
X	-0,071	0,0015960876	-0,0788377	-0,065102
X ²	-0,004	0,0003826327	-0,0058897	-0,002597

Významnost absolutního členu

Hodnota	Spodní mez	Horní mez	Závěr
1,319035542	1,313238196	1,324832889	Významný

Citlivost metody

Citlivost v nule : -0,07197035406
Citlivost v počátku : -0,03802286834
Citlivost ve středu : -0,08894409691
Citlivost na konci : -0,07197035406

Zvolený faktor K : 1,959963999
Vypočítaná sm.odch. slepého signálu : 0,001431680603

Kalibrační meze

c..kritická

d..detekce

q..kvantifikace

Metoda	Yc	Yd	Yq	Xc	Xd	Xq
Metoda podle ISO 11843-2	1,39573	1,41794	1,47242	-0,953260	-1,508	-1,906
Přímá metoda analytu	1,32622	1,32655	1,32688	-0,100541	-0,105	-0,109
Přímá metoda signálu, IUPAC	1,32622	1,33282	1,33894	-0,100541	-0,193	-0,281
Kombinovaná metoda Ebel,Kamm	1,31939	1,32655	1,32688	-0,004997	-0,105	-0,109
Metoda K*Sigma z regrese	1,32622	1,33342	1,34061	-0,100541	-0,202	-0,305
Metoda K*Sigma, ACS	1,32184	1,32464	1,32745	-0,039078	-0,078	-0,117

Kalibrační tabulka

pacient	Zpětný odhad	Spodní mez	Horní mez	Naměř. hodnoty	
A	-0.058774	0.2025177	-0.3262685	1.322	1.324
B	-0.317139	-0.059690	-0.5815151	1.340	1.342
C	0.0135673	0.5264056	-0.5275914	1.320	1.315
D	-0.798632	0.3835288	-2.1950598	1.369	1.378
E	-1.865627	-1.141375	-2.6797809	1.436	1.440

Analýza reziduí

Reziduální součet čtverců : 4,099418697E-006

Půměrné absolutní reziduum : 0,0008175024343

Korelační koeficient : 0,9999327996

měření	Naměřené X	Naměřené Y	Vypočítané Y	Reziduum
1	0	1,318480725	1,319035542	-0,0005548
2	-1	1,38774566	1,386762461	0,00098319
3	-2	1,446381812	1,446002507	0,00037930
4	-3	1,495266744	1,496755683	-0,0014889
5	-4	1,539703239	1,539021987	0,00068125

Závěr

V této úloze představují hodnoty x ředící řadu standardu v rozmezí 5 řádů. Každá hodnota y se měřila 3 krát a vypočítal se průměr. Měření vzorků od 5 pacientů se vždy 2 krát opakovalo. Data byla logaritmicky transformována.

Kalibrační model této úlohy je kvadratický (parabola). Kalibrační meze byly vypočítány pěti metodami, které nabízí modul kalibrace. Výsledek použití modelu je vidět v kalibrační tabulce. Pod kritickou úrovní (c), tj. pod nejmenší hodnotou rozeznatelnou od šumu, je zpětný odhad pro vzorek E (všechny metody) a vzorek D a B (5 metod ze 6) a vzorek A (2 ze 6 metod). Meze detekce (d), tj. hodnoty nad níž

můžeme s pravděpodobností $1-\alpha$ prokázat přítomnost vzorku, dosáhl zpětný odhad pro vzorek C (všechny metody), vzorek A (4 metody ze 6), vzorek D a B (1 metoda ze 6). Mez kvantifikace Y_q , tj. hodnota nad níž lze stanovit skutečnou hodnotu y s relativní chybou menší než α , byla dosažena u zpětného odhadu pouze pro vzorek C a A (všechny metody) a vzorek D, B a E (2 metody z 6).

A protože pod mezí kvantifikace není doporučena kvantitativní analýza, o kterou v tomto případě jde především, nelze vzorky E, B a D s jistotou použít pro další analýzy. Vzorky C a A mají dostatečnou koncentraci pro další analýzy.