Univerzita Palackého v Olomouci

Lineární regrese II

Chemometrie I (ACH/CHEX1)

1

Univerzita Palackého v Olomouci

Regresní diagnostika

2

Univerzita Palackého
v Olomouci

Úvod

- Regresní diagnostika umožňuje detailní analýzu regresního modelu s využitím grafů a statistických testů. Umožňuje během tvorby modelu interaktivní zásah uživatele, který zná „svá data“ lépe než software. Blíží se tak EDA pro analýzu jednorozměrných dat.
- Regresní diagnostika obsahuje postupy ke společnému posouzení tzv. regresního tripletu:
  - kvality dat pro regresní model (přítomnost vlivných bodů),
  - vhodnost regresního modelu pro daná data,
  - splnění předpokladů pro metodu odhadu regresních koeficientů, nejčastěji MNČ.

Regresní triplet: data + regresní model + metoda odhadu

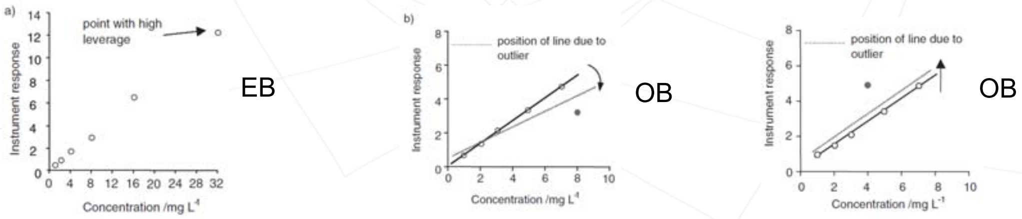
- Z praktického hlediska (využití software) budeme regresní diagnostiku dělit na 2 části:
  - metody analýzy kvality dat (vlivných bodů),
  - metody pro splnění předpokladů pro MNČ a posuzování vhodnosti modelu.

3

Univerzita Palackého
v Olomouci

Kvalita dat: vlivné body

- Vlivné body ovlivňují výsledek statistické analýzy zkreslením regresního modelu.
- Lze je rozdělit do 3 skupin:
  - hrubé chyby – důsledek chyb při manipulaci s daty,
  - body s vysokým vlivem – spolehlivě změřené body rozšiřující predikční schopnost,
  - zdánlivě vlivné body – jeví se jako vlivné, protože byl zvolen nevhodný regresní model.
- Podle místa výskytu se vlivné body dělí na:
  - **odlehlé body (OB)** – liší se v hodnotách závisle proměnné,
  - **extrémní body (EB)** – liší se v hodnotách nezávisle proměnné,
  - kombinace OB a EB, o jejich výsledném vlivu spíše rozhoduje to, že jsou EB.



4

Univerzita Palackého
v Olomouci

Indikace vlivných bodů: analýza vlivu pomocí indexů

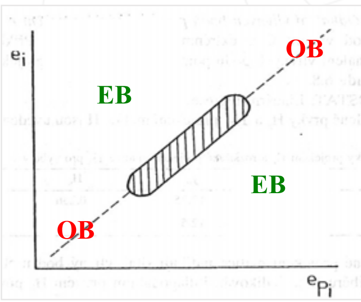
- Diagnostiky vlivných bodů pomocí indexů jsou založeny na sledování změn, ke kterým dojde při vypuštění jednoho bodu a „dopočtení“ jeho hodnot z regresního modelu. Vypuštění bodu a dopočtení se postupně provádí pro všechny body v modelu a následně se vyhodnotí jednotlivé charakteristiky.
- **Cookova vzdálenost D_i**: je-li $D_i > 1$, bod je vlivný.
- **Atkinsonova vzdálenost**: modifikace Cookovy vzdálenosti se zvýrazněnou citlivostí na EB.
- **Diagonální prvky projekční matice H_{ii}**:
  - indikují přítomnost EB, které nezachytí analýza reziduí,
  - $\mathbf{H} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$
- V QC.Expertu se používá barevné zvýraznění bodů identifikovaných jako vlivné.
- Stejnou informaci poskytnou i grafy těchto diagnostik:
  - graf Cookovy vzdálenosti.
  - graf Atkinsonovy vzdálenosti.
  - graf diagonálních prvků projekční matice **H**.

5

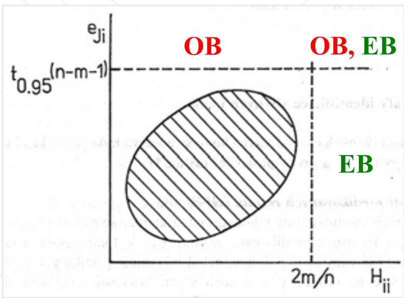
Univerzita Palackého
v Olomouci

Identifikace vlivných bodů: grafy

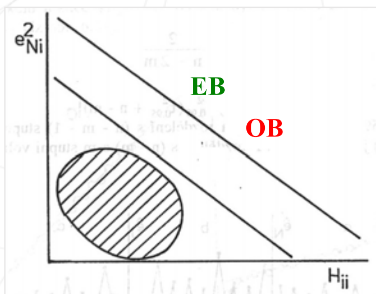
Graf predikovaných reziduí



Williamsův graf



Pregibonův graf

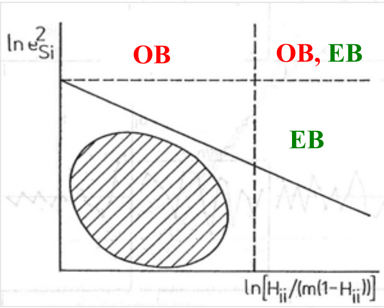


6

Univerzita Palackého v Olomouci

Identifikace vlivných bodů: grafy

McCullohův-Meeterův graf



- **L-R graf** (osa x: H_{ii}, osa y: e_{Ni}^2)
  - Hyperboly znázorňují isolinie stejného vlivu.
  - Podle polohy vůči 3 křivkám lze data rozdělit na slabě vlivná, vlivná a silně vlivná.
- **Q-Q graf** (osa x: kvantil N(0, 1), osa y: reziduum)
  - Lze konstruovat pro různá rezidua.
  - Kromě vlivných bodů slouží i k posouzení normality reziduí.

7

Univerzita Palackého v Olomouci

Analýza reziduí

- Reziduum je vyčíslená hodnota z regresního modelu a používá se při posuzování kvality modelu i kvality dat. Reziduum není to samé co chyba a dá se tedy využít ke statistickému hodnocení v lineární regresi.
- Druhy reziduí a jejich použití:
  1. Klasické reziduum $e_i = y_i - y_{i,reg}$ – využívají se pro grafické hodnocení – viz další slide
  2. Normované reziduum $e_{Ni} = e_i/\sigma$ – hodnoty větší než ±3σ indikují OB
  3. Standardizované reziduum (e_{Si}) – slouží k identifikaci heteroskedasticity $e_{Si} = \frac{e_i}{\sigma \cdot \sqrt{1-H_{ii}}}$
  4. Jackknife reziduum (e_{Ji}) – identifikuje OB
  5. Predikované reziduum (e_{Pi}) – identifikuje OB $e_{Pi} = \frac{e_i}{1-H_{ii}}$
  6. Rekurzivní reziduum (e_{Ri}) – identifikuje autokorelaci

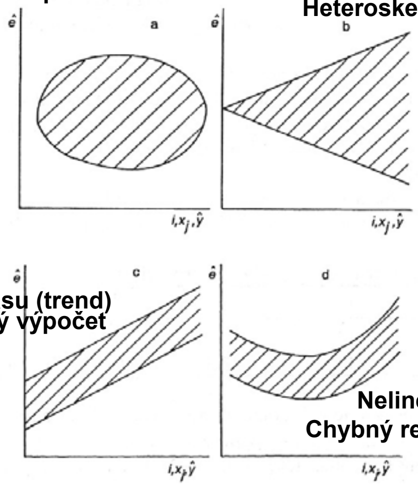
8

Univerzita Palackého v Olomouci

Analýza reziduí – grafy

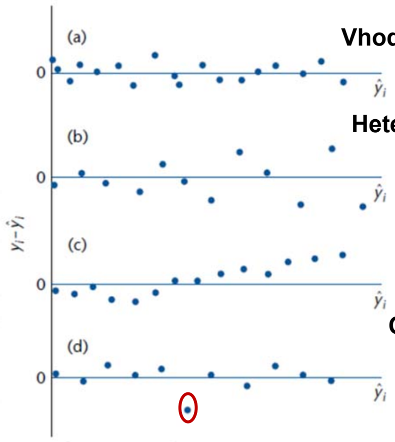
Tvar „mraku“
Vhodné použití MNČ

Tvar výseče
Heteroskedasticita



Tvar pásu (trend)
Chybný výpočet

Nelineární tvar
Chybný regresní model



Vhodné použití MNČ

Heteroskedasticita

Trend

Odlehlý bod

9

Univerzita Palackého v Olomouci

Ověření předpokladů MNČ – testování regresního tripletu

- Statistická významnost regresního modelu: **F_R test významnosti regrese** – testuje, zda použitý model je lepší než „žádný“ model.
  - *Viz prezentace CHEX1-06-LR-I*
- Multikolinearita: **Scottovo kritérium multikolinearity SC**
  - *Viz tato prezentace*
- Závislost/trend reziduí: neparametrický test ověřuje přítomnost závislostí, které nejsou postihnuty modelem – posouzení na základě počtu změn +/- reziduí.
- Normalita reziduí: **Jarque-Bearův test**; JB se srovnává s $\chi_{krit}^{2}{}_{(2)}$.
  - Test je založen na posouzení statistického rozdělení reziduí.

$$\boxed{\mathrm{JB} = n.\left(\frac{g_1}{6} + \frac{(g_2 - 3)^2}{24}\right)}$$

  - Je-li JB < χ_{krit}^2 – je prokázána normalita.
  - Normalitu reziduí lze odhalit i v Q-Q grafech reziduí.

10

Univerzita Palackého
v Olomouci

Ověření předpokladů MNČ – testování regresního tripletu

- Heteroskedasticita tj. nekonstantnost rozptylu: **Cook-Weisbergův test**; CW se srovnává s $\chi_{krit}^2(1)$.
  - Je-li CW > χ_{krit}^2 – je prokázána heteroskedasticita.

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

  - Heteroskedasticitu lze odhalit i v grafu heteroskedasticity (osa x: $(1-H_{ii})y_i$, osa y: e_{Si}^2) $\Rightarrow$ klínový tvar bodů v grafu.
- V přítomnosti heteroskedasticity je třeba uvažovat o použití metody vážených nejmenších čtverců.

- Autokorelace – v LR bývá důsledkem vynechání významné proměnné související s y: **Waldův test test**; WA se srovnává s $\chi_{krit}^2(1)$.
  - Je-li WA > χ_{krit}^2 – je prokázána autokorelace.
  - Testuje přítomnost autokorelace chyb na základě reziduí.
- Často se používá i **Durbin-Watsonův test**, který také ověřuje přítomnost autokorelace na základě reziduí.
  - $0 \leq DW < 2$ a $2 < DW < 4$ potvrzuje autokorelaci.
  - $DW \approx 2$ autokorelace není.

$$WA = \frac{n\rho_1^2}{1-\rho_1^2}$$

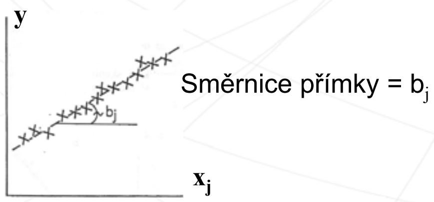
$$DW = \frac{\sum_{i=2}^{n}(e_i-e_{i-1})^2}{\sum_{i=1}^{n}e_i^2}$$

11

Univerzita Palackého
v Olomouci

Vhodnost regresního modelu

- Pomocí rozptylového regresního grafu: y = f(x).
  - Posuzujeme vhodnost a těsnost proložení bodů regresním modelem, šířku pásu spolehlivosti, přítomnost vlivných bodů.
- Pomocí parciálních regresních grafů (u vícerozměrných lineárních regresních modelů).
  - Závislost *y* na zvolené x_j s eliminací vlivu ostatních nezávisle proměnných *x*. Závislost je lineární pouze v případě, že příslušný koeficient b_j je statisticky významný.



- Pomocí charakteristik vhodnosti modelu AIC, MEP, R_p.
  - Při porovnávání regresních modelů hledáme MEP a AIC minimální a R_p maximální.

12

Univerzita Palackého v Olomouci

Vhodnost regresního modelu

- **Střední kvadratická chyba predikce** – MEP (*Mean Error of Prediction*)
  - MEP využívá princip „indexů“: predikce $y_{reg,i}$ z odhadu, při jehož konstrukci byla informace o i-tém bodu vypuštěna. Jde tedy o chybu i-tého bodu závisle proměnné spočítanou regresí právě s vypuštěním i-tého bodu.

$$MEP = \frac{1}{n}\sum_{i=1}^{n}\frac{e_i^2}{(1-H_{ii})^2}$$

- **Predikovaný koeficient determinace** R_p^2 – získáme pokud při výpočtu R^2 použijeme MEP místo RSC, je citlivější na vybočující body než klasický R^2.
  - QC.Expert používá predikovaný korelační koeficient R_p.

$$R_p^2 = 1 - \frac{n\,MEP}{\sum_{i=1}^{n} y_i^2 - n\,\bar{y}}$$

- **Akaikovo informační kriterium** AIC – je kritérium kvality regrese vycházející z RSC „penalizovaného“ počtem proměnných n.

$$AIC = n \cdot \ln\left(\frac{RSC}{n}\right) + 2m$$

13

Univerzita Palackého v Olomouci

Postup výstavby jednoduchého regresního modelu

1. Návrh předběžného regresního modelu.
2. Předběžná analýza dat (posouzení R, AIC, MEP, R_p).
3. Regresní diagnostika zaměřená zejména na kvalitu dat (OB, EB).
4. Konstrukce zpřesněného regresního modelu po případném vyloučení OB.
5. Posouzení kvality modelu s využitím testů regresního tripletu. Případné použití jiné metody odhady než je MNČ.
6. Tvorba konečného regresního modelu.

14

Univerzita Palackého v Olomouci

Polynomické regresní modely

15

Univerzita Palackého v Olomouci

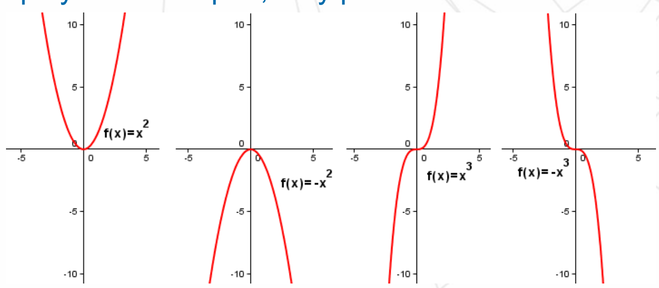
Polynomické modely, úvod

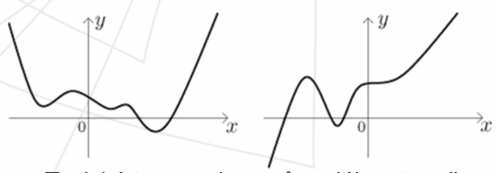
- Jedná se o regresní model, který je lineární v parametrech, ale popisuje nelineární závislost mezi proměnnými:

$$y = b_0 + b_1x + b_2x^2 + \dots + b_nx^n$$

b_i jsou regresní koeficienty, n je stupeň polynomu

- Tento model obsahuje pouze jednu nezávisle proměnnou x, která se v něm však vyskytuje v různých mocninách. Jsou zde vždy všechny mocniny od 1 do n. Speciálním případem je polynom 1. stupně, tedy přímka.





Typické tvary polynomů sudého stupně (nalevo) a lichého stupně (napravo).

16

Univerzita Palackého v Olomouci

Multikolinearita

- **MULTIKOLINEARITA** (porušení předpokladu pro MNČ) – vysoké hodnoty párových korelačních koeficientů mezi vysvětlujícími proměnnými, přibližná rovnoběžnost sloupcových vektorů v matici **X**.
- Multikolinearita způsobuje:
  - Početní problémy během MNČ: špatná podmíněnost matice $\mathbf{X}^T\mathbf{X}$ kvůli hodnotám vlastních čísel λ blízkých nule; nelze při MNČ provést inverzi matice ⇒ regresní model není jednoznačně řešitelný.
  - Statistické problémy: neúměrně vysoké rozptyly regresních koeficientů vedoucí k jejich nespolehlivému určení; nestabilita odhadů regresních koeficientů.
- Příčiny multikolinearity:
  - závislost mezi nezávisle proměnnými v LRM,
  - polynomická povaha regresního modelu,
  - přeurčenost regresního modelu – příliš mnoho nezávisle proměnných (u vícerozměrné LR).

17

Univerzita Palackého v Olomouci

Multikolinearita

- **Identifikace multikolinearity:**
  - **Číslo podmíněnosti κ (kappa):** $\kappa = \frac{\lambda_{max}}{\lambda_{min}}$

    λ_{max} a λ_{min} jsou maximální a minimální vlastní čísla. Hodnota $\kappa > 1000$ silná multikolinearita
  - **VI-faktor** (Variance Inflation Factor), VIF > 10, jde o silnou multikolinearitu:

$$VIF = \frac{1}{1 - R_j^2}$$

  - **Scottova testační charakteristika M_T:**

$$M_T = \frac{\frac{F_R}{t_S} - 1}{\frac{F_R}{t_S} + 1}$$

F_R je testační kritérium ze čtverců testačních statistik t_i (test významnosti regresního koeficientu) a t_S je průměrná hodnota čtverců testačních charakteristik t_i

$M_T > 0{,}8$ – model z hlediska multikolinearity nevyhovuje, je nezbytná úprava modelu
$0{,}33 < M_T < 0{,}8$ – model málo vyhovující, úprava však není nezbytná,
$M_T < 0{,}33$ – model není multikolinearitou ovlivněn.

18

Univerzita Palackého v Olomouci

Metoda korekce hodnosti

- Metoda odhadu regresních koeficientů používaná místo MNČ v případě polynomických regresních modelů a vyšších stupňů polynomů. Regresní koeficienty odhadnuté pomocí MNČ mají velmi vysoké rozptyly a proto neumožňují sestrojení regresního modelu.
- Postup výpočtu regresních koeficientů u metody korekce hodnosti je stejný jako u MNČ, tj. je hledán minimální RSC. Do výpočtu se vkládá omezení na velikost vlastních čísel λ.
- Omezení na velikost λ:
  - P = 0 – MNČ.
  - Hodnoty P > 0 potlačí složky odpovídající nejmenším vlastním číslům, doporučuje se P ≈ 0,1.
  - Maximální hodnota P = 1.
- Po vynechání určitého počtu malých vlastních čísel λ se vypočtou nové odhady koeficientů b_i s menším rozptylem, které jsou vychýlené.
- Tyto vychýlené odhady vyhovují lépe než nevychýlené odhady získané pomocí MNČ a lépe se interpretují v regresním modelu. Metoda korekce hodnosti tedy vede k vychýleným odhadům b_i, které jsou méně citlivé na špatnou podmíněnost matice $\mathbf{X}^T\mathbf{X}$.

19

Univerzita Palackého v Olomouci

Postup výstavby polynomického regresního modelu

1. Určení stupně polynomu (posouzení R, AIC, MEP, R_p).
2. Redukce nadbytečných členů polynomu:
   - při použití MNČ posouzení významnosti regresních koeficientů,
   - při použití metody korekce hodnosti nalezení vhodné velikosti omezení P s ohledem na významnost regresních koeficientů.
3. Návrh předběžného regresního modelu.
4. Regresní diagnostika zaměřená zejména na kvalitu dat (OB, EB).
5. Konstrukce zpřesněného regresního modelu po vyloučení OB.
6. Posouzení kvality modelu s využitím testů regresního tripletu. Případné použití jiné metody odhady než je MNČ.
7. Tvorba konečného regresního modelu.

20

Univerzita Palackého v Olomouci

Vícerozměrné lineární regresní modely

21

Univerzita Palackého v Olomouci

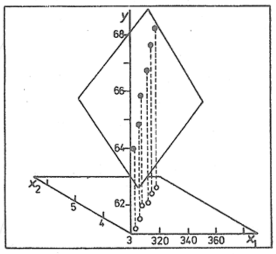
Vícerozměrné LRM

- Pokud měřené hodnoty *y* závisejí na více faktorech *x*, není vždy možné provést takovou sérii pokusů, při nichž by se měnila pouze hodnota jediného faktoru *x* a ostatní by zůstaly konstantní. Experimentální výsledky lze zpracovávat ve vztahu k současné změně více faktorů pomocí vícerozměrné lineární regrese.
- Jsou-li hodnoty *y* lineárně závislé na několika nezávislých proměnných *x*, tj. platí-li

$$y = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_k x_k$$

můžeme pro výpočet odhadů koeficientů b_i o počtu *k* použít MNČ (při splnění předpokladů pro použití).

- Úsek (absolutní člen) b_0 je průsečíkem regresní nadroviny s osou y. Koeficienty b_i jsou směrnice regresní nadroviny ve směru x_i a jsou nazývány parciálními regresními koeficienty.



Grafické znázornění vícenásobné LR pro 2 faktory x, řešením je rovina. Pokud je x > 2, řešením je nadrovina.

22

Univerzita Palackého v Olomouci

Postup výstavby vícenásobného regresního modelu

1. Návrh předběžného regresního modelu.
2. Posouzení parciálních regresních grafů a významnosti jednotlivých regresních koeficientů b_i.
   - U vícerozměrné LR nemáme k dispozici rozptylový graf regresní závislosti y na x, ale parciální regresní grafy y na x_1, y na x_2 atd. U nich posuzujeme vhodnost a těsnost proložení dat přímkou.
3. Předběžná analýza dat (posouzení vícenásobného a parciálních korelačních koeficientů, AIC, MEP, R_p).
4. Regresní diagnostika zaměřená zejména na kvalitu dat (OB, EB).
5. Konstrukce zpřesněného regresního modelu po vyloučení OB.
6. Posouzení kvality modelu s využitím testů regresního tripletu. Případné použití jiné metody odhady než je MNČ.
7. Tvorba konečného regresního modelu.

23

Univerzita Palackého v Olomouci

Kalibrace

24

Univerzita Palackého
v Olomouci

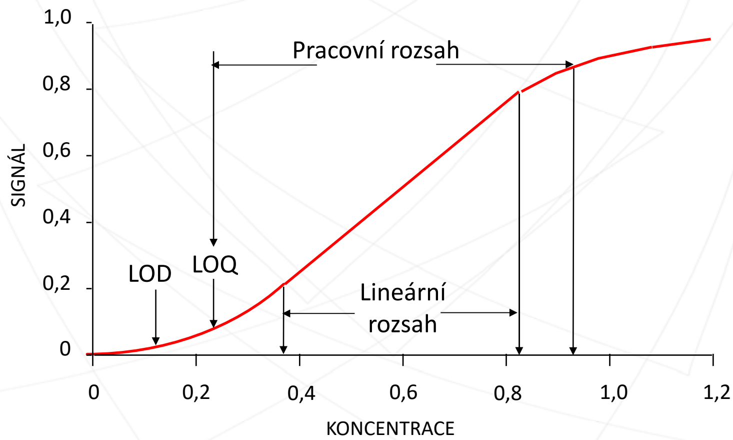
Úvod

- Kalibrace patří k základním úlohám chemické praxe, jež se řeší s využitím regresních metod.
- DEFINICE kalibrace dle VIM 3: činnost, která za specifikovaných podmínek v prvním kroku stanoví vztah mezi h**odnotami veličiny** s **nejistotami měření** poskytnutými **etalony** a odpovídajícími **indikacemi** s přidruženými nejistotami měření a ve druhém kroku použije tyto informace ke stanovení vztahu pro získání **výsledku měření** z indikace.
- Dle definice se kalibrace skládá ze dvou kroků:
  1. Sestavení kalibračního modelu – tj. sestrojení regresního modelu z výsledků analýz kalibračních standardů (etalonů).
  2. Použití kalibračního modelu: pro signál vzorku y^* se hledá odpovídající hodnota x^* (obvykle koncentrace) a interval spolehlivosti (nejistota).
- Ze statistického hlediska rozlišujeme lineární kalibraci (přímka) a nelineární kalibraci.

25

Univerzita Palackého
v Olomouci

Kalibrační závislost



Převzato z materiálů LGC Standards Limited, Velká Británie.

26

Univerzita Palackého v Olomouci

Lineární kalibrace

- Kalibrační přímka je nejpoužívanější kalibrační model. Předpokládá se, že linearita modelu platí v celém rozsahu kalibrace, což je nezbytné ověřit.

Sestavení modelu: $y = b_0 + b_1 x + \varepsilon$
Použití modelu: $y_s^* = b_0 + b_1 x_s^* + \varepsilon$

- Zpětný odhad x_s^* lze určit několika způsoby. QC Expert používá přímý odhad a nepřímý odhad. Zpětné odhady jsou hodnoty x_s^* neznámého vzorku vypočítané z naměřené odezvy y_s^* pomocí zvoleného kalibračního modelu.

- **Přímý odhad** (kombinace rovnic pro sestavení a použití modelu):

$$\boldsymbol{x_s^* = x + \frac{y_s^* - y}{b_1}}$$

- **Nepřímý** (modifikovaný) **odhad**, který koriguje vychýlení:

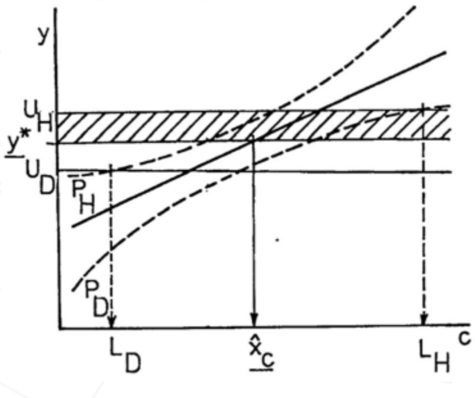
$$x_s^* = \bar{x} + \frac{(y_s^* - \bar{y})b_1}{b_1^2 + \left(\sigma^2 / \sum_i (x_i - \bar{x})^2\right)}$$

27

Univerzita Palackého v Olomouci

Interval spolehlivosti x_s^*

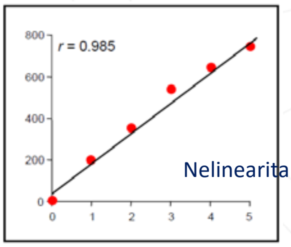
- Odhad x_s^* je odhad bodový, potřebuje znát jeho IS, nejčastěji 95%. Lze jej určit výpočtem nebo graficky.
- Grafické určení IS:
  - L_D je řešením $U_D = P_H$,
  - L_H řešením $U_H = P_D$
  - U_H, U_D IS opakovaných měření y^*
  - P_H, P_D meze pásu spolehlivosti
- IS x_s^* obecně <u>není symetrický</u>!



28

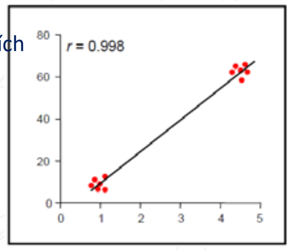
Univerzita Palackého v Olomouci

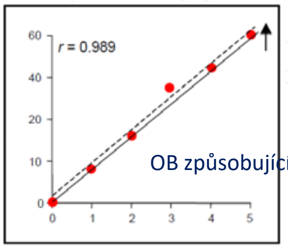
Lineární kalibrace – interpretace kalibračních grafů



Nelinearita

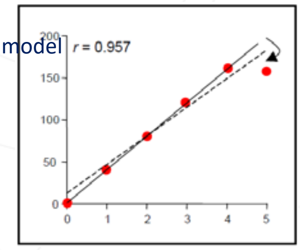
Nevhodná volba kalibračních standardů





OB způsobující odchylku

OB ovlivňující regresní model

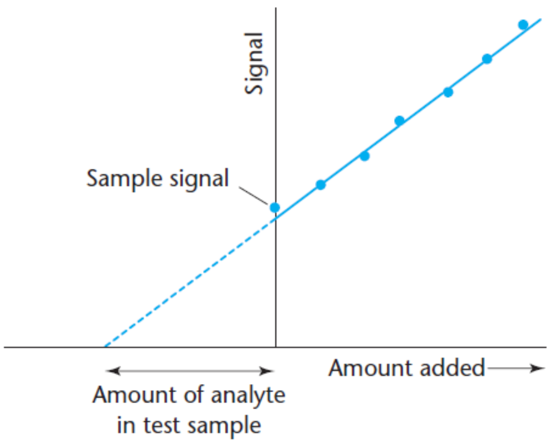


29

Univerzita Palackého v Olomouci

Metoda přídavků standardu

- *Method of standard additions* – *metoda přídavků standardu NE metoda standardního přídavku*.
- Používá se k eliminaci matričních vlivů, je časově náročná a vyžaduje větší množství vzorku.
- Koncentrace přidávaného analytu musí být v lineární části kalibrační závislosti.
- Ze statistického hlediska je vhodné analyzovat vzorek + 3 přídavky ⇒ 4 body do regresního modelu.
- Důležité je vhodná volba koncentrace přídavků, aby signál rostl rovnoměrně.

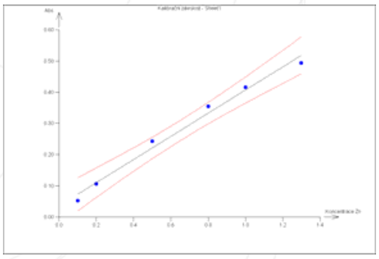


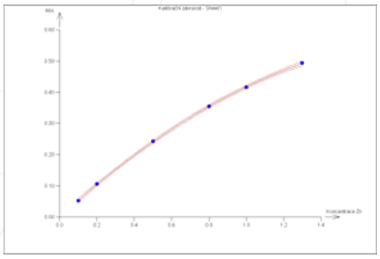
30

Univerzita Palackého
v Olomouci

Nelineární kalibrace

- Pokud přímka jako kalibrační model nevyhovuje, používáme nelineární kalibraci.
- Vhodný kalibrační model určujeme podle meze detekce. Vhodný model bude mít nejnižší mez detekce.
- Nelineární modely:
  - **Lineární spline** – proložení dat několika odlišnými přímkami spojenými uzly.
  - **Polynomický model** – obvykle dostačuje kvadratický model.
  - **Nelineární spline** – proložení dat několika odlišnými polynomickými modely spojenými uzly.





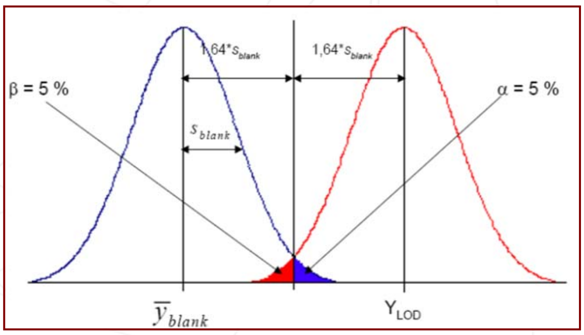
31

Univerzita Palackého
v Olomouci

Mez detekce (Limit of Detection)

- Definice z VIM 3: **naměřená hodnota veličiny** získaná daným **postupem měření**, pro kterou je pravděpodobnost nepravdivého tvrzení o nepřítomnosti složky v materiálu *β*, přičemž pravděpodobnost nepravdivého tvrzení o její přítomnosti je *α*.
  - POZNÁMKA 1 IUPAC doporučuje implicitní hodnoty pro *α* a *β* rovné 5 %.
  - POZNÁMKA 2 Někdy se používá zkratka LOD.
  - POZNÁMKA 3 Termín „citlivost“ se nedoporučuje používat pro ‚mez detekce‘.

$$y_{LOD} = \bar{y}_{blank} + (1{,}64 s_{blank} + 1{,}64 s_{blank})$$



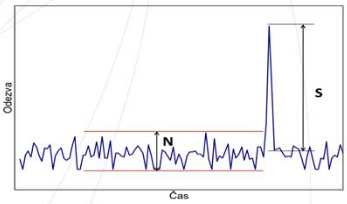
Předpoklad: normální rozdělení
a shoda rozptylů.

32

Univerzita Palackého v Olomouci

Mez detekce – přístupy k určení

- Hodnocení variability měření slepého vzorku (pokusu) zavedl v 60. letech 20. stol. Kaiser.
  - Je třeba vhodně zvolit slepý vzorek (pokus).
  - Způsob výpočtu *s*: za podmínek opakovatelnosti či lépe za podmínek mezilehlé preciznosti .
- Vyhodnocení z kalibrační křivky (např. v QC.Expert)
  - Základem jsou normy ISO 11483-2 a DIN 32645.
  - V praxi je třeba vhodně zvolit koncentrační rozsah a složení standardů pro určení LOD. Doporučuje se max. 1 řád nad očekávanou hodnotou LOD.
- Mez detekce ve vztahu k poměru signálu k šumu
  - U instrumentálních technik, u kterých je kontinuálně registrována nulová linie jako např. u chromatografie, může být LOD odvozena od koncentrace analytu v dávkovaném vzorku, který v použitém detekčním systému vykazuje zvolený poměr signálu k šumu. Používá se hodnota poměru 2 až 5, nejčastěji 3.



Vyhodnocení poměru signálu k šumu S/N u chromatografické analýzy

33

Univerzita Palackého v Olomouci

Mez detekce: Kaiserův přístup

- Použití (nezávislé) korekce pozadí/baseline

$$y_{LOD} = K \cdot s_{blank}$$

- Není použita korekce pozadí/baseline

$$y_{LOD} = \overline{y}_{blank} + K \cdot s_{blank}$$

- Určení K:
  - Kaiser (normální rozdělení): K = 3 (3,28 zaokrouhleno),
  - případně kvantil Studentova *t* rozdělení pro nízká *n*.
- **Mez stanovitelnosti, K = 10:**
  - Pro potřeby kvantitativní analýzy definujeme mez stanovitelnosti **LOQ** jako parametr určující obvykle počátek pracovního rozsahu metody.
  - Není definována statisticky, ale konvenčně jako hodnota obsahu složky (koncentrace analytu), při které je nejistota stanovení vyjádřená jako relativní směrodatná odchylka rovna předem určené hodnotě (doporučení IUPAC 10 %).

$$y_{LOQ} = 10 \cdot s_{blank} \text{ nebo } y_{LOQ} = \overline{y}_{blank} + 10 \cdot s_{blank}$$

34

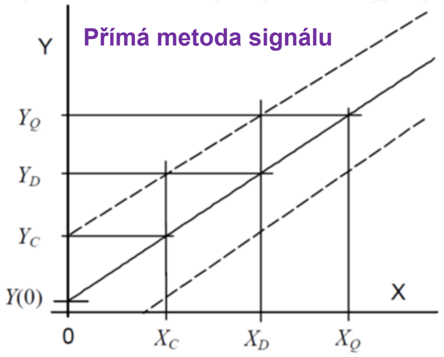
Univerzita Palackého
v Olomouci

Mez detekce – vyhodnocení z kalibrační křivky

- Využívá statistických vlastností (pásu spolehlivosti, směrnice) kalibračního modelu.

$$X_{LOD} = s_{x0} \cdot t \cdot \sqrt{\frac{1}{m} + \frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}}$$

$$X_{LOQ} = k.s_{x0} \cdot t \cdot \sqrt{\frac{1}{m} + \frac{1}{n} + \frac{(x_Q - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}}$$



- Y_C … kritická úroveň na ose y
- Y_{LOD} … mez detekce na ose y
- Y_{LOQ} … mez stanovitelnosti na ose y
- X_C … kritická úroveň na ose x
- X_{LOD} … mez detekce na ose x
- X_{LOQ} … mez stanovitelnosti na ose x

- Hodnoty nižší než kritická úroveň se považují za šum.

35

Univerzita Palackého
v Olomouci

Prezentování výsledků okolo meze detekce

Hodnota signálu *y*	Statistický přístup prezentování výsledku	„Laboratorní přístup“ prezentování výsledku
> LOQ	Stanoveno, konkrétní hodnota výsledku	Stanoveno, konkrétní hodnota výsledku
> LOD, < LOQ	Detekováno, nestanoveno, příp. hodnota LOQ	Nestanoveno, příp. hodnota LOQ
< LOD	Nedetekováno, příp. hodnota LOD	Nedetekováno, příp. hodnota LOD

36