

Nejistoty měření v metrologii

Jiří Militký¹, Vladimír Bajzík¹, Milan Meloun²

¹ Katedra textilních materiálů, Textilní fakulta, Technická universita v Liberci, Liberec

² Katedra analytické chemie, Universita Pardubice, Pardubice

Motto: „*The only relevant thing is uncertainty - the extent of our knowledge and ignorance. The actual fact of whether or not the events considered are in some sense determined, or known by other people, and so on, is of no consequence*“

[Bruno deFinetti]

1. Úvod

Je známo, že měření a interpretace výsledků měření je základem jak přírodních tak i technických věd. Žádné měření není úplně perfektní, protože probíhá na přístrojích s omezenou přesností konstruovaných podle přibližných měřicích principů a v průběhu měření se vyskytuje řada nekonstantních podmínek. V řadě případů je integrální součástí měřicího řetězce také člověk jako zdroj subjektivity resp. nepřesnosti. V praxi jsou tedy měření zatížena celou řadou různých šumů označovaných obvykle jako chyby resp. systematických vychýlení (bias). Tyto šумы pak způsobují rozptýlení měřených hodnot a jsou zdrojem nepřesnosti výsledků. Způsob kombinace jednotlivých chyb je specifikován modelem jejich působení.

Účelem měření je v nejjednodušším případě stanovení jedné (měřené) veličiny. Výsledky měření jsou pak vyjádřeny pomocí vhodného odhadu skutečné (neznámé) hodnoty a odpovídající míry nejistoty, které souvisejí s modelem působení chyb resp. vychýlením.

Klasická statistika (vycházející z definice pravděpodobnosti jako limity relativní četnosti) poskytuje aparát pro vyjádření nejistoty jako intervalu spolehlivosti parametru μ .

Vyjádření nejistot publikované v příručkách [1,2] je filosoficky blíže subjektivní definici **pravděpodobnosti jako stupni důvěry (víry)**. Tato pravděpodobnost pak souvisí spíše s nedostatkem znalostí než s výsledkem opakovaného experimentu.

Nejistoty ←→ **Intervaly spolehlivosti**
ISO, EURACHEM ←→ **Statistika klasická**

V této práci je porovnání přístupu presentovaného v příručce EURACHEM [1] resp. příručce NIST [2] s klasickým statistickým přístupem vycházejícím z metody maximální věrohodnosti [5,6] a Bayesovským přístupem vycházejícím ze subjektivní pravděpodobnosti [3]. Jsou ukázány způsoby vyjádření výsledků měření pro komplikovanější praktické situace. Je navržen způsob výpočtu nejistot nepřímých měření založený na použití simulace typu Bootstrap.

2. Nejistoty výsledků měření

Nejistota je celkem známý statistický pojem související s odhadováním parametrů: Omezme se na základní model aditivních šumů. Pro tento model je výsledek měření ve tvaru

$$x = \mu + \varepsilon \quad (1)$$

Parametr (střední hodnota) μ je hodnota odhadovaná na základě **výsledků měření** x . O náhodné **chybě** ε se předpokládá, že má nulovou střední hodnotu ($E(\varepsilon)=0$) a konstantní rozptyl $D(\varepsilon)=\sigma^2$. **Nejistota jednotlivého měření** se pak vyjadřuje intervalem: $x \pm u$, kde u je násobek σ , tj. $u = k \sigma$. Při volbě $k=1$ odpovídá tento interval přibližně 65 %-nímu intervalu spolehlivosti a pro $k=2$ odpovídá tento interval přibližně 95 %-nímu intervalu spolehlivosti. V řadě případů je situace složitější a některé zdroje chyb jsou nenáhodné. Této situaci lépe vyhovuje rozšířený model

$$x = \mu + b + \varepsilon \quad (2)$$

Systematické vychýlení (bias) b se často pouze odhaduje na základě expertních odhadů vhodným intervalem $a - d \leq b \leq a + d$. Expertní odhad parametrů a, d pak umožňuje konstrukci konzervativních nejistot měření

$$(x-a) \pm (1.96\sigma + d) \quad (3)$$

Tento tzv. *ortodoxní přístup* byl svého času doporučen jak americkou NIST tak i anglickou NPL. Základní nevýhodou je, že náhodná a systematická složka se zpracovávají zvlášť **Nejistota výsledku**, tj. střední hodnoty μ odhadované jako $\hat{\mu}$ je vyjádřena intervalem spolehlivosti střední hodnoty, pro který platí

$$P(a^- \leq \mu \leq a^+) = 1 - \alpha \quad (4)$$

BIMP (International Bureau of weights and measures) 1980 doporučil pět pravidel pro vyjadřování nejistoty a na jejich základě pak ISO doporučilo postup založený obecně na nepřímých měřeních, kdy platí

$$\mu = f(\bar{x}, \bar{z}) \quad (5)$$

kde $\bar{x} = (x_1, \dots, x_p)$ jsou měřené veličiny (odpovídající nejistoty u_A **typu A** jsou určeny standardními statistickými postupy z odhadů rozptylů $s_{x_i}^2$) a $\bar{z} = (z_1, \dots, z_R)$ jsou neměřené resp. neměřitelné veličiny. Odpovídající nejistoty u_B **typu B** se určují z expertních odhadů rozptylů $s_{z_i}^2$

Odhad $D(\hat{\mu})$ se počítá z Taylorova rozvoje do lineárních členů

$$D(\hat{\mu}) = \sum_i c_i^2 s_{x_i}^2 + \sum_j e_j^2 s_{z_j}^2 \quad (6)$$

$$\text{kde } c_i = \frac{\partial f(\bar{x}, \bar{z})}{\partial x_i} \quad \text{a} \quad e_j = \frac{\partial f(\bar{x}, \bar{z})}{\partial z_j}$$

Pro určení $(1-\alpha)\%$ -ních intervalů spolehlivosti se pak využívá tzv. *efektivních stupňů volnosti*

$$v_{\text{eff}} = \frac{D(\hat{\mu})^2}{\sum_i c_i^4 s_{x_i}^4 / v_i} \quad (7)$$

Nejistotu výsledku je pak možno využít vztah

$$\hat{\mu} \pm t(v_{\text{eff}})_{1-\alpha/2} * \sqrt{D(\hat{\mu})} \quad (8)$$

Popsaný postup lze použít pro vyjádření nejistoty rozšířeného modelu viz rov. (2). Zřejmě zde platí, že $\mu = x_1 - z_1$. Pokud je vychýlení z_1 z intervalu $a - d \leq z_1 \leq a + d$ vyjde

$$\text{Nejistota měření ISO} \quad (x_1 - a) \pm 1.96 * \sqrt{(\sigma_{x_1}^2 + d^2 / 3)}$$

Po dosazení do rov(3) vyjde konzervativní odhad

$$\text{Nejistota měření ortodoxní} \quad (x_1 - a) \pm (1.96 * \sigma_{x_1} + d)$$

Je patrné, že postup podle ISO zpracovává systematickou složku stejně jako náhodnou. Je užitečné porovnat postupný vývoj vyjadřování celkové nejistoty U jako kombinace nejistot typu A (statistická) a typu B (nestatistická). V roce 1969 byl navržen (US Air Force) vztah

$$U = \pm [u_B + s_{\bar{x}} * t_{0.95}] \quad (9)$$

kde druhý člen odpovídá 95 % nímu intervalu spolehlivosti střední hodnoty určené z nejistot typu A. Mísí se tedy míry rozptýlení a intervalové odhady. Navíc není celková nejistota intervalem. V roce 1985 byl (ASME) navržen výraz

$$U = \pm \sqrt{[u_B^2 + (s_{\bar{x}} * t_{0.95})^2]} \quad (10)$$

Také zde se mísí bodové odhady a intervalové odhady. Celková nejistota je však již intervalem. Konečně v roce 1993 navrhla ISO unifikovaný postup vedoucí k intervalu

$$U = \pm k \sqrt{[u_B^2 + u_A^2]} \quad (11)$$

Tato definice umožňuje využití neexperimentální informace. Na druhé straně je problémem, jak stanovit zdroje nejistot a jejich variabilitu.

3. Zpracování nepřímých měření

Výsledek analýzy lze v tomto případě vyjádřit jako

$$y = f(\mu_1, \dots, \mu_M) \quad (12)$$

Zde $f(\mu_1, \dots, \mu_M)$ je známá funkce skutečných hodnot výsledků přímých měření μ_1 až μ_M (např. měříme poloměr a chceme znát plochu příčného řezu kruhových vláken). K dispozici jsou odhady parametrů $(\hat{\mu}_1, \hat{\mu}_2, \dots, \hat{\mu}_M)$ a příslušné odhady rozptýlů resp. čtverců nejistot $D(\hat{\mu}_1), D(\hat{\mu}_2), \dots, D(\hat{\mu}_M)$.

Standardní statistická analýza [5,6].

- a) Odhad y z odhadů $\hat{\mu}_i \quad i=1, \dots, M$
- b) Odhad rozptylu $D(\hat{y})$
- c) Odhad intervalu spolehlivosti pro y

Analýza nejistot podle EURACHEM (viz [1,2].)

a) Odhad y z odhadů $\hat{\mu}_i \quad i=1, \dots, M$: Neřeší se přímo, ale zřejmě se příliš aproximativně předpokládá $\hat{y} = f(\hat{\mu}_1, \hat{\mu}_2, \dots, \hat{\mu}_M)$.

b) Odhad rozptylu $D(\hat{y})$: Je vlastně rozšířená nejistota $u(y)$. Vychází se z předpokladu, že $f(\underline{x})$ lze nahradit linearizací Taylorovým rozvojem v okolí $\underline{\mu}$.

$$y = f(\underline{x}) \approx f(\underline{\mu}) + \sum_{i=1}^M \left(\frac{\partial f(\cdot)}{\partial x_i} \right) (x_i - \mu_i) \quad \underbrace{D(y)}_{u^2(y)} \approx \sum_{i=1}^M \left(\frac{\partial f(\cdot)}{\partial x_i} \right)^2 \cdot \underbrace{D(x_i)}_{u^2(x_i)} + \text{cov}(\dots)$$

$D(y)$ se nesprávně označuje jako zákon šíření nejistot. V případě že zdroje nejistot jsou lineárně závislé provádí se korekce s využitím kovariancí $\text{cov}(\cdot)$.

Linearizace může být v řadě případů velmi nepřesná, zejména co se týče intervalů spolehlivosti (rozšířené nejistoty).

Příklad na nepřesnost linearizace [3].

Energie protonu E [GeV] se dá určit z jeho rychlosti c podle vztahu

$$E = \frac{m \cdot c^2}{\sqrt{1 - (v/c)^2}}$$

Pro případ, že rychlost je měřena s relativní přesností 0,2% a $v/c = 0,9971$ je třeba odhadnout 95 %-ní konfidenční interval

Linearizace
 $0.7 \leq E \leq 24$

Korektní řešení
 $7.2 \leq E \leq \infty$

c) Odhad intervalu spolehlivosti pro y : Předpokládá se téměř vždy nekorektně přibližná normalita. (Nelineární funkce normálně rozdělených náhodných veličin již normální rozdělení nemá !!). Polovina 95 %-ního intervalu spolehlivosti, resp. rozšířená nejistota je pak $U = 2 \cdot u(y)$. Zde 2 resp. přesněji 1,98 je kvantil normovaného normálního rozdělení. Pro nelineární transformaci však rezultují nesymetrická rozdělení, což vede k nesymetrickému intervalu spolehlivosti. Ve speciálních případech (např. stopová analýza) to může výrazně ovlivnit závěry (pro pozitivně zešikmená rozdělení vyjde ve směru k nižším hodnotám korektnější interval užší a ve směru k vyšším hodnotám širší).

4. Výhrady k nejistotám

A. Terminologické

<u>EURACHEM</u>	<u>Statistika klasická</u>
Standardní nejistota A	směrodatná odchylka měřené šumové složky
Standardní nejistota B	směrodatná odchylka (odhadnutá) šumové složky
Kombinovaná nejistota	směrodatná odchylka funkce y
Rozšířená nejistota	polovina intervalu spolehlivosti
Faktor pokrytí	kvantil normovaného normálního rozdělení

Terminologické nepřesnosti nejsou na závadu, pokud se naleznou a přesně uvedou rozumné důvody proč je potřeba volit vlastní názvosloví.

B. Statistické

Vychází se z těchto striktních předpokladů bez ověření:

- aditivní model měření resp. působení šumových složek (zdrojů nejistot)
- konstantní rozptyl měření (resp. zdrojů nejistot)
- normalita nelineární funkce normálně rozdělených proměnných (pro určení rozšířené nejistoty resp. intervalu spolehlivosti - IS)
- nekorelovanost měření
- malá nelinearita $f(\underline{x})$ umožňující použití linearizace.

Dále je zde nekorektnost při konstrukci a interpretaci U (resp. IS). Klasická statistika vede k tomu, že pro $n \rightarrow \infty$ je $100(1-\alpha)\%$ -ní interval spolehlivosti parametru μ roven [5,6].

$$\hat{\mu} \pm n_{1-\alpha/2} \cdot \sqrt{D(\hat{\mu})}.$$

Při výpočtu pomocí nejistot není vlastně kombinovaná nejistota u_c^2 pouze odhadem rozptylu $D(\hat{\mu})$, ale obsahuje další složky. Pak tedy vyjde rozšířená nejistota **systematicky vyšší** než polovina intervalu spolehlivosti, hodnota 2 nezajišťuje přibližně 95%-ní pokrytí a interpretace takového intervalu je nesnadná.

C. Výpočetní

Místo náhrady derivací diferencemi, jak se doporučuje v příručkách, by bylo podstatně jednodušší užít simulace nebo tzv. Bootstrap odhadů (zejména tam, kde se používá pro výpočty počítač).

5. Bayesovský přístup k nejistotám

Vychází se z Bayesovy definice podmíněné pravděpodobnosti a využívá se a' priori informací (označené indexem o). Věrohodnostní funkce je obecně sdružená hustota pravděpodobnosti $f(\underline{x} / \mu, \underline{h})$, kde $\underline{x} = (x_1, \dots, x_n)$ jsou měření a $\underline{h} = (h_1, \dots, h_n)$ jsou hodnoty ovlivňujících proměnných (zdrojů nejistot - t.j. externích nesledovaných parametrů, systematických odchylek, kalibračních konstant atd.). A' priori informace o μ jsou vyjádřeny hustotou pravděpodobnosti $f_o(\mu)$.

Podle Bayesovy formule pak lze a' posteriorní hustotu pro μ t.j. $f(\mu / \underline{x})$ vyjádřit jako

$$f(\mu / \underline{x}) = \frac{f(\underline{x} / \mu, \underline{h}_o) \cdot f_o(\mu)}{\int f(\underline{x}, \underline{h}_o) \cdot f_o(\mu) d\mu} \quad (13)$$

Tato rovnice vychází z předpokladu, že všechny ovlivňující veličiny nabývají hodnot h_o .

Při znalosti a' posteriorní hustoty pravděpodobnosti $f(\mu / \underline{x})$ (které je kombinací a' priori pravděpodobnosti $f_o(\mu)$ a informací skrytých v experimentu) se snadno z definice určí odhady střední hodnoty

$$\hat{\mu} = \int \mu \cdot f(\mu / \underline{x}) d\mu \quad (14)$$

a rozptylu

$$D(\hat{\mu}) = \int (\mu - \hat{\mu})^2 \cdot f(\mu / \underline{x}) d\mu \quad (15)$$

Pro velké rozsahy experimentu je a' priori informace málo důležitá a $f(\mu / \underline{x})$ je pak úměrné přímo věrohodnostní funkci. Maximálně věrohodné odhady pak odpovídají odhadům Bayesovským.

Bayesovy formule lze použít také v případech, kdy se vyskytují systematické odchylky. (šumové složky nemají nulovou střední hodnotu). Předpokládejme, že veličina h , ovlivňující velikost μ má rozdělení $h \sim N(h_0, \sigma_h^2)$ a měření x_i mají rozdělení $x_i \sim N(\mu, \sigma_M^2)$. Pak na základě Bayesova vztahu určíme, že $\hat{\mu} \approx N(\bar{x} - h_0, \sqrt{\sigma_M^2 + \sigma_h^2})$. To znamená, že výsledek je korigován o hodnotu systematické odchylky a celkový rozptyl je součtem rozptylu měření (určitelný z opakovaných měření) a rozptylu ovlivňující veličiny (může být určen někým jiným - konstrukce přístroje, resp. simulačně atp.). Zde σ_M je nejistota typu A a σ_h je nejistota typu B.

6. Modely měření

Podle působení šumové složky rozlišujeme tyto modely:

Aditivní	A
Multiplikativní	M
Obecné	O

Podle předpokladu o šumové složce ε se rozlišuje:

Konstantní rozptyl	K
Nekonstantní rozptyl	N
Autokorelace	R

Podle předpokladu o rozdělení chyb ε resultují dva typy modelů

I. symetrické rozdělení: Pro tento typ rozdělení má hustota pravděpodobnosti tvar

$$p_\varepsilon(\varepsilon_i) = Q_N \exp\left(-\frac{|\varepsilon_i|^P}{\alpha}\right) \quad (16)$$

Zde Q_N je normalizační konstanta, α je úměrné rozptylu a P specifikuje typ rozdělení.

$P=1$ Laplaceovo rozdělení

$P=2$ Normální rozdělení

$P \rightarrow \infty$ Rovnoměrné rozdělení

II. nesymetrické rozdělení: Jedna z cest nápravy je zde použití vhodné symetrizační transformace $h(\cdot)$ (tzv. TBS modely) ve které přibližně platí, že

$$h(x_i) = h(\mu) + \varepsilon_i \quad (17)$$

kde $E(\varepsilon_i) = 0$, $D(\varepsilon_i) = \sigma^2$. Pak lze pro odhad střední hodnoty použít vztah

$$\hat{\mu} = h^{-1}\left[\frac{1}{n} \sum h(x_i)\right] \quad (18)$$

Pro mocninnou transformaci $h(x) = x^\lambda$, $\lambda \neq 0$ resp $h(x) = \ln x$, $\lambda = 0$ je aproximativně

$$\hat{\mu} = \left[\frac{1}{n} \sum |x_i|^\lambda \right]^{1/\lambda} \quad (19)$$

Podle velikosti λ pak resultují různé průměry

Harmonicky \leq Geometricky \leq Aritmeticky \leq Kvadraticky

6.1 Metoda maximální věrohodnosti:

Tato metoda slouží pro odhad parametrů při známém rozdělení chyb měření. Za předpokladu nezávislých ε_i je věrohodnostní funkce resp. její logaritmus

$$\ln L(\mu) = \sum \ln p(x_i) \quad (20)$$

Maximálně věrohodný odhad je pak

$$\hat{\mu} = \max(\ln L(\mu)) \text{ ..vypočtený např. z podmínky.. } \frac{\partial \ln L(\mu)}{\partial \mu} = 0$$

Obecný model působení poruch zahrnuje jak aditivní tak multiplikativní modely

$$x_i = g(\mu, \varepsilon_i), \quad i = 1 \dots n \quad (21)$$

Při znalosti $p_\varepsilon(\varepsilon)$ je

$$p(x_i) = p_\varepsilon \left[g^{-1}(\mu, x_i, \varepsilon_i) \right] \left[\frac{\partial g^{-1}(\cdot)}{\partial x_i} \right] \quad (22)$$

A. Modely aditivní

$$g^{-1}(\cdot) = x_i - \mu$$

$$\frac{\partial g^{-1}(\cdot)}{\partial x_i} = 1$$

B. Modely multiplikativní

$$g^{-1}(\cdot) = \ln x_i - \ln \mu$$

$$\frac{\partial g^{-1}(\cdot)}{\partial x_i} = \frac{1}{x_i}$$

6.2 Typické modely

1. AKP = 2 :

Model měření má tvar

$$x_i = \mu + \sigma \varepsilon_i^*, \quad \varepsilon_i^* \approx N(0,1)$$

Rozptyl měření je

$$D(x_i) = \sigma^2$$

Odhadem parametru μ je $\hat{\mu} = \frac{1}{N} \sum x_i$ a rozptyl je $D(\hat{\mu}) = \frac{\sigma^2}{n}$

Tento model se používá jako prakticky jediný při výpočtech „standardních nejistot“.

2. ANP = 2

Model měření má tvar

$$x_i = \mu + w_i \varepsilon_i^*$$

Rozptyl měření je

$$D(x_i) = w_i^2$$

Odhadem parametru μ je $\hat{\mu} = \frac{\sum w_i x_i}{\sum w_i^{-2}}$ a rozptyl je $D(\hat{\mu}) = \frac{\sum w_i^2}{n^2}$

Speciálním případem je měření s konstantní relativní odchylkou (relativní přesností) běžné u řady měřicích přístrojů. Pak je

relativní odchylka $V = \frac{\sigma}{X} \rightarrow \sigma^2 = V^2 X^2$ a váhy $w_i = V^2 x_i^2$

Rozptyl měření je $D(x_i) = V^2 x_i^2$

Odhadem parametru μ je $\hat{\mu} = \frac{\sum x_i^{-1}}{\sum x_i^{-2}}$ a rozptyl je $D(\hat{\mu}) = \frac{V^2}{n^2} \sum x_i^2$

3. ARP = 2

Model měření má tvar

$$x_i = \mu + W_i$$

$$\text{kde } W_i = \rho W_{i-1} + \sigma \varepsilon_i^*$$

Rozptyl měření je $D(x_i) = \frac{\sigma^2}{1-\rho^2}$ a rozptyl μ je $D(\hat{\mu}) = \frac{1}{n} \left[\frac{\sigma^2}{1-\rho^2} \right]$

4. MKP = 2

Model měření má tvar

$$x_i = \mu \exp(\sigma \varepsilon_i^*)$$

Rozptyl měření je $D(\ln x_i) = \sigma^2 \rightarrow D(x_i) \approx \sigma^2 x_i^2$

Odhadem parametru μ je geometrický průměr $\hat{x} = \exp\left(\frac{1}{N} \sum \ln x_i\right)$

a rozptyl je $D(\hat{\mu}) = \frac{2}{n^2} \sum x_i^2$

Tento model lépe vystihují fyzikální měření, kdy výsledky z měřicích přístrojů jsou pouze kladné.

Pro $p \neq 2$ tj. nenormální rozdělení šumové složky je metoda maximální věrohodnosti komplikovanější. Jednoduché výrazy rezultují pro :

AKP = 1 $\hat{\mu} = \text{med}(x_i)$

AKP $\rightarrow \infty$ $\hat{\mu} = \frac{1}{2}(x_{(1)} + x_{(n)})$

7. Simulace pro výpočet nejistot nepřímých měření

Jak bylo ukázáno výše, závisí přesnost výpočtu nejistot nepřímých měření na nelinearitě funkce (viz rov (5)). Pro silně nelineární funkce $\mu = f(\bar{x}, \bar{z})$ je výpočet odpovídajícího rozptylu odhadu $D(\hat{\mu}) = \sum_i c_i^2 s_{x_i}^2 + \sum_j e_j^2 s_{z_j}^2$ silně zkreslený. S výhodou se v těchto situacích

dá použít simulační výpočet založený na myšlenkách metod Bootstrap. Principy metod Bootstrap lze jednoduše demonstrovat na příkladu konstrukci intervalu spolehlivosti populačního parametru p_s . Pro tento účel je obecně třeba znát rozdělení $g(p)$ jeho odhadu p .

Pro některá rozdělení (např. normální) a parametry (střední hodnota, rozptyl) jsou rozdělení odhadů nebo jejich funkcí známy a intervaly spolehlivosti je možné konstruovat relativně snadno.

Pro neznámé rozdělení výběru $\mathbf{x} = (x_1..x_N)$ a libovolný parametr p_s lze s výhodou použít technik Bootstrap, které umožňují jak nalezení rozdělení výběrové statistiky p , tak i konstrukci intervalu spolehlivosti. Základní myšlenka metod Bootstrap je jednoduchá[6-8]. Spočívá v generaci M -tice simulovaných výběrů $\mathbf{v}_1.. \mathbf{v}_M$ označovaných jako Bootstrap výběry. Jejich rozdělení odpovídá rozdělení původního výběru \mathbf{x} , charakterizovaného hustotou pravděpodobnosti $g(x)$. Z těchto výběrů se určí M -tice odhadů $p_i = p(\mathbf{x})$ hledaného parametru p_s . Z této M -tice hodnot lze počítat intervaly spolehlivosti pomocí celé řady metod.

A. Odhad z asymptotické normality

Jde o nejjednodušší postup založený na představě, že M je dostatečně veliké a p_i $i = 1..N$ lze zpracovat jako výběr z normálního rozdělení. Pro tzv. Bootstrap odhad střední hodnoty parametru p_s platí

$$p_B = \frac{1}{M} \sum_{i=1}^M p_i \quad (23)$$

a odpovídající rozptyl má tvar

$$s_B^2 = \frac{1}{M} \sum_{i=1}^M (p_i - p_B)^2 \quad (24)$$

Pro $100(1-\alpha)$ %ní interval spolehlivosti parametru p_s se pak použije známý vztah

$$p_B - u_{1-\alpha/2} * s_B \leq p_s \leq p_B + u_{1-\alpha/2} * s_B \quad (25)$$

kde $u_{1-\alpha/2}$ je kvantil normovaného normálního rozdělení.

B. Percentilový odhad

Tento postup je založen na neparametrickém odhadu mezi intervalu spolehlivosti vycházejícím z pořádkových statistik $p_{(i)}$, kde $p_{(i)} \leq p_{(i+1)}$ jsou pořádkové statistiky, pro které platí, že jsou d %ním kvantilem rozdělení odhadu p pro

$$d = \frac{i}{M+1}$$

Dolní mez $100(1-\alpha)$ %ní intervalu spolehlivosti je pak

$$LC = p_{(k1)} \text{ kde } k1 = \text{int}[\alpha * (M+1) / 2] \quad (26)$$

a pro horní mez platí

$$UC = p_{(k2)} \text{ kde } k2 = \text{int}[(1-\alpha/2) * (M+1)] \quad (27)$$

Zde $\text{int}(x)$ je celá část čísla x .

C. Studentizovaný odhad

Tento odhad vychází z jednoduché transformace vedoucí na Studentizovanou náhodnou veličinu t_i

$$t_i = \frac{p_i - p_B}{s_{Bi}}$$

kde s_{B_i} je výběrová směrodatná odchylka počítaná pro i -tý Bootstrap výběr v_i . Pro $100(1-\alpha)$ %ní interval spolehlivosti pak platí

$$p_B - t_D * s_B \leq p_S \leq p_B + t_D * s_B \quad (28)$$

kde pořádková statistika $t_D = t_{(int[\alpha*(M+1)/2])}$ a pořádková statistika $t_H = t_{(int[(1-\alpha/2)*(M+1)])}$

D. Vyhlazený odhad

Obecně lze na základě hodnot p_i sestavit odhad hustoty pravděpodobnosti jejich rozdělení $fe(p)$ např. s využitím histogramu nebo jádrového odhadu. Při znalosti funkce $fe(p)$ se snadno konstruuje interval spolehlivosti přímo z definice. Pro meze tohoto intervalu pak platí, že

$$\alpha / 2 = \int_{-\infty}^{LC} fe(p) dp$$

a

$$\alpha / 2 = \int_{UC}^{\infty} fe(p) dp$$

Podle typu odhadu fe může jít o úlohu numerické nebo analytické integrace.

Základním předpokladem úspěšnosti celého postupu je **generace Bootstrap výběrů**. Pro tento účel je třeba buď znát nebo volit rozdělení $g(x)$. Standardní technika **neparametrického Bootstrap** vychází z neparametrického odhadu $g(x)$ ve tvaru

$$g(x) = \frac{1}{N} \delta(x - x_i)$$

kde Diracova funkce $\delta(x - x_i) = 1$ pro $(x = x_i)$ a všude jinde je. $\delta(x - x_i) = 0$.

Toto rozdělení pokládá pravděpodobnost $1/N$ v každém bodě. Simulované výběry se pak realizují jako náhodné výběry složené z prvků původního výběru x s vrácením (tj. jeden prvek původního výběru se může v simulovaném výběru vyskytovat i opakovaně). Tato technika se pro účely výpočtu nejistot nepřímých měření nehodí, protože nezohledňuje nejistoty typu **B**.

Další možností je konstruovat vhodný **parametrický model** $g(x)$, odhadnout jeho parametry a generovat simulované výběry standardními postupy. Tento přístup naráží na celou řadu problémů souvisejících s možnou nehomogenitou, vybočujícími body, heteroskedasticitou a autokorelací. Parametrický model se dá použít i pro nejistoty typu B (kde vlastně při opakování experimentů za stejných podmínek jsou odpovídající příspěvky k celkovému rozptylu nulové).

Bootstrap metody obecně poskytují informace jak o bodových odhadech, tak i intervalech spolehlivosti.

Uvažujme standardní neparametrický Bootstrap (v_i jsou výběry s vrácením) pro $p_S = \mu$, tj. jde o střední hodnotu a její interval spolehlivosti střední hodnoty. Lze snadno určit, že v tomto případě je Bootstrap průměr totožný s aritmetickým průměrem původních dat a Bootstrap rozptyl je M -krát menší než rozptyl původních dat. Liší se však intervaly spolehlivosti zejména tam, kde se rozdělení dat výrazně odchyluje od normálního rozdělení.

Kromě standardního Bootstrap lze použít také **dvojitý Bootstrap** (Bootstrap aplikovaný na výběry v_i), **blokový Bootstrap** (realizace výběru s vrácením na bloky homogenních dat a sestavení celkového Bootstrap výběru spojením výsledků). [7]

Z hlediska realizace metod Bootstrap na počítači je základem generace simulovaných výběrů. Velmi jednoduše se dá tato operace provést v jazyku MATLAB s využitím vektorového triku.

Úsek programu má tvar

```
ar=load('dat.txt');[c s]=size(ar); b=800;
if c ==1
    ar=ar';c=s;
end
B=ar(ceil(c*rand(c,b)));
```

Předpokládá se, že n -tice dat je v souboru *dat.txt* a b – tice Bootstrap výběrů je v poli B . Pro výpočet odhadu p_i se používá standardních postupů. Výpočet intervalů spolehlivosti je pak závislý na volbě přístupu.

Při výpočtu nejistot nepřímých měření lze v zásadě použít parametrický Bootstrap s tím, že pro nejistoty typu A se v každém simulovaném výběru určí \bar{x} z předpokládaného rozdělení a pro nejistoty typu B se určí příspěvky k chybě měření ε_i (standardní předpoklad je, že $E(\bar{z}) = 0$). Konkrétní kombinace $f(\bar{x}_i)$ a ε_i se zvolí na základě vybraného modelu měření (viz. kap. 7).

8. Závěr

Je patrné, že výpočet nejistot, jak je navržen ISO a EURACHEM je použitelný jen za speciálních předpokladů o působení poruch, typu modelované funkce a zdrojích nejistot. Pro složitější situace je vždy lépe nejdříve nalézt vhodný model měření a v jeho rámci pak provádět stanovení intervalu neurčitosti. Také problém náhodných a systematických neexperimentálních chyb není ještě uspokojivě dořešen.

Poděkování: Tato práce vznikla s podporou grantů MŠMT VCT II No. 1M0553 a CQR No. 1M06047.

9. Literatura

- [1] Quantifying Uncertainty in Analytical Measurement, EURACHEM 1995
- [2] Taylor B., Kuyatt CH.E. : Guidelines for Evaluation and Expressing the Uncertainty of NIST Measurement Results, NIST Tech. Note 1297, 1994
- [3] D Agostini G. : Probability and Measurement Uncertainty in Physic, Rept. DESY 95-242, Roma December 1995
- [4] Phillips S.D., Eberhart K. R., Parry B.: Guidelines for Expressing the Uncertainty of Measurement Results Containing Uncorrected Bias, J. Res. Natl. Inst. of Standards **102**, 577 (1997)
- [5] Meloun M., Militký J., Forina M.: Chemometrics for Analytical Chemistry, vol I, Ellis Horwood, Chichester, 1992
- [6] Meloun M., Militký J.: Statistická analýza experimentálních dat, Academia Praha 2004
- [7] Wekrens, R. a kol.: Chem.Int. Lab. Systems **54**, 35-52 (2000)
- [8] Davidson, A., Hinkley, D.V.,: Bootstrap Methods and Their Applications, Cambridge Univ. Press, Cambridge, 1997