



Editor: Zbyněk Plzák (plzak@iic.cas.cz)

Odhad směrodatné odchylky preciznosti z duplicitních výsledků

Mezi základní charakteristiky výkonnosti analytického měřicího postupu patří preciznost. Charakterizuje se směrodatnou odchylkou, která se často vyhodnocuje analýzou dat z opakovaných měření. V laboratoři jsou často k dispozici rozsáhlé soubory dat duplikátních měření vzorků. Jak taková data použít k odhadu směrodatné odchylky preciznosti, o tom pojednává tento metodický list.

Úvod

Rozptýlení opakovaných výsledků je jedním ze základních charakteristik měřicího postupu, je charakterizováno precizností, kterou definuje Mezinárodní metrologický slovník VIM3 [1] jako *těsnost shody mezi indikacemi nebo naměřenými hodnotami veličiny získanými opakovanými měřeními na stejném objektu nebo na podobných objektech za specifikovaných podmínek*. Preciznost měření je zpravidla vyjádřena číselně mírami nepřeciznosti, jako například směrodatnou odchylkou za specifikovaných podmínek měření. Opakovatelnost měření a reprodukovatelnost měření představují dvě hraniční míry preciznosti v závislosti na specifikovaných podmínkách měření. Opakovatelnost odráží variabilitu získanou za podmínek, kdy jsou měření prováděna jedním analytikem na stejném vybavení v krátkém časovém intervalu. Reprodukovatelnost odráží variabilitu výsledků získaných v různých laboratořích a její vyhodnocení vyžaduje mezilaboratorní experiment. Mezi těmito mírami se nachází mezilehlá preciznost, která charakterizuje rozptýlení výsledků získaných v jediné laboratoři za podmínek, které jsou volnější než u opakovatelnosti (např. provádění analýz různými analytiky, na odlišném vybavení a v delším časovém úseku) a musí být specifikovány. Preciznost se vyhodnocuje v rámci validační studie [2] typicky jako vyhodnocení série opakovaných měření, jako jedna ze základních charakteristik měřicího postupu. Slouží i k odhadu nejistoty měření, obvykle jako jedna z významných složek bilance nejistoty. Preciznost obecně závisí na koncentraci analytu a měla by se tedy stanovovat pro několik koncentrací v rozmezí pracovního rozsahu metody. Pokud k jejímu vyjádření použijeme směrodatnou odchylku, můžeme ji kvantifikovat jako jedinou, konstantní hodnotu obvykle jen pro určité úzké rozmezí koncentrací. Často je preciznost úměrná koncentraci analytu a v takových případech je vhodné vyjadřovat preciznost jako relativní směrodatnou odchylku, která je pak konstantní pro určené širší rozmezí koncentrací, ev. v celém pracovním rozsahu měřicího postupu. Tento metodický list hovoří o koncentraci, uváděný postup lze však analogicky použít i pro další veličiny měřené opakovaně v analytických i zdravotnických laboratořích.

V rutinních laboratořích se často provádějí analýzy určitého podílu vzorků i kontrolních vzorků duplicitně. Je tak k dispozici často rozsáhlý soubor dat, který lze použít k věrohodnému odhadu směrodatné odchylky, který může sloužit jako statisticky vydatný odhad pro porovnání s hodnotou preciznosti získanou v prvotní validační studii nebo pro zpřesnění bilance při odhadu nejistoty měření. Vzhledem k tomu, že vyhodnocení rozptýlení statistického výběru stanovení prováděných duplicitně nebývá často v statistických příručkách pro takový datový soubor popisováno, zabývá se jím tento metodický list.

Zpracování datového souboru duplicitních měření

Mějme tedy soubor k dvojic naměřených hodnot duplicitních měření (x_{i1}, x_{i2}) různých vzorků. Většinou můžeme předpokládat, že závislost preciznosti na koncentraci lze v celém pracovním rozmezí nebo alespoň v širokém vybraném koncentračním rozmezí vystihnout přímou úměrou. Pak je vhodné rozdíly

„normalizovat“ průměry daných dvojic a z nich počítat relativní směrodatnou odchylku, jak je to popsáno např. v publikaci [3]. Tento způsob zpracování dat popíšeme jako obecný přístup. U dat, kde preciznost nezávisí na koncentraci, můžeme postupovat analogickým způsobem (viz dále).

Postup bude sestávat ze čtyř kroků:

1. normalizace primárních rozdílů duplicitních stanovení
2. zjištění přítomnosti či absence odlehlých hodnot ve zkoumaném výběru pomocí průzkumové diagnostiky – např. histogramu, ev. statistických testů
3. výpočtu směrodatné odchylky normalizovaných rozdílů
4. výpočtu směrodatné odchylky jednotlivého stanovení jako vyjádření preciznosti

Rozdíly duplicitních stanovení je nutno normalizovat vydělením jejich průměrnými hodnotami.

$$D_i = \frac{(x_{i1} - x_{i2})}{\frac{(x_{i1} + x_{i2})}{2}} \quad (1)$$

Ke zjištění, zda soubor rozdílů obsahuje odlehlé hodnoty lze použít grafickou diagnostiku např. histogram zkonstruovaný v tabulkovém editoru. K tomu lze využít např. tabulkový editor MS Excel s aktivovaným doplňkem Analýza dat (do karty DATA se aktivuje přes MOŽNOSTI>Doplňky) nebo jiný software či statistické testy, na jejichž použití pro tento účel je laboratoř zvyklá. Existenci odlehlých hodnot je třeba podrobit odborné analýze, může se jednat např. o vzorek s odlišnou matricí, nebo jiný problém při měření. Vyloučení takové dvojice je třeba vždy pečlivě zvážit.

Takto prověřený soubor normalizovaných rozdílů D_i lze použít k výpočtu relativní směrodatné odchylky s_D podle následujícího vzorce [odkaz 3]:

$$s_D = \sqrt{\frac{\sum_{i=1}^k (D_i - \bar{D})^2}{k-1}} \quad (2)$$

kde \bar{D} je průměr rozdílů k dvojic. Při výpočtu dle vzorce 2 nesmí být v dvojicích měněno pořadí první a druhé hodnoty, musí být totiž zachována náhodnost znamének rozdílů dle výskytu v průběhu měření, (nesmí být také počítáno s absolutními hodnotami diferencí), v opačném případě by byla změněna hodnota průměru \bar{D} , a tudíž i odhadu směrodatné odchylky. Mělo by být k dispozici tolik vzorků, aby počet dvojic k byl větší či roven 10, což je v praxi většinou bohatě splněno. Počet stupňů volnosti při tomto výpočtu $v = k - 1$.

Rozptyl (čtverec směrodatné odchylky) rozdílů je dvojnásobkem rozptylu individuálních hodnot, a je proto pro odhad relativní směrodatné odchylky jednotlivých hodnot s_{rel} nutno podělit směrodatnou odchylku normalizovaných rozdílů $\sqrt{2}$:

$$s_{rel} = \frac{s_D}{\sqrt{2}} \quad (3)$$

Výpočet tohoto typu demonstruje tabulka A4.2. odkazu 3, který lze použít jako vzorový příklad. Výpočet lze provést výhodně ve shodném rozfázování za použití nástrojů tabulkového editoru MS Excel, aplikujeme funkci směrodatné odchylky výběru SMODCH.VÝBĚR

Pokud je zřejmé, že průměr \bar{D} se významně neliší od nuly, lze počítat relativní směrodatnou odchylku s_D postupem uváděným např. v dokumentech IUPAC [4] a v učebnicích K. Eckschlagera [např. 5]:

$$s_D = \sqrt{\frac{\sum_{i=1}^k D_i^2}{k}} \quad (4)$$

Počet stupňů volnosti je v tomto případě roven hodnotě k . (Signifikantně nenulový průměr \bar{D} by odporoval podmínce, že střední hodnota diferencí je rovna nule, za které byl vzorec 4 odvozen [6].) Výpočet lze provést za použití funkcí v tabulkovém editoru MS Excel: =ODMOCNINA(SUMA.ČTVERCŮ(adresa pole relativních rozdílů)/POČET(adresa pole relativních rozdílů)) (5)

Nalezený soubor rozdílů duplicitních výsledků může v některých případech vykazovat nezávislost rozptylu absolutních rozdílů na koncentraci. I v takovém případě je možné počítat relativní směrodatnou odchylku za použití rovnice 2 nebo 4 s relativními rozdíly, ovšem za předpokladu, že se krajní hodnoty koncentračního rozmezí, v kterém byly výsledky naměřeny, relativně příliš neliší, tj. poměr krajních hodnot koncentrací není příliš odlišný od 1 (na úzkém koncentračním rozmezí se koncentrační závislost neprojeví). Pokud se však krajní hodnoty relativně značně liší, např. řádově, nelze počítat s relativními rozdíly, protože vydělením absolutních rozdílů tak odlišnými průměry dostaneme relativní rozdíly, jejichž rozptýlení bude na tomto rozmezí výrazně klesat s rostoucí koncentrací. V tomto případě je třeba počítat z dat směrodatnou odchylku (nikoliv relativní směrodatnou odchylku) za použití vzorců 2 nebo 4 s tím, že dosazujeme za D_i absolutní rozdíly místo relativních (tj. bez dělení průměry dvojic dle vztahu 1) a výpočet směrodatné odchylky pak dokončíme dle vzorce 3.

Při rozhodování, zda zvolit postup s relativními nebo absolutními rozdíly a na jakém koncentračním rozmezí, nám může pomoci grafické vyšetřování souboru rozdílů dvojic. Sestrojíme grafy závislosti absolutních hodnot rozdílů na koncentraci - na vodorovnou osu vynášíme průměry dvojic, na svislou osu jednak absolutní hodnoty rozdílů dvojic (tj. všechny rozdíly bereme kladně), jednak absolutní hodnoty relativních rozdílů. Podle toho, jak se mění rozptýlení diferencí a relativních diferencí s rostoucím průměrem, můžeme usoudit, zda je možné počítat např. jedinou relativní směrodatnou odchylku pro celé koncentrační rozmezí, nebo je třeba hodnotit preciznost zvlášť pro několik koncentračních rozmezí (většinou stačí zvlášť počítat společnou směrodatnou odchylku pro koncentrace v blízkosti meze detekce a zvlášť společnou relativní směrodatnou odchylku pro zbylou horní oblast koncentračního rozmezí).

Tímto způsobem získáme směrodatnou odchylku, která odráží variabilitu měřených hodnot za podmínek, za nichž byly získávány dvojice duplicitních výsledků. Dva odečítané výsledky pro daný vzorek bývají velmi často naměřeny za podmínek opakovatelnosti, ovšem v důsledku použití dvojic pro různé vzorky měřené v delším časovém rozmezí pak do jisté míry odrážejí náhodné jevy dané různými variantami daného typu vzorku (variabilita matrice a koncentrací), různými sériemi rozkladu a měření na přístroji a různými operátory. Směrodatná odchylka vypočtená ze souboru takových rozdílů nemůže být tedy považována za míru opakovatelnosti, ale za určitou míru mezilehlé preciznosti.

Závěr

Zpracování rozdílů duplicitních stanovení, která jsou často bez zvláštního úsilí v laboratořích k dispozici v dostatečném počtu, lze získat věrohodný odhad směrodatné odchylky jako kvantitativního vyjádření mezilehlé preciznosti. S rostoucím počtem dvojic měření totiž rychle klesá vychýlení odhadu a současně se snižuje i jeho rozptyl. Doporučený postup výpočtu umožňuje tedy prověřit jednu ze základních charakteristik výkonnosti měřicího procesu a získat tak i další vstupní kvantitativní údaj pro odhad nejistoty měření.

Zpracoval: V. Synek a Z. Plzák

Znění Metodického listu prošlo v roce 2018 revizí. Verze z roku 2016 je nadále platná beze změn.

Literatura

1. TNI 01 0115:2009 *Mezinárodní metrologický slovník - Základní a všeobecné pojmy a přidružené termíny* (VIM).
2. Kvalimetrie 20. *Vhodnost analytických metod pro daný účel*. M. Suchánek a D. Milde (eds.) Eurachem-ČR Praha 2015, ISBN 978-80-86322-00-1.
3. Kvalimetrie 19. *Stanovení nejistoty analytického měření*, str.75. M. Suchánek a D. Milde (eds.) Eurachem-ČR Praha 2014, ISBN 978-80-86322-07-0.
4. IUPAC (1998) *Compendium of analytical nomenclature - Definitive Rules 1997 („The Orange Book“)*, 3rd edn. ch. 2.3. Dostupné z http://old.iupac.org/publications/analytical_compendium/.
5. Doerffel K., Eckschlager K. *Optimální postup chemické analýzy*. SNTL, Praha 1985, str.148.
6. Synek V.: *Evaluation of the standard deviation from duplicate results*. Accred. Qual. Assur. **13**, 335 (2008).