STATISTICKÉ ZPRACOVÁNÍ VÝSLEDKŮ PROGRAMŮ ZKOUŠENÍ ZPŮSOBILOSTI

Poskytovatel programů zkoušení způsobilosti při SZK FAST
Veveří 95, 602 00 Brno
Czech Republic

www.szk.fce.vutbr.cz

Datum vydání: 28. 2. 2017

Vypracoval (28. 2. 2017):	Ing. Petr Misák, Ph.D.
Kontroloval (28. 2. 2017):	doc. Ing. Tomáš Vymazal, Ph.D.
Schválil za PoZZ (28. 2. 2017):	doc. Ing. Tomáš Vymazal, Ph.D.

Obsah

1 Postupy statistické analýzy experimentu shodnosti

K popisu přesnosti metod měření se využívá termínů správnost a shodnost. Správnost se týká těsnosti shody mezi aritmetickým průměrem velkého počtu výsledků zkoušek a pravou nebo přijatou referenční hodnotou. Shodnost se týká těsnosti shody mezi výsledky zkoušek. Nutnost uvažování shodnosti vzniká ze skutečnosti, že zkoušky, o nichž se předpokládá, že jsou provedeny na stejném materiálu za stejných podmínek, neposkytují obecně stejné výsledky. Příčinou jsou náhodné chyby, kterým se nelze vyhnout. Tyto chyby jsou nedílnou součástí každého zkušebního postupu a nelze je nikdy v plném rozsahu ovládat. Analýza experimentu shodnosti není zaměřena na zkoumání správnosti výsledků zkoušek, ale především na jejich shodnost. Výsledky se tedy posuzují vzájemně mezi sebou a nikoli vzhledem k nějaké referenční nebo pravdivé hodnotě.

Základem statistické analýzy je kritické zhodnocení údajů podle ČSN EN 5725-2 [1], tedy zjištění a ošetření podezřelých a odlehlých hodnot a dalších nepravidelností. Toto zhodnocení se provádí prostřednictvím Mandelových statistik (grafické zhodnocení) a především pomocí Grubbsových a Cochranových testů (numerické zhodnocení). Dalšími sledovanými statistickými parametry jsou mezilaboratorní rozptyl, rozptyl opakovatelnosti a reprodukovatelnosti a na ně navazující charakteristiky opakovatelnost a reprodukovatelnost. Výsledkem programu MPZ je vyhodnocení výkonnosti zúčastněných laboratoří (účastníků) podle ČSN EN ISO/IEC 17043 [2], které se skládá z určení vztažných hodnot a jejich nejistot a následného porovnání s výsledky zkoušek účastníků MPZ.

Předpokladem pro použití těchto metod je jednovrcholové rozdělení pravděpodobnosti naměřených dat. Dále označme p počet účastnících se laboratoří označených indexem $i = 1, \dots, p$, z nichž každá provedla n zkoušek.

1.1 Numerický postup zjišťování odlehlých hodnot

Ke zjišťování odlehlých hodnot se používají dva základní statistické testy. Prvním z nich je Cochranův test, který je testem vnitrolaboratorních variabilit (je-li počet měření jedné veličiny v jedné laboratoři > 2) a používá se jako první. Pokud tento test označí výsledky jedné z laboratoří jako odlehlé, musí se laboratoř vyřadit a test zopakovat. Druhý test (Grubbsův) je předně testem mezilaboratorní variability a lze ho rovněž použít, když Cochranův test vzbudí podezření, zda vysoké vnitrolaboratorní rozptýlení lze připsat na vrub pouze jednoho z výsledků zkoušek. Oba tyto testy předpokládají vyváženost experimentu, tedy mělo by platit, že počet zkoušek v jedné laboratoři pro stanovení jedné veličiny je konstantní.

Při zjišťování vybočujících nebo odlehlých hodnot mohou nastat tři případy:

- Je-li testová statistika menší než 5% kritická hodnota nebo je-li této hodnotě rovna, považuje se testovaná entita za správnou;
- Je-li testová statistika větší než 5% kritická hodnota a menší než 1% kritická hodnota nebo je-li této hodnotě rovna, nazve se testovaná entita **vybočující**;
- Je-li testová statistika větší než 1% kritická hodnota, nazve se testovaná entita **odlehlou** hodnotou.

1.1.1 Cochranův test

Cochranova statistika C je dána vztahem

$$C = \frac{s_{max}^2}{\sum_{i=1}^{p} s_i^2} \tag{1}$$

kde s_{max} je největší výběrová směrodatná odchylka, s_i jsou výběrové směrodatné odchylky stanovené na základě výsledků ve všech laboratořích a p je počet účastnících se laboratoří v experimentu.

Výběrová směrodatná odchylka se stanovuje ze vztahu

$$s_i = \sqrt{\frac{1}{n_i - 1} \sum_{k=1}^{n_i} (y_k - \bar{y})^2}, \tag{2}$$

kde n_i je počet výsledků zkoušek stanovení jedné veličiny v i-té laboratoři, y_k je k-tá hodnota a $bary_i$ je aritmetický průměr hodnot změřených v i-té laboratoři. Jsou-li pro sledovanou veličinu naměřeny pouze dva výsledky, je možné použít zjednodušeného vztahu

$$s_i = \frac{|y_1 - y_2|}{\sqrt{2}}. \tag{3}$$

1.1.2 Grubbsův test – jedno odlehlé pozorování

Z dané množiny údajů x_i pro $i = 1, 2, ..., p$, uspořádané vzestupně podle velikosti, se pro určení použitím Grubbsova testu, zda je největší pozorování odlehlou hodnotou, vypočte Grubbsova statistika G_p

$$G_p = \frac{x_p - \bar{x}}{s}, \tag{4}$$

přičemž $\bar{x}$ je aritmetický průměr sledovaného znaku. Sledovaným znakem může být průměrná hodnota určované veličiny v rámci laboratoře. Dále je s výběrová směrodatná odchylka sledovaného znaku, tedy v tomto případě směrodatná odchylka počítána přes všechny laboratoře.

Pro test významnosti nejmenšího pozorování se vypočte testová statistika

$$G_p = \frac{\bar{x} - x_p}{s}. \tag{5}$$

1.2 Mandelovy statistiky

Pro zjišťování konzistence dat se použily dvě míry, nazývané Mandelovy statistiky h a k. Běžně se tyto míry používají pro grafické hodnocení laboratoří podobně jako popis variability.

1.2.1 Mezilaboratorní statistika konzistence h

Pro každou laboratoř se vyhodnotila mezilaboratorní statistika konzistence h podle vzorce

$$h_i = \frac{\bar{y}_i - \bar{\bar{y}}}{\sqrt{\frac{1}{p-1}\sum_{i=1}^{p_j}\left(\bar{y}_i - \bar{\bar{y}}\right)^2}}. \tag{6}$$

1.2.2 Vnitrolaboratorní statistika konzistence k

Vnitrolaboratorní statistika konzistence k se vypočítá podle vztahu

$$k_i = \frac{s_i\sqrt{p}}{\sqrt{\sum_{i=1}^{p} s_i^2}}. \tag{7}$$

kde s_i je výběrová směrodatná odchylka hodnot naměřených v i-té laboratoři. Stejně jako statistika h se hodnoty k vynášejí do grafů. Prohlídka grafů hodnot h a k může naznačovat, že u určitých laboratoří se ukazuje podstatně odlišné uspořádání výsledků než u ostatních studovaných laboratoří. Je to důsledkem trvale velkého a/nebo malého rozptylu výsledků nebo extrémních průměrů výsledků napříč úrovněmi.

1.3 Výpočet odhadů rozptylů

Po vyřazení odlehlých hodnot (laboratoří) je možné přikročit k výpočtu základních charakteristik variability, a to rozptylu opakovatelnosti, mezilaboratorního rozptylu a rozptylu reprodukovatelnosti. Tyto charakteristiky se uvádějí ve formě směrodatných odchylek, tedy po odmocnění. Výhodou je stejný fyzikální rozměr charakteristiky variability a sledované veličiny.

1.3.1 Rozptyl opakovatelnosti

$$s_r^2 = \frac{\sum_{i=1}^{p}(n_i - 1)s_i^2}{\sum_{i=1}^{p}(n_i - 1)} \tag{8}$$

1.3.2 Mezilaboratorní rozptyl

$$s_L^2 = \frac{s_d^2 - s_r^2}{\bar{\bar{n}}}, \tag{9}$$

kde

$$s_d^2 = \frac{1}{p-1}\sum_{i=1}^{p} n_i\left(\bar{y_i} - \bar{\bar{y}}\right)^2 \tag{10}$$

a

$$\bar{\bar{n}} = \frac{1}{p-1}\left[\sum_{i=1}^{p} n_i - \frac{\sum_{i=1}^{p} n_i^2}{\sum_{i=1}^{p} n_i}\right]. \tag{11}$$

1.3.3 Rozptyl reprodukovatelnosti

$$s_R^2 = s_r^2 + s_L^2, \tag{12}$$

kde s_r^2 je rozptyl opakovatelnosti a s_l^2 je mezilaboratorní rozptyl.

1.4 Opakovatelnost a reprodukovatelnost

Opakovatelnost vyjadřuje, že rozdíl mezi dvěma výsledky zkoušek z téhož vzorku, provedených stejným pracovníkem, na tomtéž zařízení, v nejkratším možném časovém intervalu nebude překračovat hodnotu opakovatelnosti r v průměru ne více než jednou ve 20 případech při běžném a správném provádění metody.

Hodnota opakovatelnosti je vyjádřena vztahem

$$r = 2,8 s_r, \tag{13}$$

kde $s_r = \sqrt{s_r^2}$ je směrodatná odchylka opakovatelnosti.

Reprodukovatelnost vyjadřuje, že výsledky zkoušek na tomtéž vzorku, získané v nejkratším možném časovém intervalu dvěma pracovníky, kteří použili každý své zařízení, se nebudou lišit hodnotou reprodukovatelnosti R v průměru ne více než jednou ve 20 případech při běžném a správném provádění metody.

Hodnota reprodukovatelnosti je vyjádřena vztahem

$$R = 2,8 s_R, \tag{14}$$

kde $s_R = \sqrt{s_R^2}$ je směrodatná odchylka reprodukovatelnosti.

1.5 Vztažná hodnota

Organizátor MPZ zajistí stanovení vztažné hodnoty X a její nejistoty pro každý program MPZ. Vztažné hodnoty jsou vždy účastníkům MPZ sdělovány až po dodání výsledků MPZ a to tak, aby účastníci nemohli získat žádnou výhodu z jejich předčasného zveřejnění.

Vztažné hodnoty organizátor stanovuje jako konsenzuální hodnotu účastníků podle přílohy B normy ČSN EN ISO/IEC 17043 [2] za použití statistických metod popsaných v ISO 13528 [3] a ČSN ISO 5725-5 [4]. Vztažná hodnota X je tedy určena jako robustní odhad hodnoty průměru x^* (**Algoritmus A** uvedený v [3] a [4]).

Vypočtou se počáteční hodnoty x^* a s^* (robustní směrodatná odchylka) jako

$$x^* = \text{medián } x_i, \tag{15}$$

$$s^* = 1,483 \cdot \text{medián} \left|x_i - x^*\right|, \tag{16}$$

kde $i = 1, \ldots, p$. Hodnoty x^* a s^* se upraví následovně. Vypočte se $\varphi = 1,5 \cdot s^*$. Pro každou hodnotu x_i se vypočte

$$x_i^* = \begin{cases} x^* - \varphi & \text{jestliže } x_i < x^* - \varphi, \\ x^* + \varphi & \text{jestliže } x_i > x^* + \varphi, \\ x_i & \text{v ostatních případech.} \end{cases} \tag{17}$$

Vypočtou se nové hodnoty x^* a s^* ze vztahu

$$x^* = \sum_{i=1}^{p} \frac{x_i^*}{p}, \tag{18}$$

a

$$s^* = 1,134 \cdot \sqrt{\sum_{i=1}^{p} \frac{(x_i^* - x^*)^2}{p-1}}. \tag{19}$$

Robustní odhady se odvozují iterací, dokud nejsou změny odhadů od jednoho výpočtu k dalšímu malé.
Standardní nejistota u_X takto stanovené vztažné hodnoty se určuje podle vztahu

$$u_X = 1,25 \frac{s^*}{\sqrt{p}}. \tag{20}$$

V případě malého počtu účastníků MPZ stanovuje organizátor vztažné hodnoty jako konsenzuální hodnoty získané od expertních účastníků, kteří prokázali kompetenci ke stanovení měřené veličiny, která je předmětem zkoušky.

Při nízkém počtu účastníků ($4 \leq p \leq 20$) může organizátor dále zvážit využití tzv. **Hornova postupu** pro stanovení vztažných hodnot. Tento postup spočívá ve stanovení tzv. pivotů, na jejichž základě se určí odhad polohy a variability. Nejdříve se provede vzestupné seřazení posuzovaných dat. Dolní pivot se poté určí ze vztahu

$$x_D = x_{(H)}, \tag{21}$$

kde H je pořadový index daný rovnicí $H = \frac{\mathrm{int}\left(\frac{p+1}{2}\right)}{2}$ nebo $H = \frac{\mathrm{int}\left(\frac{p+1}{2}+1\right)}{2}$.
Horní pivot se poté určí ze vztahu

$$x_H = x_{p+1-H}. \tag{22}$$

Vztažná hodnota je prostřednictvím Hornova postupu určena jako odhad polohy, tedy tzv. pivotová polosuma

$$x^* = \frac{x_D + x_H}{2}. \tag{23}$$

Odhad variability se stanovuje jako tzv. pivotové rozpětí

$$R_L = x_H - x_D \tag{24}$$

a nejistota takto určené vztažné hodnoty jako 95% intervalový odhad střední hodnoty

$$u_X = R_L t_{L;0,95}(p), \tag{25}$$

kde $t_{L;0,95}(p)$ je $(1-\alpha)$ kvantil rozdělení T_L s p stupni volnosti.

1.6 Výpočet statistik výkonnosti

Výsledky zkoušek se musí pro interpretaci a porovnání se stanovenými cíli převést na tzv. výkonnostní statistiky. Účelem je vyjádřit odchylku od vztažné hodnoty takovým způsobem, který umožňuje porovnání s kritérii výkonnosti. Podle normy ČSN EN ISO/IEC 17043 [2] se výkonnost účastnících se pracovišť hodnotí podle tzv. z-score a ζ-score (zeta-score).

Pro každou neodlehlou laboratoř se z-score vypočte podle vztahu

$$z_i = \frac{|\bar{x_i} - x^*|}{s^*}. \tag{26}$$

ζ-score (zeta-score) se vypočítá pomocí rovnice

$$\zeta_i = \frac{|\bar{x_i} - x^*|}{\sqrt{u_i^2 + u_X^2}}, \tag{27}$$

kde u_i je standardní kombinovaná nejistota i-té laboratoře. Standardní kombinované nejistoty měření lze získat podělením rozšířené nejistoty U koeficientem rozšíření k, který má pro normální rozdělení pravděpodobnosti hodnotu $k = 2$. Pokud účastník neuvedl rozšířenou nejistotu měření na záznamovém listu výsledků zkoušek, není možné ζ-score určit. Více o nejistotách měření lze nalézt v dokumentu [5].

Pro z-score a ζ-score (pro jednoduchost je uvedeno pouze z-score) platí následující stupnice:

$$z\text{-score} = \begin{cases} |z| \leq 2 & \text{ukazuje, že výkonnost laboratoře je } \textbf{vyhovující}, \\ 2 \leq |z| \leq 3 & \text{ukazuje, že výkonnost laboratoře je } \textbf{problematická} \text{ a vytváří varovný podnět,} \\ 3 \leq |z| & \text{ukazuje, že výkonnost laboratoře je } \textbf{nevyhovující} \text{ a vytváří podnět k akci.} \end{cases} \qquad (28)$$

Odkazy

[1] ČSN ISO 5725-2. *Přesnost (správnost a shodnost) metod a výsledků měření – Část 1: Základní metoda pro stanovení opakovatelnosti a reprodukovatelnosti normalizované metody měření*. 1997.

[2] ČSN EN ISO/IEC 17043. *Posuzování shody - Všeobecné požadavky na zkoušení způsobilosti*. 2010.

[3] ISO 13 528. *Statistical methods for use in proficiency testing by interlaboratory comparisons*. 2005.

[4] ČSN ISO 5725-5. *Přesnost (správnost a shodnost) metod a výsledků měření – Část 5: Alternativní metody pro stanovení shodnosti normalizované metody měření*. 1999.

[5] EA 4/02. *Vyjadřování nejistot měření při kalibracích*. 2000.