



Seriál: Zpracování dat fyzikálních měření

V minulém díle seriálu jsme si rozebrali, jak se matematicky modeluje experimentální měření fyzikálních veličin. Řekli jsme si, že hodnotu, kterou nám ukáže měřicí přístroj, považujeme za náhodnou veličinu se střední hodnotou rovnou skutečné hodnotě měřené fyzikální veličiny – tedy skutečná hodnota měřené fyzikální veličiny je pevná a neměnná, zatímco naše naměřená data jsou v určitém smyslu náhodná (toto si připomeňte, budeme to potřebovat). Dále jsme si odvodili, jak se v nejjednodušším modelu, kdy uvažujeme, že měřená data mají normální rozdělení, konstruuje odhad měřené fyzikální veličiny a interval spolehlivosti pro měřenou fyzikální veličinu (což fyzikové nazývají nejistota měření).

V tomto díle seriálu se budeme zabývat dvěma věcmi. V první části si rozebereme, jak postupovat v obecnějším případě, kdy naše měřená data nemají normální rozdělení. Uvidíme, že se v takovémto případě postupuje velmi podobně jako v případě normálně rozdělených dat (což už bychom měli umět). Ve druhé části tohoto dílu seriálu se budeme zabývat situací, kdy potřebujeme naměřené fyzikální veličiny (které jsou naměřeny s určitou nejistotou) dosazovat do dalších fyzikálních vzorců. Odvodíme si, jak v takovémto případě přepočítat nejistotu měření, abychom mohli konstruovat intervaly spolehlivosti pro naše transformovaná data.

Centrální limitní věta (CLV)

V této kapitole si povíme, jak zpracovávat naměřená data, která se neřídí normálním rozdělením (nebo to o nich nemůžeme s jistotou prohlásit). V případě bodových odhadů střední hodnoty a rozptylu není žádný rozdíl oproti případu, kdy by námi naměřená data měla normální rozdělení. Odhady uvedené v minulém díle, tedy výběrový průměr

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}_n,$$

výběrový rozptyl

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}_n)^2$$

a výběrová směrodatná odchylka

$$S_n = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}_n)^2}$$

budou mít i v tomto obecnějším případě stejné vlastnosti jako v případě normálního rozdělení (tedy, že pro velký počet měření bude jejich hodnota s největší pravděpodobností velmi blízká skutečným hodnotám střední hodnoty resp. rozptylu, resp. směrodatné odchylky).

Problém nastane u konstrukce intervalového odhadu. V minulém díle jsme předpokládali, že pokud mají naše měřená data normální rozdělení (s libovolnými parametry) a my na ně aplikujeme následující transformaci

$$\sqrt{n} \frac{\bar{x}_n - \mu}{S_n}, \quad (1)$$

kde μ je skutečná střední hodnota rozdělení našich dat (tedy skutečná hodnota měřené fyzikální veličiny), potom má takto vzniklá náhodná veličina (je důležité si uvědomit, že jde o náhodnou veličinu, neboť dopředu neznáme její hodnotu) studentovo rozdělení o $n - 1$ stupních volnosti. Toto tvrzení neplatí, pokud měřená data nemají normální rozdělení.

V obecnějším případě, kdy nemají měřená data normální rozdělení, budeme vycházet ze stejné transformace našich naměřených dat, akorát musíme zjistit, jaké bude mít tato náhodná veličina v tomto obecnějším případě rozdělení. O tom hovoří centrální limitní věta. Centrální limitní věta říká, že ať mají naše jednotlivá měření jakékoliv rozdělení (ve skutečnosti je tady podmínka na konečný rozptyl, ale tím se zatěžovat nebudeme), potom bude platit, že takto transformovaná data konvergují v distribuci k normálnímu rozdělení $N(0, 1)$. Matematicky zapsáno

$$\sqrt{n} \frac{\bar{x}_n - \mu}{S_n} \xrightarrow{D} N(0, 1).$$

Co je to ta konvergence v distribuci? Intuitivně řečeno to znamená, že pro velký počet měření bude hustota náhodné veličiny (1) čím dále více podobná hustotě rozdělení $N(0, 1)$. Podobná v tomto případě znamená, že se v každém bodě bude hustota pravděpodobnosti této náhodné veličiny blížit hustotě pravděpodobnosti rozdělení $N(0, 1)$. Pozorného čtenáře by ihned měly napadnout následující dvě otázky:

1. Co je to velký počet měření? Toto je poměrně složitá otázka, ale v základním modelu, který budeme používat, budeme za dostatečně velký počet měření označovat situaci, kdy budeme mít alespoň 30 měření. V takovémto případě už bude aproximace pomocí CLV velice přesná. Pokud budeme mít méně měření, musíme vzít v úvahu fakt, že aproximace normálním rozdělením pomocí CLV nemusí být tak přesná (o této problematice více dále).
2. Není znění centrální limitní věty ve sporu s tím, co jsme probírali v minulém díle seriálu? V minulém díle seriálu jsme uvažovali, že naše pozorování mají normální rozdělení (což spadá do kolonky „jakékoliv rozdělení“), takže by rozdělení transformace námi naměřených dat (1) podle CLV mělo konvergovat k rozdělení $N(0, 1)$, ale v minulém dílu seriálu se tvrdilo, že tato transformace dat má rozdělení t_{n-1} . Vysvětlení je jednoduché, studentovo rozdělení t_n pro velká n konverguje v distribuci k rozdělení $N(0, 1)$, což si můžete ověřit pomocí vykreslení grafů hustot pravděpodobnosti obou rozdělení.

Velice podobným způsobem jako v minulém díle seriálu můžeme potom pracovat s pravděpodobnostmi a odvodit tak intervalový odhad. Jediný rozdíl je, že místo kvantilů studentova t_{n-1} rozdělení budeme muset používat kvantily normálního rozdělení $N(0, 1)$, které budeme značit u (tedy α -kvantil normálního rozdělení $N(0, 1)$ zapíšeme jako u_α). Pro jistotu ještě připomeneme, co jsou to kvantily rozdělení. Pokud si zvolíme číslo α z intervalu $(0, 1)$, potom α -kvantilem normálního rozdělení (značíme ho u_α) bude takové číslo, které splňuje podmínku, že pravděpodobnost, že naše náhodná veličina řídící se rozdělením $N(0, 1)$ bude menší než u_α , bude rovna α . Matematicky zapsáno

$$P(X < u_\alpha) = \alpha.$$

Pro jistotu stručně zopakujeme postup odvození podoby intervalového odhadu pro střední hodnotu (tedy měřenou fyzikální veličinu) v případě obecného rozdělení měřených dat za použití CLV (postup je naprosto stejný jako v minulém díle seriálu).

$$\begin{aligned}
 1 - \alpha &\simeq \text{P} \left(u_{\frac{\alpha}{2}} < \sqrt{n} \frac{\bar{x}_n - \mu}{S_n} < u_{1-\frac{\alpha}{2}} \right) = \\
 &= \text{P} \left(u_{\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}} < \bar{x}_n - \mu < u_{1-\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}} \right) = \\
 &= \text{P} \left(u_{\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}} - \bar{x}_n < -\mu < u_{1-\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}} - \bar{x}_n \right) = \\
 &= \text{P} \left(\bar{x}_n - u_{\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}} > \mu > \bar{x}_n - u_{1-\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}} \right) = \\
 &= \text{P} \left(\bar{x}_n + u_{1-\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}} > \mu > \bar{x}_n - u_{1-\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}} \right) = \\
 &= \text{P} \left(\bar{x}_n - u_{1-\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}} < \mu < \bar{x}_n + u_{1-\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}} \right).
 \end{aligned}$$

Všechny prováděné úpravy byly víceméně triviální, jen je nutné si uvědomit, že můžeme upravovat obě nerovnosti naráz. Mezi čtvrtým a pátým řádkem jsme využili vztahu

$$u_{\frac{\alpha}{2}} = -u_{1-\frac{\alpha}{2}},$$

což plyne ze symetrie hustoty rozdělení $N(0, 1)$ kolem osy y .

Tímto jsme podobně jako v minulém díle seriálu zkonstruovali intervalový odhad (v tomto případě však pouze asymptotický) pro střední hodnotu náhodné veličiny, tedy vlastně pro měřenou fyzikální veličinu ve tvaru

$$1 - \alpha \simeq \text{P} \left(\bar{x}_n - u_{1-\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}} < \mu < \bar{x}_n + u_{1-\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}} \right).$$

Tento intervalový odhad lze interpretovat tak, že pravděpodobnost pokrytí skutečné hodnoty měřené fyzikální veličiny tímto intervalem je pro velký počet měření přibližně $1 - \alpha$ (uvědomme si, že náhodné jsou meze intervalu, nikoliv hodnota měřené fyzikální veličiny). Tento intervalový odhad lze zapsat jako

$$\bar{x}_n \pm u_{1-\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}}.$$

Fyzikové většinou zapisují ještě více zkráceně a vynechají i kvantil $u_{1-\frac{\alpha}{2}}$, tedy intervalový odhad zapíšou jako

$$\bar{x}_n \pm \frac{S_n}{\sqrt{n}},$$

kde se člen $\frac{S_n}{\sqrt{n}}$ nazývá výběrová směrodatná odchylka průměru (odhaduje to směrodatnou odchylku výběrového průměru) a značí se s_n . Někdy se tomuto členu také říká standardní odchylka.

Tento nejvíce zjednodušený zápis má dvě možné interpretace

- Buď se tím myslí, že uvažujeme hladinu spolehlivosti $\alpha \simeq 0,32$ (tedy pravděpodobnost pokrytí je přibližně 0,68), potom totiž platí

$$u_{1-\frac{\alpha}{2}} = 1.$$

- Pokud čtenář chce uvažovat jinou hladinu spolehlivosti, může si podle tabulek doplnit příslušný kvantil a tím intervalový odhad ze zkráceného zápisu zrekonstruovat.

Několik poznámek k intervalovým odhadům založeným na CLV

Zde uvedeme několik důležitých poznámek.

- Jak jsme psali výše, pro velké hodnoty počtu měření n , je studentovo t_n rozdělení téměř identické s rozdělením $N(0, 1)$. Pro velký počet měření je tedy v podstatě jedno, jestli používáme kvantily rozdělení $N(0, 1)$ nebo kvantily studentova t_n rozdělení. To je v souladu s intuitivní představou, že by obě metody (ta uvedená v minulém díle seriálu a ta uvedená dnes) měly pro velký počet měření dávat stejné (nebo alespoň podobné) výsledky.
- Někdy se postupuje tak, že se místo kvantilů normálního $N(0, 1)$ rozdělení uvažují kvantily studentova t_n rozdělení. Tedy zpracováváme naměřená data jakoby měřené hodnoty měli normální rozdělení (viz minulý díl seriálu), ačkoliv víme, že tomu tak není nebo být nemusí. Toto není nutně chybný postup, neboť studentovo t_n rozdělení (pro libovolné $n \in \mathbb{N}$) je více zploštělé než rozdělení $N(0, 1)$ (ověřte si sami vykreslením grafů hustot pravděpodobnosti), tedy jeho kvantily budou vždy v absolutní hodnotě větší než příslušné kvantily (tj. kvantily pro stejnou hladinu spolehlivosti α) normálního rozdělení $N(0, 1)$, tedy budeme tímto způsobem vždy dostávat širší intervaly spolehlivosti. Budeme se tedy chovat více konzervativně tj. výslednou nejistotu měření si tímto způsobem vždy zvětšíme, nikdy si ji nemůžeme zmenšit. Jak už bylo zmíněno v předchozí poznámce, pro velký počet měření tento postup dává téměř identické výsledky jako postup založený na kvantilech $N(0, 1)$ rozdělení. V řešení úloh ovšem používejte klasický postup (tedy ten založený na kvantilech $N(0, 1)$ rozdělení).
- Centrální limitní větu lze použít jen pro dostatečně velký počet měření. Pokud máme alespoň 30 měření, můžeme CLV bez obav použít, aproximace bude už velmi přesná. Pokud máme alespoň 10 měření, můžeme CLV použít, ale musíme mít na paměti, že aproximace normálním rozdělením nebude dokonalá a příslušné intervalové odhady budou jen přibližné. Pokud máme méně než 10 měření, aproximace normálním rozdělením může být už značně nepřesná. Zároveň ale platí, že nic lepšího zkonstruovat nelze. Proto se výsledek většinou uvádí stejný, ale je nutné do diskuze přidat upozornění, že může jít o značně nepřesný výsledek, neboť je založen na malém počtu měření ¹!
- V minulém dílu seriálu jsme popisovali nejistoty typu A (pocházející z náhodnosti měřených dat) a nejistoty typu B (pocházející z ostatních příčin, typicky nepřesnosti použitého měřidla). Nyní uvedeme, že výběrová směrodatná odchylka průměru (neboli standardní

¹Pokud jsme si jisti, že námi měřená data mají normální rozdělení – což prakticky nikdy nejsme – potom je intervalový odhad založený na studentově t_n rozdělení (tj. bez použití CLV) přesný pro libovolný počet měření (viz minulý díl seriálu).

odchylka) představuje nejistotu typu A a pokud pracujeme ještě s nejistotou typu B, potom ji k nejistotě typu A přičteme pomocí stejného vzorce

$$s_K = \sqrt{s_A^2 + s_B^2},$$

čímž získáme kombinovanou nejistotu měření. Intervalové odhady potom konstruujeme za použití této kombinované nejistoty měření.

- Cílem fyzikálního experimentu je mít co nejmenší kombinovanou nejistotu měření (tj. naměřit fyzikální veličinu co možná nejpřesněji). Podobně jako v případě normálně rozdělených dat toho lze dosáhnout buď větším počtem měření (tj. zvětšit člen \sqrt{n} ve jmenovateli zlomku, tedy zmenšit nejistotu typu A vyjádřenou pomocí s_n), nebo přesnější měřicí aparaturou (tj. snížit rozptyl měřených dat, tedy snížit člen S_n , který představuje odhad rozptylu).

Vícerozměrný případ

Nyní se zaměříme na případ, kdy chceme změřit více fyzikálních veličin (které budou samozřejmě zatíženy nejistotou) a tyto změřené veličiny následně dosadit do určitého vzorce. Náš úkol bude zkoumat vlastnosti výsledku dosazení (zejména budeme chtít umět vyjádřit nejistoty určení a konstruovat intervalové odhady).

Uvažujme, že cílem našeho experimentu je změřit fyzikální veličiny $v^{(1)}, \dots, v^{(k)}$ a následně je dosadit do vzorce

$$v = f(v^{(1)}, \dots, v^{(k)}),$$

čímž získáme výslednou fyzikální veličinu v . Odvodíme si nyní, jaká bude nejistota určení veličiny v a jak pro ni konstruovat intervalové odhady.

Na tomto místě si musíme uvědomit, že v případě měření více fyzikálních veličin najednou nastává problém se závislostí jednotlivých měření. V úplně nejjednodušším modelu můžeme uvažovat, že jsou všechna prováděná měření na sobě nezávislá, což bude obvykle opodstatněný předpoklad. V nějakých případech ovšem takto jednoduchý model použít nelze a budeme muset do určování výsledné nejistoty a konstrukce intervalových odhadů také zahrnout možnou závislost našich měření. Nyní se budeme věnovat odvození složitějšího modelu se závislými měřeními a jednodušší model pro nezávislá měření dostaneme jako jeho speciální případ. Jako první si musíme zadefinovat, jak budeme matematicky popisovat závislost našich měření (jelikož výsledek jednoho měření modelujeme jako náhodnou veličinu, budeme vlastně popisovat závislosti náhodných veličin).

Kovariance a korelace

Za účelem měření míry závislosti dvou náhodných veličin se zavádí kovariance (též koeficient kovariance, kovarianční koeficient), která je pro dvě náhodné veličiny X a Y definována jako

$$\text{cov}(X, Y) = E[(X - EX)(Y - EY)],$$

kde E značí střední hodnotu. Koeficient kovariance může obecně nabývat libovolné reálné hodnoty. Intuitivně si lze kovarianci představovat tak, že vyjadřuje střední hodnotu výrazu

$$(X - EX)(Y - EY),$$

který bude nabývat nejčastěji kladných hodnot (potom jeho střední hodnota potom bude kladné číslo), pokud na sobě budou náhodné veličiny závislé takovým způsobem, že pokud jedna náhodná veličina nabývá vysoké hodnoty (tj. vyšší než je její střední hodnota), potom i druhá náhodná veličina bude s největší pravděpodobností nabývat vysoké hodnoty (tj. vyšší než její střední hodnota). Podobně bude tento výraz nabývat nejčastěji záporných hodnot (potom jeho střední hodnota bude záporné číslo), pokud bude závislost našich náhodných veličin taková, že vysoké hodnoty jedné veličiny (tj. vyšší než její střední hodnota), se budou s největší pravděpodobností vyskytovat právě v případě, že druhá náhodná veličina bude nabývat nízkých hodnot (tj. nižších než její střední hodnota). Také musíme uvést fakt, že pokud budou náhodné veličiny X a Y nezávislé, potom bude hodnota jejich kovariančního koeficientu vždy 0. To lze odvodit přímo z definice následovně

$$\begin{aligned} \text{cov}(X, Y) &= E[(X - EX)(Y - EY)] = E[X - EX] \cdot E[Y - EY] = \\ &= (EX - EX) \cdot (EY - EY) = 0. \end{aligned}$$

Úpravy, které jsme prováděli, můžeme provést jen za předpokladu, že náhodné veličiny X a Y jsou nezávislé, jinak takto postupovat nemůžeme.

Pro lepší popis závislosti náhodných veličin se zavádí ještě korelace (též koeficient korelace, korelační koeficient), která je definovaná jako

$$\text{corr}(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X) \cdot \text{var}(Y)}}.$$

Díky tomu, že koeficient korelace je definován jako vhodně normovaný koeficient kovariance, může nabývat pouze hodnot od -1 do 1 . Pro ideálně lineárně závislé veličiny (tj. pokud platí $X = aY + b$, pro nějaké konstanty a, b) bude korelační koeficient nabývat hodnot ± 1 (bude mít stejné znaménko jako konstanta a). Pokud budou náhodné veličiny X a Y nezávislé, potom bude jejich korelační koeficient nulový (triviální důsledek definice korelačního koeficientu a vlastnosti kovariance, kterou jsme odvodili dříve). Hodnoty korelačního koeficientu blízko nuly ukazují na malou závislost náhodných veličin a hodnoty korelačního koeficientu blízko 1 nebo -1 ukazují na vysokou míru závislosti náhodných veličin.

Kovarianční a korelační koeficient našich naměřených dat v praxi většinou není dopředu znám (podobně jako střední hodnota nebo rozptyl). Pokud si můžeme být jisti, že jsou všechna naše měření fyzikálních veličin nezávislá, potom můžeme tvrdit, že je příslušný kovarianční a korelační koeficient nulový. Tuto jistotu získáme jen díky znalosti toho, jak jsme naše data měřili. V některých případech se stane, že si nemůžeme být jisti nezávislostí našich měření a musíme tedy z naměřených dat odhadovat míry závislosti.

Pokud máme k dispozici měření dvou fyzikálních veličin (vlastně realizace dvou náhodných veličin), můžeme odhadovat jejich kovarianční a korelační koeficienty. Potřebujeme k tomu vždy dvojice měření $(x_1, y_1), \dots, (x_n, y_n)$, kde x_i je i -té měření první fyzikální veličiny a y_i je i -té měření druhé fyzikální veličiny. Zde je nutné poznamenat, že nelze postupovat tak, že si naměřené hodnoty libovolně spárujeme, dvojice hodnot (x_i, y_i) musí vždy pocházet z odpovídajících si měření! Pokud máme takovéto dvojice odpovídajících si měření, můžeme kovarianční koeficient odhadnout pomocí tzv. výběrového kovariančního koeficientu koeficientu definovaného jako

$$\widehat{\text{cov}}(X, Y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n) (y_i - \bar{y}_n),$$

kde $\overline{x_n}, \overline{y_n}$ mají význam výběrových průměrů. Podobně lze odhadovat korelační koeficient pomocí výběrového korelačního koeficientu definovaného jako

$$\widehat{\text{corr}}(X, Y) = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \overline{x_n})(y_i - \overline{y_n})}{\sqrt{S_{X,n}^2 \cdot S_{Y,n}^2}},$$

kde $S_{X,n}^2$ je výběrový rozptyl zkonstruovaný z realizací náhodné veličiny X , $S_{Y,n}^2$ je výběrový rozptyl zkonstruovaný z realizací náhodné veličiny Y a členy $\overline{x_n}$ a $\overline{y_n}$ mají význam výběrových průměrů.

Na tomto místě si opět musíme rozdíel mezi odhadem kovariančního (resp. korelačního) koeficientu a jeho skutečnou hodnotou (kterou většinou neznáme). Skutečná hodnota kovariančního (resp. korelačního) koeficientu je konstanta (není na ní nic náhodného), zatímco odhad kovariančního (resp. korelačního) koeficientu (tj. výběrový kovarianční a korelační koeficient) je náhodná veličina (protože je zkonstruován na základě naměřených dat, která jsou náhodná, nemůžeme dopředu určit jeho hodnotu). Pro velký počet měření bude mít výběrový kovarianční (resp. korelační) koeficient podobně jako předchozí odhady tu vlastnost, že bude s největší pravděpodobností velice blízko skutečným hodnotám kovariančního (resp. korelačního) koeficientu. Nyní už víme, jak matematicky měřit závislosti dvou náhodných veličin a můžeme přistoupit k formulaci vícerozměrné centrální limitní věty.

Vícerozměrná verze CLV

Nyní už můžeme uvést znění vícerozměrné centrální limitní věty². Uvažujme tedy situaci jako výše, tedy že měříme k různých fyzikálních veličin $v^{(1)}, \dots, v^{(k)}$, které naměříme s určitou nejistotou. Máme tedy naměřeno

$$\left(\overline{v^{(1)}} \pm s_{n_1}^{(1)} \right), \dots, \left(\overline{v^{(k)}} \pm s_{n_k}^{(k)} \right),$$

kde $\overline{v^{(i)}}$ je výběrový průměr i -té fyzikální veličiny a $s_{n_i}^{(i)}$ je standardní odchylka i -té fyzikální veličiny tak, jak byly definovány výše (uvažujeme tedy, že každá fyzikální veličina mohla být měřena různým počtem měření, a uvažujeme také, že rozdělení měřených dat nemusí být normální a ani měření různých fyzikálních veličin nemusí být vzájemně nezávislé). Tyto naměřené veličiny následně chceme dosadit do vzorce

$$v = f(v^{(1)}, \dots, v^{(k)}) \quad (2)$$

a určit výslednou nejistotu a zkonstruovat intervalové odhady.

V tomto případě využijeme vícerozměrnou verzi centrální limitní věty, která tvrdí, že platí

$$\frac{f\left(\overline{v_n^{(1)}}, \dots, \overline{v_n^{(k)}}\right) - f\left(v^{(1)}, \dots, v^{(k)}\right)}{\sqrt{S^2}} \xrightarrow{D} N(0, 1),$$

²Uvedeme ji zde v trochu neobvyklém tvaru, který je ale velice užitečný pro výpočty. Pokud byste si o tomto tématu četli i v jiné literatuře, je pravděpodobné, že tam bude CLV pro vícerozměrný případ formulována jinak, nicméně tyto dvě formulace jsou ekvivalentní.

kde $v^{(i)}$ je skutečná hodnota i -té fyzikální veličiny a S^2 je v tomto případě odhad rozptylu náhodné veličiny představující celkový výsledek měření³, který se spočte podle vzorce⁴

$$S^2 = \left(\frac{\partial f}{\partial v^{(1)}}(\bar{v}) \quad \dots \quad \frac{\partial f}{\partial v^{(k)}}(\bar{v}) \right) \begin{pmatrix} S_{n_1}^2 & \dots & \widehat{\text{cov}}(v^{(1)}, v^{(k)}) \\ \frac{\widehat{\text{cov}}(v^{(2)}, v^{(1)})}{\sqrt{n_2 n_1}} & \dots & \frac{\widehat{\text{cov}}(v^{(2)}, v^{(k)})}{\sqrt{n_2 n_k}} \\ \vdots & \ddots & \vdots \\ \frac{\widehat{\text{cov}}(v^{(k)}, v^{(1)})}{\sqrt{n_k n_1}} & \dots & S_{n_k}^2 \end{pmatrix} \begin{pmatrix} \frac{\partial f}{\partial v^{(1)}}(\bar{v}) \\ \vdots \\ \frac{\partial f}{\partial v^{(k)}}(\bar{v}) \end{pmatrix}. \quad (3)$$

Poznamenejme na tomto místě, že výrazem

$$\frac{\partial f}{\partial v^{(i)}}(\bar{v})$$

se myslí hodnota parciální derivace funkce f podle i -té proměnné vyčíslená v bodě

$$\bar{v} = \left(\overline{v_{n_1}^{(1)}}, \dots, \overline{v_{n_k}^{(k)}} \right),$$

tedy jde o číslo.

Klasickým způsobem, který jsme zde už několikrát viděli, můžeme zkonstruovat asymptotický intervalový odhad pro

$$v = f(v^{(1)}, \dots, v^{(k)}).$$

Uvedeme zde jen hlavní kroky odvození⁵

$$\begin{aligned} 1 - \alpha &\simeq \text{P} \left(u_{\frac{\alpha}{2}} < \frac{f(\overline{v_n^{(1)}}), \dots, \overline{v_n^{(k)}}) - f(v^{(1)}, \dots, v^{(k)})}{\sqrt{S^2}} < u_{1 - \frac{\alpha}{2}} \right) = \\ &= \text{P} \left(-f(\overline{v_n^{(1)}}), \dots, \overline{v_n^{(k)}}) + u_{\frac{\alpha}{2}} S < -f(v^{(1)}, \dots, v^{(k)}) < -f(\overline{v_n^{(1)}}), \dots, \overline{v_n^{(k)}}) + u_{1 - \frac{\alpha}{2}} S \right) = \\ &= \text{P} \left(f(\overline{v_n^{(1)}}), \dots, \overline{v_n^{(k)}}) - u_{1 - \frac{\alpha}{2}} S < f(v^{(1)}, \dots, v^{(k)}) < f(\overline{v_n^{(1)}}), \dots, \overline{v_n^{(k)}}) + u_{1 - \frac{\alpha}{2}} S \right). \end{aligned}$$

V nejvíce zkrácené podobě se tento intervalový odhad bude zapisovat stejně jako v předchozích případech pouze jako

$$f(\overline{v_n^{(1)}}), \dots, \overline{v_n^{(k)}}) \pm S,$$

kde S se označuje jako standardní odchylka.

³Tj. dosazení odhadů našich fyzikálních veličin do vzorce (2).

⁴Tento vzorec využívá parciální derivace a maticové násobení. Pokud toto neumíte, můžete se podívat na poslední kapitulu tohoto dílu seriálu, kde je vše stručně vysvětleno, nebo můžete pokračovat dále, neboť jak uvidíme, v praxi nejčastěji používaných případech se tento vzorec značně zjednoduší.

⁵Používají se jen algebraické úpravy nerovností a nahrazení $u_{\frac{\alpha}{2}}$ pomocí $-u_{1 - \frac{\alpha}{2}}$.

Několik poznámek k intervalovým odhadům založeným na vícerozměrné CLV

Zde uvedeme několik důležitých poznámek.

- V odvození jsme uvažovali vícerozměrný případ, tedy obecně $k \in \mathbb{N}$. Celou odvozenou teorii lze samozřejmě aplikovat i na speciální případ, kdy zvolíme $k = 1$ (tedy jen jednorozměrný případ). Toto se při zpracování fyzikálních experimentů vyskytuje poměrně často (měříme jen jednu veličinu a chceme ji dosadit do určitého fyzikálního vzorce).
- Vícerozměrnou centrální limitní větu lze použít jen pro dostatečně velký počet měření. Pokud jsme provedli více než 30 měření každé fyzikální veličiny, potom už bude aproximace pomocí vícerozměrné CLV velmi přesná. Pokud jsme provedli alespoň 10 měření každé fyzikální veličiny aproximace pomocí vícerozměrné CLV nebude tolik spolehlivá, ale stále bude poměrně přesná. Pro menší počet měření nějaké veličiny než 10 platí, že použití vícerozměrné CLV už může dávat značně nepřesné výsledky. Zároveň ale platí, že žádný lepší postup než výše uvedený neexistuje. Proto se i pro malý počet měření (menší než 10) výsledek většinou uvádí stejný, ale je nutné přidat upozornění, že může jít o značně nepřesný výsledek⁶!
- V rámci typů nejistot měření můžeme standardní odchylku S považovat za kombinovanou nejistotu měření, pokud jsme jako $s_{n_i}^{(i)}$ používali kombinované nejistoty měření, nebo ji můžeme považovat pouze za nejistotu typu A, pokud jsme jako $s_{n_i}^{(i)}$ používali pouze nejistoty typu A.

Zákon šíření nejistot

V nejobecnějším modelu (kdy uvažujeme závislá měření různých fyzikálních veličin) nemůžeme vzorec (3) nijak zjednodušit a musíme opravdu provést celý takto složitý výpočet. V určitých speciálních případech se ovšem tento vzorec značně zjednoduší.

Ve speciálním případě, kdy budeme uvažovat, že jsou všechna měření na sobě nezávislá (tj. nebudeme ani odhadovat kovariance měření různých fyzikálních veličin, ale rovnou budeme předpokládat, že jsou nulové), se nám vzorec (3) zjednoduší na tvar

$$S = \sqrt{\left(\frac{\partial f(\bar{v})}{\partial v^{(1)}}\right)^2 s_{n_1}^{(1)2} + \dots + \left(\frac{\partial f(\bar{v})}{\partial v^{(k)}}\right)^2 s_{n_k}^{(k)2}}.$$

Tomuto vzorci se také někdy říká zákon šíření nejistot (případně zákon propagace nejistot).

V praxi se vyplatí znát přesnou podobu zákona šíření nejistot pro nejčastěji se vyskytující funkce, které můžeme vidět v tabulce 1. Pilný čtenář si může platnost všech uvedených vzorců ověřit dosazením příslušných funkcí do zákona šíření nejistot.

Občas se lze v praxi setkat také s pojmem relativní nejistota měření, která je definována jako

$$\delta_x = \frac{\bar{x}_n}{s_n}.$$

Je důležité poznamenat, že relativní nejistota měření je bezrozměrná veličina (na rozdíl od klasické nejistoty). Přejít od relativní nejistoty měření ke klasické nejistotě měření si jistě

⁶Pokud bychom věděli, že námi měřená data mají normální rozdělení, lze odvodit přesný postup (tj. bez jakékoli aproximace) analogický jednorozměrnému případu. Tento postup je ovšem příliš složitý a nebudeme ho zde uvádět (v praxi se stejně téměř nepoužívá).

Tab. 1: Nejdůležitější příklady použití zákona šíření nejistot.

Použitá funkce	Vzorec na výslednou nejistotu měření
$f(v^{(1)}, v^{(2)}) = v^{(1)} + v^{(2)}$	$S = \sqrt{s_{n_1}^{(1)2} + s_{n_2}^{(2)2}}$
$f(v^{(1)}, v^{(2)}) = v^{(1)} - v^{(2)}$	$S = \sqrt{s_{n_1}^{(1)2} + s_{n_2}^{(2)2}}$
$f(v^{(1)}, v^{(2)}) = v^{(1)} \cdot v^{(2)}$	$S = \sqrt{\frac{v_{n_2}^{(2)2}}{v_{n_1}^{(1)2}} s_{n_1}^{(1)2} + \frac{v_{n_1}^{(1)2}}{v_{n_2}^{(2)2}} s_{n_2}^{(2)2}}$
$f(v^{(1)}, v^{(2)}) = \frac{v^{(1)}}{v^{(2)}}$	$S = \sqrt{\frac{1}{v_{n_2}^{(2)2}} s_{n_1}^{(1)2} + \frac{v_{n_1}^{(1)2}}{v_{n_2}^{(2)4}} s_{n_2}^{(2)2}}$
$f(v) = v^a, a \in \mathbb{R}, a \neq 0$	$S = a \cdot \overline{v_n}^{a-1} s_n$

už každý čtenář dokáže odvodit sám. Celá teorie by se dala podobným způsobem vybudovat, pokud bychom používali relativní nejistoty měření, ale až na výjimečné příklady⁷ by to nevedlo k jednodušším výsledkům. V dalším pokračování seriálů i v řešení seriálových úloh se nebudeme relativními nejistotami měření dále zabývat, zde je uvádíme jen jako drobné rozšíření vykládané látky.

Parciální derivace a násobení matic (vektorů)

Jelikož se ve vzorcích (3) vyskytly parciální derivace a násobení matic, což možná není všem úplně známé, uvedeme zde stručné objasnění použitých pojmů.

V tomto díle seriálu bohužel nemůžeme podrobně vysvětlit, jak je definována derivace (resp. parciální derivace) a jaké má vlastnosti. Proto pouze uvedeme, že parciální derivace funkce k proměnných podle i -té proměnné se spočítá jako normální derivace funkce, když budeme na všechny proměnné kromě i -té pohlížet jako na konstanty.

Matice jsou v podstatě jen tabulky čísel, nic více. U matice rozlišujeme její rozměry (tj. kolik má řádků a sloupců), říkáme, že matice A je typu k krát n (píšeme $A_{k \times n}$ – první číslo udává počet řádků, druhé počet sloupců). Matice typu 1×1 je pouze číslo a matici typu $1 \times n$ (resp. $n \times 1$) se říká řádkový (resp. sloupcový) vektor.

Násobení matic je definováno poněkud složitěji. Uvažujme, že máme matice $A_{k \times n}$ a $B_{n \times l}$ (aby byl součin těchto matic definován musí být počet sloupců první matice roven počtu řádků

⁷Pokud bychom formulovali zákon šíření nejistot v řeči relativních nejistot, potom by například pro volbu funkce $f(u, v) = u \cdot v$ výsledná relativní nejistota měření byla jen kvadratickým součtem relativních nejistot měření veličin u a v . Ve většině ostatních případů by ale používání relativních nejistot vedlo spíše ke složitějším výsledkům.

druhé matice), potom součin AB je definován následovně⁸

$$\begin{aligned}
 A \cdot B &= \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k,1} & a_{k,2} & \cdots & a_{k,n} \end{pmatrix} \cdot \begin{pmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,l} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,l} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n,1} & b_{n,2} & \cdots & b_{n,l} \end{pmatrix} = \\
 &= \begin{pmatrix} \langle a_{1,\bullet}, b_{\bullet,1} \rangle & \langle a_{1,\bullet}, b_{\bullet,2} \rangle & \cdots & \langle a_{1,\bullet}, b_{\bullet,l} \rangle \\ \langle a_{2,\bullet}, b_{\bullet,1} \rangle & \langle a_{2,\bullet}, b_{\bullet,2} \rangle & \cdots & \langle a_{2,\bullet}, b_{\bullet,l} \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle a_{k,\bullet}, b_{\bullet,1} \rangle & \langle a_{k,\bullet}, b_{\bullet,2} \rangle & \cdots & \langle a_{k,\bullet}, b_{\bullet,l} \rangle \end{pmatrix},
 \end{aligned}$$

kde $a_{i,\bullet}$ označuje i -tý řádek matice A , $b_{\bullet,j}$ označuje j -tý sloupec matice B a výraz $\langle a_{i,\bullet}, b_{\bullet,j} \rangle$ označuje skalární součin vektorů $a_{i,\bullet}$ a $b_{\bullet,j}$ definovaný jako

$$\langle a_{i,\bullet}, b_{\bullet,j} \rangle = \sum_{g=1}^n a_{i,g} \cdot b_{g,j}.$$

Výsledkem součinu matic $A_{k \times n}$ a $B_{n \times l}$ je tedy matice tvaru $C_{k \times l}$. Pokud chceme násobit vektor s maticí (nebo 2 vektory spolu) stačí na vektory pohlížet jako na speciální případ matice a řídit se pravidly pro maticové násobení.

Fyzikální korespondenční seminář je organizován studenty MFF UK. Je zastřešen Oddělením pro vnější vztahy a propagaci MFF UK a podporován Ústavem teoretické fyziky MFF UK, jeho zaměstnanci a Jednotou českých matematiků a fyziků.

Toto dílo je šířeno pod licencí Creative Commons Attribution-Share Alike 3.0 Unported.
Pro zobrazení kopie této licence navštivte <http://creativecommons.org/licenses/by-sa/3.0/>.

⁸Je důležité poznamenat, že maticové násobení není komutativní, tedy obecně platí $A \cdot B \neq B \cdot A$, ale je asociativní, tedy platí $A \cdot (B \cdot C) = (A \cdot B) \cdot C$ pro libovolné matice A, B, C , pro které je takový součin definován.