Koncepciók a metrológiában

A mérési bizonytalanság fogalmának fejlődése

A hibáktól a valószínűségsűrűség- függvényekig

WALTER BICH, INSTITUTO NAZIONALE DI RICERCA METROLOGICA (INRIM), TORINO, OLASZORSZÁG

WILLEM KOOL, BUREAU INTERNATIONAL DE METROLOGIE LEGALE (BIML), PÁRIZS, FRANCIAORSZÁG

Mérés, tulajdonságok és mennyiségek

A Nemzetközi Metrológiai Értelmező Szótár (VIM) [1] legutóbbi kiadása a következőképpen határozza meg a *mérést*: „egy vagy több olyan mennyiségérték kísérleti megszerzésének folyamata, amelyek indokoltan tulajdoníthatók egy mennyiségnek", a *mérendő mennyiséget* pedig így, „az a mennyiség, amelyet mérni szándékoznak". A mennyiség meghatározása a következő: „jelenség, test vagy anyag olyan tulajdonsága, amelynek egy számmal és egy hivatkozással kifejezhető nagysága van". A *tulajdonság* és a *hivatkozás* fogalma nincs meghatározva, jóllehet a hivatkozást a mennyiség meghatározásához fűzött megjegyzés értelmezi. Az értelmezés így hangzik: „A hivatkozás lehet egy mértékegység, egy mérési eljárás, egy anyagminta, vagy ilyenek társítása".

A legtöbb metrológus a mennyiséget úgy képzeli el, mint valami olyat, aminek van egy mértékegysége. A *mértékegység* a VIM szerint „megállapodással meghatározott és elfogadott valós skalár mennyiség, amellyel bármely más, vele azonos fajtájú mennyiség összehasonlítható úgy, hogy a két mennyiség aránya egy számmal legyen kifejezhető." Ez a meghatározás nem alkalmas olyan tulajdonságokra, mint a keménység vagy a szín. A keménységet és sok más tulajdonságot azonban „mérhetőnek" tekintenek abban az értelemben, hogy egy megállapodásos eljárást (a referenciát) követve, pontos (nagyság szerinti) sorrendi relációt lehet felállítani. Az ilyen mennyiségeket *sorrendi mennyiségeknek* nevezik. A sorrendi mennyiségek különbségeinek és hányadosainak nincs fizikai értelmezése.

A sorrendi mennyiségeknek ez a tulajdonsága nem akadályozza meg a nemzeti metrológiai intézeteket (NMI-ket) abban, hogy összehasonlítsák keménységmérési képességeiket. Az összehasonlítás általában egy átlag-jellegű referenciaérték meghatározását, és az összehasonlításban részt vevők eredményeinek ettől a referenciaértéktől való eltéréseit foglalja magába. Ezen kívül a részt vevő NMI-k és a vezetőlaboratórium rutinszerűen kiszámítják a szórásokat abból a célból, hogy ki tudják számítani az egyes egyedi eredményekhez és a referenciaértékhez társított bizonytalanságokat.

Itt kell megjegyezni, hogy ezeket az algebrai műveleteket ugyanannak a mérendő mennyiségnek a mérési eredményeire alkalmazzák. Míg világos, hogy ugyanazon tárgy keménységének két mérési eredménye eltérhet, mondjuk 2 HRC-vel, egymástól, annak az állításnak, hogy egy kés éle 2 HRC-vel keményebb, mint egy másik késé, nincs hasonlóan kétségtelen mennyiségi jelentése.

Az átlagokat és a szórásokat illető figyelmeztetést már S. S. Stevens pszichológus is megfogalmazott 1946-ban, úttörő tanulmányában [2], amelyben megalapozta a „mérések skáláinak” ma is széles körben alkalmazott osztályozását.

A korszerű nyelvhasználatban (és a VIM-ben) a *tulajdonság* magasabb rendű fogalom, amelynek a *mennyiség* egy alárendelt fogalma. A mennyiségek olyan tulajdonságok egy alkészletét képezik, amelyeknek a Stevens-féle osztályozásnak megfelelő arány-, intervallum- és sorrendi skálák használatával „mérhető” nagysága van. Stevens „abszolút skálája” az arány skála egy speciális esetének tekinthető, amelynél az „egység” az „1” (számlálás). Stevens osztályozását kiegészíti a „névleges skála” (nominal scale), ahol az osztályok (számok vagy nevek: például „Alan”, „John”, „Pierre”) sorrendje magának a tulajdonságnak (ebben a példában a „személynévnek”) egyetlen saját jellemzőjéhez sem kapcsolódik. Bár a személyek csoportjának egyedi nevei ábécé-sorrendben is előállíthatók, ez a sorrend nem képvisel semmiféle „nagyságot” vagy relatív fontosságot. A „személynév” tulajdonság tehát „névleges tulajdonság” (nominal property).

A *mérés* meghatározása a VIM-ben egy megjegyzéssel egészül ki: „A mérés névleges tulajdonságokra nem alkalmazható,.” azt hangsúlyozva ezzel, hogy a mérés elvégezhetőségéhez a *mérendő mennyiségnek* valóban *mennyiségnek* kell lennie, vagyis olyan tulajdonságnak, amelynek nagysága van.

Mennyiségértékek és valódi érték

A klasszikus fizikában egy fizikai rendszerben lévő mennyiségről hagyományosan feltételezik, hogy annak egyedi értéke van. Az egyedi értéknek a jellemzésére a „valódi”jelzőt fogadták el. Ennek a jelzőnek a használata már régóta vita tárgya.

A VIM második kiadásában már megjelent az elképzelés, hogy egy mennyiségnek több valódi értéke is lehet, attól függően, hogy a mennyiség mennyire jól van definiálva. Ez a felfogás még erőteljesebben megnyilvánul a VIM jelenlegi, harmadik kiadásában, amely szerint egy mérendő mennyiség esetében a valódi értékek egész készlete létezik. Ezt a készletet a *meghatározási (definiálási) bizonytalanság* számlájára lehet írni, amelynek a meghatározása a következő: „ A mérési bizonytalanságnak az az összetevője, amely abból ered, hogy a mérendő mennyiség meghatározásának részletezése véges”. A definiálási bizonytalanság fogalmának megvannak az előnyei, de nehézséget is okoz nemcsak a bizonytalanságértékelés elvi sémájában, hanem a gyakorlati mérések szempontjából is. Szerencsére a VIM-ben a *mennyiség valódi értékének* meghatározása egy megjegyzéssel egészül ki: „Ha a mérendő mennyiséghez társított definiálási bizonytalanság a mérési bizonytalanság többi összetevőjéhez képest elhanyagolható, akkor a mérendő mennyiség úgy tekinthető, hogy „lényegében véve egyetlen” valódi mennyiségértéke van. Ez a megközelítés található a *Guide to the expression of uncertainty in measurement*-ben (GUM) [3] és a kapcsolódó dokumentumokban, amelyekben a „valódi” szót redundánsnak tekintik.

Számos olyan példát lehet találni, amelyekben a mérendő mennyiség nincs túl jól meghatározva. Például a „szén atomsúlya” változik a minta izotópösszetételétől függően, és a „Párizs és Róma közötti távolság” is függ attól, hogy a két városban milyen vonatkoztatási pontot választanak ki a méréshez.

A GUM megállapítja: „Ez az útmutató elsősorban olyan jól definiált fizikai mennyiségek - a mérendő mennyiségek - mérési bizonytalanságának megadásával foglalkozik, amelyek lényegében véve egyetlen értékkel jellemezhetők. Ha az érdeklődésre számot tartó jelenség csak az értékek egy eloszlásával írható le, vagy ha egy vagy több olyan paramétertől függ, mint amilyen az idő, akkor a mérendő mennyiség leírásához az eloszlást vagy a függést leíró egész készletére van szükség.” Más szóval, a GUM elismeri a „meghatározási (definiálási) bizonytalanságot”, de nem ad meg sem pontosan kifejezett értékelési módot, sem speciális példákat. A GUM alkalmazásához jól definiált mérendő mennyiségre van szükség.

Hiba és bizonytalanság

A mérés elvégzését történetileg olyan feladatnak tekintették, amelynek célja a mérendő mennyiség (ismeretlen) „valódi” értékéhez a lehető legközelebb levő számérték megtalálása. A két érték különbsége a *mérési hiba*. Minthogy azonban a „valódi” érték ismeretlen, a mérési hiba meghatározásában ’valódi mennyiségérték’ helyett a *referencia mennyiségértéket* használják.

A következő lépés az volt, hogy megkíséreljék mennyiségileg meghatározni (kvantifikálni) a mérési eredménynek a „valódi” értékhez való közeliségét. Ebben a tekintetben az a probléma, hogy ha a mérési hiba akár csak közelítőleg ismert lenne, akkor össze lehetne kapcsolni a mérési eredménnyel ahhoz, hogy egy jobb értéket kapjanak. Ez a nyilvánvalóan naív felfogás azonban egyike a modern bizonytalanságról kialakított sarkalatos elgondolásoknak, és ezért mélyebben meg kell vitatni. Mindenesetre, mivel a valódi hiba ismeretlen, a mérési eredmény valódi értékhez való közeliségének egyetlen lehetséges mértéke valószínűségi természetű.

A mai uralkodó nézetek szerint a mérés célja a mérendő mennyiségre vonatkozó ismertállapot javítása. A bizonytalanság úgy tekinthető, mint ennek az ismeretállapotnak az ellentéte (reciproka). Minél jobban ismerjük a mérendő mennyiséget, annál kisebb a bizonytalanság. Ahogyan a mennyiség értékére vonatkozó ismeretállapot javul, mind szűkebbé válik az a tartomány, amelyről úgy hisszük, hogy a mennyiség értéke benne fekszik.

Közvetett mérések

A legtöbb mérés közvetett, ami azt jelenti, hogy a mérendő mennyiséget nem közvetlenül észleljük, hanem más, közvetlenül mért vagy észlelt mennyiségek függvényeként. A mérési hiba így a közvetlenül mért/észlelt mennyiségek hibáinak függvénye lesz. A GUM-ban alkalmazott közelítésmód egy valószínűségi mértéket (a szórást) tulajdonít minden bizonytalanság-összetevőnek, és az összetevők szórásait kombinálja annak érdekében, hogy megkapja a mérési eredmény bizonytalanságát.

Véletlen és rendszeres hibák

A GUM-ban elfogadott közelítésmód jól működik a véletlen hibák esetében, és a használatában nincsen semmi új. Probléma a rendszeres hibáknál lép fel, amelyekre nincs ismert terjedési szabály. Még ha lenne is ilyen szabály, gondot jelentene a kétféle típusú hiba kombinálása. Ennek a problémának a megoldására a GUM egyszerű meggondolás alapján tesz javaslatot. Ha egy rendszeres hiba azonosítható, akkor az értékének is becsülhetőnek kell lennie. Ez a becslés az úgynevezett „torzítás” (bias).

A szokásos eljárás az volt (és jelenleg is az), hogy a torzítást beveszik a bizonytalansáilistába.
A torzítás azonban nem bizonytalanság-összetevő, hanem egy járulék a mérési eredmény becsléséhez, és azt megfelelő előjellel, mint korrekciót, be kell venni a mérési modellbe. A

torzítások azonban néha „okos becslések” (educated guess) alakját öltik, amelyeknek szintén vannak társított bizonytalanságai, és ezeket úgy kell kezelni, mint a véletlen hibák bizonytalanságait. A GUM szerint az „okos becslések” semmivel sem alábbvalók, mint az „objektív becslések”, értékük egyetlen mértéke egyaránt a bizonytalanság.

Mivel a véletlen hibákat egy sokaságnak tekintik (tipikusan Gauss-félének), a bizonytalanság mértékének megadásához statisztikai eszközök alkalmazhatók. Ez nem tehető meg a torzítási bizonytalanság esetén, mert a rendszeres hiba hátterében nincs semmiféle sokaság.(halmaz). Ha a statisztikai eszközök a rendszeres hibákra nem alkalmazhatók, akkor más alkalmazható valószínűségszámítási eszközöket kell igénybe venni, vagy legalább is a valószínűség számos értelmezésének egyikét. Ezt az értelmezést általában bayesi-nek[1] nevezik, a „frekvenciánus” („frequentist”) felfogástól való megkülönböztetés érdekében.

A bayesi értelmezés mindent véletlennek tekint, ami nem pontosan ismert, és a mennyiségre vonatkozó ismeretállapot - akár kísérletből vagy más forrásból származik, beleértve a szubjektív becslést is – valószínűség-sűrűségfüggvénnyel[2] modellezhető. Ennek a szemléletnek megfelelően csaknem minden véletlen. Például: tetszés szerint sok számjegyet tulajdoníthatunk a π-nek, a rávonatkozó ismertünk hiányos (nem teljes) marad. Ezt a nem teljes ismeretet egy (nagyon keskeny) valószínűség-sűrűségfüggvény fejezi ki, úgyhogy a π véletlen változónak tekinthető.

Össze nem illő (nemkompatibilis) fogalmak (Incompatible concepts)

Amikor a bayesi felfogást bevezették a mérésekbe, a *valódi érték* fogalma több bírálat célpontjává vált. A valódi értéken és a hibán alapuló felfogást megkérdőjelezték, mivel az meg nem ismerhető mennyiségekre, vagyis idealizált koncepciókra épült. Magukat a fogalmakat is kiirtották az irodalomból, és aki használni merészelte azokat, azt az elavult elgondolások hívének tekintették. Akkoriban (és még ma is) a statisztikai oktatás, amit a fizikusok és főként a metrológusok kaptak, lényegében véve frekvenciánus volt, úgyhogy számukra nem volt könnyű elfogadni az „újszerű” és gyakran intuícióellenes (counterintuitive) bayesi felfogást.

Nyilvánvaló, hogy a valódi érték fogalma mind filozófiai, mind logikai szempontból össze nem illő (inkompatibilis) a *véletlen mennyiség* fogalmával. Gyakorlati szempontból azonban ezek a fogalmak egyáltalán nem inkompatibilisek. A gyakorlatban tekinthetjük úgy, hogy a π véletlen változó, mert a rá vonatkozó ismeretünk nem teljes. Nem vonjuk azonban kétségbe, hogy értéke egy (ma még megismerhetetlen) valós szám.

Így hát az a kérdés, hogy a *valódi érték* és a *hiba* logikailag érvényes fogalmak-e, a formálisan korrekt bizonytalanságelmélet kifejlesztése szempontjából lényegtelen. Annak az eldöntése, hogy a π véletlen mennyiség-e, vagy van egy valódi értéke egy hozzá társított valószínűség-sűrűségfüggvénnyel együtt, ami leírja a rá vonatkozó ismertállapotot, lehet

---

[1] lásd: http://en.wikipedia.org/wiki/Bayesian_probability

[2] A valószínűség sűrűségfüggvény olyan függvény, amely leírja annak a relatív likelihoodját, hogy egy véletlen változó felvesz egy adott értéket. Annak a valószínűségét, hogy a véletlen változó egy adott tartományba esik, a változó sűrűségének az erre a tartományra vett integrálja adja meg. A valószínűség sűrűségfüggvény mindenütt nemnegatív, és integrálja a teljes tér felett az egységgel egyenlő. [forrás: Wikipedia]

gyökeres filozófiai probléma, de a gyakorlati metrológus számára, aki megalapozott bizonytalanságszámításokat akar végezni, aligha több mint ízlés dolga.

A bizonytalan bizonytalanság

Ha a valószínűség-sűrűségfüggvényt használjuk a mennyiségre vonatkozó ismeretállapot modellezésére, akkor a varianciák (pontosabban szólva azok pozitív négyzetgyökei vagy a szórások) használhatók a bizonytalanság mértékeként. Ez a GUM által előírt módszer, és megengedi, hogy minden bemenő mennyiség hibáira ugyanazt a terjedési szabályt alkalmazzuk, így jutva el a mérendő mennyiség bizonytalanságának egységes megállapításához. Általánosan előforduló helyzetekre vonatkozóan a GUM ad némi útmutatást a valószínűség-sűrűségfüggvény kijelöléséhez és a megfelelő varianciák kiszámításához.

A kijelzések (észlelési eredmények) mintájából kapott variancia és a bemenő mennyiségnek tulajdonított valószínűség-sűrűségfüggvény azonban egy lényeges tulajdonságban különbözik egymástól. A korábbi a valódi halmaz szórásbecslése, és a minta nagyságától függő bizonytalansága van, ami a szabadságfokkal mérhető[3]. Ez azt jelenti, hogy a mérendő mennyiség varianciája (vagyis a mérési bizonytalanság) maga is bizonytalanná válik.

Hogy csökkentse az összhang hiányát (inkonzisztenciát), amit a minta varianciájának és a valószínűség-sűrűségfüggvény-varianciáknak a keveredése okoz, a GUM az utóbbiakat mesterségesen a becslésekre korlátozza, amelyekhez egy szabadságfokot rendel. Ez nem a GUM főszövegében található, hanem a *G* mellékletben, ahol a *kiterjesztett bizonytalanságot* tárgyalják, ami a standard bizonytalanságon alapul, de különbözik is attól (és használhatóbb a végfelhasználó számára). Az érv az, hogy a becslésnek tulajdonított valószínűség-sűrűségfüggvény nem teljesen megbízható, és ezért szükséges a varianciájához egy szabadságfokot csatolni. Ily módon a GUM két lehetséges nézőpontja összebékül, a bayesi felfogást a frekvenciánus felé terelve. Ugyanakkor látható, hogy a GUM most az ellenkező irányzatot részesíti előnyben.

A biztos bizonytalanság

Egy olyan bizonytalanság, aminek magának is bizonytalansága van, magába foglalja a valódi, ismeretlen bizonytalanság gondolatát. A mérendő mennyiség bizonytalanságának a valódi bizonytalanság becslésének kell lennie. Ez a nézőpont bizonyára figyelemre méltó, de az utóbbi években mind több és több kritika éri. Ráadásul, nem tűnik meggyőzőnek szabadságfok csatolása a valószínűség-sűrűségfüggvény-varianciákhoz.

A Bayesi paradigma (szabályrend) szigorú kereteket nyújt, amelyeken belül a bizonytalanságoknak nincs bizonytalansága. Ebben a rendszerben minden bemenő mennyiségnek egy valószínűség-sűrűségfüggvényt tulajdonítanak, hogy azok reprezentálják a rájuk vonatkozó ismeretállapotot, függetlenül attól, hogy az ismeret az adatokból (a mérésből) származik, vagy nem. Fordítsuk figyelmünket az első kategóriára! Ha az ismeret *m* adat

[3] Szokásos módszer a szabadságfokot úgy értelmezni, mint azoknak a független információadagoknak (pieces of information) a számát, amelyek rendelkezésre állnak egy másik információadag becsléséhez. Konkrétabban: a szabadságfokok száma az adatmintában levő független észlelések száma, amelyek rendelkezésre állnak ahhoz, hogy becsüljék annak a halmaznak egy paraméterét, amelyből a mintát vették. Ha például két észlelésünk van, akkor az átlagérték kiszámításához két független észleléssel rendelkezünk, de ha a varianciát számítjuk ki, akkor csak egy független észlelésünk, mert a két észlelés egyenlő távolságra van az átlagtól. [forrás: Wikipedia]

halmazából származik, s szórás mellett, akkor a GUM jelenlegi megközelítésével a megfelelő variancia a minta varianciája (s^2/m, amit a GUM az „átlagérték tapasztalati varianciájának" nevez) olyan szabadságfokkal, ami 1-el kisebb ($\nu = m$ -1) mint a mintában levő adatok száma. Ez becsli a hipotetikus adatok végtelen halmazának V varianciáját, amely halmazról feltételezik, hogy az esetek többségében Gauss eloszlású. A Bayes-féle értelmezésben, ahogyan az a GUM 1. és 2. Kiegészítésében [4,5] meg van fogalmazva, a megfelelő variancia a skálázott és eltolt Student-féle t-eloszlás varianciája, elkerülve így annak a szükségességét, hogy a valószínűség-sűrűségfüggvény-varianciákhoz megkérdőjelezhető szabadságfokokat illesszenek.

A Student-féle t-eloszlásnak m = 2 vagy m = 3 esetében nincs definiált varianciája. Ez megfelel a józanésznek. Ki bízna meg egy olyan tapasztalati szórásban, amit mindössze 2 vagy 3 észlelési eredményből határoztak meg? Ilyen esetekben előnyösebb valamilyen előzetes (prior) ismeret felhasználása, ahogyan azt a jelenlegi GUM javasolja.

A Bayes-féle paradigma előnyei kézenfekvőek: a leginkább szembetűnő az, hogy a mérendő mennyiséghez társított bizonytalanságnak nincs további bizonytalansága. Ez a pont azonban mélyebb megvitatást igényel. Az az elgondolás, hogy a mérési eredményhez társított „valódi"bizonytalanság egy becslés lehet, nem hozható összhangba az egyetlen értékkel jellemzett mérendő mennyiség elgondolásával. Felvetődik azonban, hogy mi a jelentése ebben a gondolatkörben a mérési eredményhez társított bizonytalanságnak. Például kör-összehasonlító méréseknél, amelyeknél a különféle résztvevők ugyanarra a mérendő mennyiségre nemcsak különböző becsléseket, hanem különböző bizonytalanságokat is adnak meg, melyik értelmezést kell alkalmazni? A cikk szerzőinek az a véleménye, hogy a válasz a koncepció szubjektivitásában rejlik. Mindegyik résztvevő a mérendő mennyiségre vonatkozó személyes ismeretállapota alapján nyilatkozik a bizonytalanságról, amit az általa elvégzett speciális mérések alapján alakított ki. Alternatívaként az is megtehető, hogy a mérendő mennyiség lehetséges értékeinek intervallumában a „hit mértéke" kifejezését használják. Mindkét koncepció, az „ismeretállapot" és a „hit mértéke", a szubjektivitás egy fokát foglalja magába. Így a Bayes-féle gondolatkörben kiszámított bizonytalanság biztos, de bizonyos mértékig szubjektív. Ezek a figyelembe veendő kívánatos tulajdonságok.

Jobb mértékek

A szórás a mennyiségre vonatkozó ismeretállapot egyik jó mértéke. Számos alkalmazásban azonban nem elegendő, például akkor, ha a felhasználó egy olyan tartományt akar megállapítani, amelyben a mérendő mennyiség értéke a kívánt valószínűséggel benne fekszik. Ezt az intervallumot a frekvenciánus gondolatkörben *konfidencia tartománynak* nevezik (a GUM-ban *intervallumnak*) vagy - a GUM 1. és 2. Kiegészítésében – *fedési tartománynak* (illetve *megbízhatósági tartománynak*). E felfogások között kisebb különbségek vannak, így a jellegi irányzatot követve, az utóbbit fogjuk megtárgyalni.

A megbízhatósági tartomány annak a tartománynak egy része, amelyre a mennyiség valószínűség-sűrűségfüggvénye definiálva van, lévén a megbízhatósági valószínűség a valószínűség-sűrűségfüggvény és a megbízhatósági tartomány által átfogott terület, míg a valószínűség-sűrűségfüggvény alatti teljes terület az egységgel egyenlő. Ezért ahhoz, hogy adott valószínűség mellett megadjanak egy megbízhatósági tartományt, ismerni kell a valószínűség-sűrűségfüggvényt. Ha a bemenő mennyiségekre vonatkozó valószínűség-sűrűségfüggvények adottak, akkor már csak az a probléma, hogy a bemenő mennyiség valószínűség-sűrűségfüggvényeiből felépítsék a mérendő mennyiség valószínűség-sűrűségfüggvényét.

A matematikai statisztikában ennek a problémának egyenes megoldása van, amit azonban, néhány legegyszerűbb és érdektelen esettől eltekintve, analitikus módon csaknem lehetetlen alkalmazni. A GUM 1. Kiegészítése egy numerikus szimulációs módszert javasol (a Monte Carlo-t), amely hatékonyan és konzisztens módon megadja a mérendő mennyiség valószínűség-sűrűségfüggvényét, és azt is, hogyan lehet a kívánt megbízhatósági tartományt felépíteni. A GUM is ajánl a G mellékletében egy megoldást, ami azonban csak számos feltétel fennállása esetén ad helyes eredményt, és ez korlátozza az alkalmazhatóságát. További fontos tulajdonsága az 1. Kiegészítésnek, hogy listázza a bemenő mennyiségeknek való valószínűség-sűrűségfüggvény tulajdonítás választékának sokkal gazdagabb receptjeit, mint maga a GUM, és hogy két módszeren alapul: az adatmintákra épülő ismeret kezelésének Bayes-féle eljárásán és a maximális entrópia elvén.[4]

Az következő GUM

Az eddig elmondottak alapján úgy tűnhet, hogy a GUM már elavult. De nem ez a helyzet: a GUM, mint közelítés, számos esetben jól használható és olyannyira elterjedt, hogy nem lenne bölcs dolog mellőzni. Inkább az a probléma, hogy a GUM nincs összhangban a saját Kiegészítéseivel, amelyekben a bayesi felfogás érvényesül. Ezért kell a GUM-ot felülvizsgálni. A folyamat már elkezdődött a JCGM (Metrológiai Útmutatók Vegyes Bizottsága) 1. munkacsoportjában, és feltehetőleg néhány évig fog tartani. Az azonban biztos, hogy a következő GUM-ot inkább az evolúció, mint a revolúció fogja jellemezni.

A Bayes szabály

A Bayes szabály[5] egy olyan mechanizmus, ami lehetővé teszi friss ismeret ráépítését a prior ismeretre. Ez úgy történik, hogy a mérendő mennyiségre vonatkozó ismeretet megadó és rendszerint széles prior (előzetes) valószínűség-sűrűségfüggvényt megfelelő módon kombinálják a mérési adatokkal abból a célból, hogy a javult ismeretállapotot leíró, rendszerint keskenyebb posterior valószínűség-sűrűségfüggvényt megkapják. A nagyon megbízható mérési adatok a prior valószínűség-sűrűségfüggvényt elhanyagolhatóvá teszik, míg a gyengébb mérési adatok keveset adnak hozzá a meglévő ismerethez, úgyhogy a posterior valószínűség-sűrűségfüggvény csak kevésé különbözik a prior valószínűség-sűrűségfüggvénytől.

Példaként tekintsünk egy 1 kilogrammos tömegetalont, amit egy nagyon pontos mérleggel kell kalibrálni. Az etalon gyártója azt állítja, hogy az etalon megfelel az OIML E_2 osztálynak, amelyre a megengedett legnagyobb hiba 1,6 mg. A legkézenfekvőbb eljárás ahhoz, hogy ezt az ismeretet a megfelelő valószínűség-sűrűségfüggvénybe „becsomagoljuk”, ha olyan egyenletes eloszlást választunk prior valószínűség-sűrűségfüggvényként, amelynek a középpontja 1 kg, a szélessége pedig 3,2 mg. Ha a kalibrálással kapott eredmény bizonytalansága, mondjuk, 10 μg, akkor belátható, hogy a prior ismeretnek nem lesz hatása,

---

[4] A Bayes-féle valószínűségelméletben a maximális entrópia elve egy axióma. Megállapítja, hogy pontosan ismert prior adatok esetén melyik az a javaslat, ami kifejezi a tesztelhető információt, melyik az a legtöbb információt tartalmazó valószínűségeloszlás, ami a legjobban fejezi ki az aktuális ismeretállapotot

.

[5] Lásd: http://en.wikipedia.org/wiki/Bayes_rule

úgyhogy a Bayes szabály csaknem egybeesik a hagyományos eljárással, amelyben a prior ismeretet nem veszik figyelembe. A helyzet megváltozik az etalon újrakalibrálásakor, amelynél a rendelkezésre álló prior ismeret sokkal jobb, mint ami a tényleges kalibrálással kapható. Ebben az esetben a friss információ elhanyagolható a prior ismerethez képest, és a kalibrálás jobbára úgy tekinthető, mit az etalon stabilitásának ellenőrzése.

A Bayes szabály gyakorlati mérésekben való alkalmazását a legutóbbi tíz évig nem szorgalmazták a számítások bonyolultsága miatt. Majd a Markov Lánc Monte Carlo numerikus szimulációs technikával sikerült megoldani a mágikus számítási problémákat. Ezzel a módszerrel a mérésben szereplő minden mennyiségre vonatkozó prior ismeret, beleértve a mérendő mennyiséget is, egyenes és kényelmes módon belefoglalható a mérési modellbe. Ezt a teljesen bayesi következtetés-levonást széles körben alkalmazzák a geológiától a meteorológiáig, de viszonylag kevéssé ismert a metrológiában.

Következtetés

Ebben a cikkben áttekintettük a metrológia néhány alapvető fogalmát (*mérés, mennyiség, mennyiségérték, valódi érték…*) és megnéztük, hogyan fejlődött a mérési bizonytalanság fogalma a *hiba* fogalmától (ami az észlelt érték és a valódi érték közötti különbség) addig, hogy a valószínűségen alapuló mérték legyen. Most egy átmeneti időszakban vagyunk, amikor a frekvenciánus (gyakoriságra alapozott) felfogástól a Bayesi felfogás felé mozdulunk el, amely az „ismeretállapotunkat” vagy a „hit mértékét” a valószínűség-sűrűségfüggvénnyel írja le. A GUM felülvizsgálata, amit jelenleg a JCGM/WG 1 folytat, ennek a folyamatnak egy újabb mérföldköve lesz.

A *hiba* és a *valódi érték* fogalmával összeegyeztethetetlen új felfogást, a bizonytalanságnak az ismeretállapotot leíró valószínűség-sűrűségfüggvényre alapozott értékelését úgy kell tekintenünk, mint az ezekből a fogalmakból eredő természetes fejlődést.

Köszönetnyilvánítás

Ez a cikk Walter Bich „A hibáktól a valószínűség-sűrűségfüggvényekig – A mérési bizonytalanság fogalmának fejlődése” tanulmányának adaptálása, ami az IEEE Transactions on Instrumentation and Measurements legközelebbi számában fog megjelenni. A szerzők köszönetet mondanak az IEEE-nek, amiért megengedte a cikk közzétételét az OIML Bulletinben. A szerzők: WB a JCGM WG1 (GUM) elnöke, WK a tagja. Az ebben a cikkben kifejtett vélemény nem feltétlenül fejezi ki a munkacsoport véleményét.

A szerzők hálás köszönetüket fejezik ki a JCGM WG 1 tagjainak a hasznos általános vitákért, amelyek hozzájárultak a témában kialakított véleményük megformálásához.

Irodalom:

[1] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPSP és OIML, *International Vocabulary of metrology – Basic and general concepts and associate terms (VIM)* 3rd edition, JCGM 200:2012. Published by OIML as OIML V 2-200:2012 (Letöltésre elérhető: http://www.oiml.org/publications/).

[2] S.S. Stevens, „A skálák és a mérések elméletéről” *Science* **103**, no. 2684, pp. 677-680, June 1946.

[3] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPSP és OIML, *Guide to the expression of uncertinty in measurement (GUM),* JCGM 100:2008. Published by OIML as OIML g 1 – 100:2008 ((Letöltésre elérhető: http://www.oiml.org/publications/).

[4] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPSP és OIML, *Evaluation of measurement data – Supplement 1 to „Guide to the expression of uncertinty in measurement – Propagation of distributions using Monte Carlo method,* JCGM 101:2008. Published by OIML as OIML G 1 – 101:2008 ((Letöltésre elérhető: http://www.oiml.org/publications/).

[5] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPSP és OIML, *Evaluation of measurement data – Supplement 2 to the „Guide to the expression of uncertinty in measurement – Extension to any number o faoutput quantities,* JCGM 102:2011. Published by OIML as OIML G 1 – 102:2008 ((Letöltésre elérhető: http://www.oiml.org/publications/).

Fordította: Bánkuti László