SZISZTEMATIKUS HIBÁK KIKÜSZÖBÖLÉSÉNEK EGY MÓDSZERÉRŐL

Írta: CSISZÁR IMRE és DOBÓ ANDOR

Bevezetés

A gyakorlatban számos esetben előfordul, hogy valamely mérés vagy termelési eljárás stb. során fellépő hiba egy esetről esetre változó véletlen hibán kívül valamely beállítási pontatlanságból adódó szisztematikus hibából is származik. Minthogy ezt a szisztematikus hibát magában nem tudjuk megmérni, kiküszöbölése csakis a teljes hibának, azaz a véletlen és rendszeres hiba összegének a megfigyelése alapján történhet. Ez a probléma lép fel például a következő esetekben:

a) Valamely gyártási folyamatnál a gyártott termékeknek az előírt mérettől való eltérése egyrészt a gyártó gép nem teljesen pontos beállításából, másrészt pedig a gyártás folyamata során fellépő véletlen hibákból tevődik össze. Jelöljük az előírt méretet c-vel, a beállítási hibát ξ-vel. Itt a ξ hiba nagysága a véletlentől függ, tehát ξ valószínűségi változó, de a gyártási folyamat egy szakaszában (pl. a gép újabb beállításáig) értéke állandó. A gép tehát a $c+\xi$ méretre van beállítva, ez azonban nem jelenti azt, hogy a gyártott termék minden egyes darabjának az előírt c mérettől való eltérése pontosan ξ lesz, mert a gyártás során esetről esetre változó véletlen hibák is fellépnek. Jelöljük a $c+\xi$ beállítási értéktől való méreteltérést η-val. Ekkor a teljes hiba $\xi+\eta=\zeta$ lesz. A gyártásközi ellenőrzés során a termelt alkatrésznek az előírt c mérettől való eltérését mérjük, azaz a ζ értékeket figyeljük meg s ezek alapján kell következtetnünk arra, hogy a ξ beállítási hiba nem volt-e a megengedhetőnél nagyobb, illetve kell a gép beállítását módosítanunk.

b) Tekintsünk egy skálabeosztással ellátott mutatós műszert; egyszerűség kedvéért tegyük fel, hogy a mutató kitérésének nagysága arányos a mért mennyiség értékével. A műszer legyen olyan, hogy a kitérés nagyságát szabályozni tudjuk (de a szabályozás után is megmarad a kitérésnek a mért mennyiség értékével való arányossága). Ha már most ezzel a műszerrel mérést végzünk, akkor a mérési hiba a műszer beállítási hibájából (ξ) és egy véletlen hibából (η) tevődik össze. A beállítási hibát úgy küszöbölhetjük ki, hogy ugyanazt a mennyiséget egy pontos műszerrel is megmérjük, miáltal a $\zeta=\xi+\eta$ értéket, vagyis a teljes mérési hibát határozzuk meg. A ζ-ra vonatkozó megfigyelésekből kell következtetnünk ξ nagyságára és ennek alapján kell a műszer beállítását úgy módosítanunk, hogy a ξ beállítási hiba kiküszöbölődjék.

Tulajdonképpen ez az eset lép fel az órák szabályozásánál is. Itt a mérendő mennyiség az idő. A beállítás hibáját, azaz a szisztematikus késést, ill. sietést a sebességszabályozó kar állításával lehet kiküszöbölni. A szisztematikus hibát azonban tisztán itt sem figyelhetjük meg, a beszabályozást csak a teljes késés, ill. sietés megfigyelése alapján végezhetjük, melynek nagysága a beállítás hibáján kívül véletlenszerű körülményektől (hőmérséklet, levegő páratartalma stb.) is függ.

c) Egy lövegnek valamely célpontra való beállítása úgy történik, hogy a cél távolságát bemérik; minthogy ez a mérés nem teljesen pontos, így a tényleges c távolság helyett valamely $c+\xi$ távolságra történik az irányzás. Egyszerűség kedvéért tegyük fel, hogy iránybeli hiba nem lép fel. Ekkor a becsapódás helyének a céltól való távolsága $\zeta=\xi+\eta$, ahol η a becsapódás véletlen hibája, amely a légköri viszonyoktól, a lőszer minőségétől stb. függ. A bemérés ξ hibáját közvetlenül nem tudjuk megállapítani, azonban rendszerint mérni lehet a becsapódási helynek a céltól való távolságát, vagyis $\xi+\eta=\zeta$ értékét. Ennek megfigyelése alapján kell következtetnünk ξ nagyságára és ennek megfelelően korrekciót hajtani végre a löveg beállításában.

Mindezekben a példákban az a közös vonás, hogy egy kísérletsorozaton belül állandó ξ hiba hatását mindig úgy kell kiküszöbölnünk, hogy egy $\zeta=\xi+\eta$ valószínűségi változó értékeit figyeljük meg, s ennek alapján becsült ξ értékkel korrigálunk. Ennek során mindig feltételezzük, hogy a valószínűségi változónak tekintett ξ beállítási hiba a priori eloszlása ismeretes. A következőkben ennek a korrekciónak a módjaival foglalkozunk.

A gyakorlatban általában úgy szokás eljárni, hogy a ζ teljes hibát többször mérik, s a mérési eredmények átlagával korrigálják a beállítási értéket. Ennek a módszernek a használata indokolt is abban az esetben, ha a ξ beállítási hiba eloszlásáról semmiféle információnk nincs. Igen gyakran előfordul azonban, hogy előző mérések alapján ξ a priori eloszlása (legalábbis közelítőleg) ismeretes. Természetesen ennek felhasználásával általában a fentinél célravezetőbb korrigálási eljárás is megadható, amennyiben a BAYES-tétel segítségével a korrigálásnál figyelembe vehetjük a korábbi mérések által szolgáltatott (az a priori eloszlás ismeretében megnyilvánuló) információkat is. A beállítási hiba a priori eloszlását figyelembe vevő korrigálási eljárás ipari bevezetése pl. tömeggyártás során számottevő gyakorlati haszonnal járna, minthogy lehetővé tenné a selejtarány csökkenését, a minőség javítását stb.

1. §.

Tegyük fel, hogy valamely mérésnél vagy termelési folyamatnál stb... kétféle hiba lép fel: egy „szisztematikus", beállítási pontatlanságból származó, valamely kísérletsorozaton belül állandó ξ hiba és egy kísérletenként változó η „véletlen" hiba. Tehát a mérés eredménye, illetve a gyártott termék jellemző mérete stb... $\tilde{c}=c+\xi+\eta=c'+\eta$, ahol c a mérendő mennyiség tényleges értéke, illetve a gyártott munkadarabok előírt mérete stb., $c'=c+\xi$ pedig c-nek a ξ szisztematikus hibával terhelt értéke, mely a műszer, ill. a gép beállítási pontosságától függ. A továbbiakban c'-t beállítási értéknek fogjuk nevezni. (Itt ξ-t és η-t 0 várható értékű valószínűségi változóknak tekintjük.)

Tegyük fel, hogy mérni tudjuk a $\zeta=\xi+\eta$ teljes hiba értékét és ilyen mérések alapján akarjuk (a mérőműszer, illetve a gyártó gép stb. beállításának módosításával) kiküszöbölni a ξ szisztematikus hibát.

Ez úgy történik, hogy ζ-ra vonatkozóan n megfigyelést végzünk, azaz mérjük a $\zeta_1=\xi+\eta_1,\ \zeta_2=\xi+\eta_2, \ldots, \zeta_n=\xi+\eta_n$ értékeket, ahol $\eta_1, \eta_2, \ldots, \eta_n$ függetlenek és egyforma eloszlásúak. Ezt követően a beállítási értéket az

$$m_n=\mathbf{M}(\xi|\zeta_1, \zeta_2, \ldots, \zeta_n) \tag{1}$$

értékkel korrigáljuk, azaz

$$c'' = c' - \mathbf{M}(\xi|\zeta_1, \zeta_2, \dots, \zeta_n) \tag{2}$$

lesz az új beállítási érték.

Mint ismeretes, ez a korrekció a legjobb abban az értelemben, hogy $\mathbf{M}((\xi - \xi^*(\zeta_1, \zeta_2, \dots, \zeta_n))^2)$

$$\xi^*(\zeta_1, \zeta_2, \dots, \zeta_n) = \mathbf{M}(\xi|\zeta_1, \zeta_2, \dots, \zeta_n) \text{ esetén minimális.}$$

A ξ hiba kiküszöbölése céljából úgy is eljárhatunk, hogy a beállítási értéket nem csak akkor korrigáljuk, amikor már mind az n megfigyelést elvégeztük, hanem minden egyes mérés után azonnal korrekciót hajtunk végre. Az első megfigyelés után, melynek eredménye $\zeta_1' = \xi + \eta_1'$ a korrekció $\mu_1 = \mathbf{M}(\xi|\zeta_1')$.

Ekkor a második megfigyelésnél (mérésnél) a beállítási érték már $c_1' = c' - \mu_1$ lesz és ehhez adódik hozzá az η_2' véletlen hiba. Ily módon a második megfigyeléskor észlelt teljes hiba $\zeta_2' = \xi - \mu_1 + \eta_2'$. Így a beállítási értékre vonatkozó második korrekció

$$\mu_2 = \mathbf{M}(\xi - \mu_1|\zeta_1', \zeta_2')$$

és így tovább. A k-adik megfigyelés a $c_{k-1}' + \eta_k'$ valószínűségi változóra vonatkozik; pontosabban annak a c-től való eltérését észleljük, és a

$$\zeta_k' = c_{k-1}' + \eta_k' - c = \xi - \sum_{i=1}^{k-1} \mu_i + \eta_k' \tag{3}$$

megfigyelt érték után a beállítási értéket a

$$\mu_k = \mathbf{M}\left(\xi - \sum_{i=1}^{k-1} \mu_i|\zeta_1', \zeta_2', \dots, \zeta_k'\right) \tag{4}$$

mennyiséggel korrigáljuk.

Az n-edik megfigyelés után ily módon nyert korrigált beállítási érték

$$c_n' = c' - \sum_{i=1}^{n} \mu_i. \tag{5}$$

Itt felvetődik az a kérdés, hogy az n megfigyelés elvégzése után kapott kétféle korrigált eredmény hogyan viszonylik egymáshoz, ha csak az összes megfigyelés után, illetőleg ha minden megfigyelés után külön-külön korrigálunk.

Erre ad választ a következő, matematikai szempontból egyszerű, de gyakorlatilag fontos tétel:

1. TÉTEL: Valamely c konstans ξ hibával terhelt $c' = c + \xi$ értékét korrigáljuk egyrészt a $c' + \eta_i$ értékeknek c-től való $\zeta_i = \xi + \eta_i$ $(i = 1, 2, \dots, n)$ eltérésének megfigyelése alapján (tehát az $m_n = \mathbf{M}(\xi|\zeta_1, \zeta_2, \dots \zeta_n)$ értékkel), másrészt pedig úgy, hogy minden egyes megfigyelés után korrigálunk a $\mu_k = \mathbf{M}\left(\xi - \sum_{i=1}^{k-1} \mu_i|\zeta_1', \zeta_2', \dots, \zeta_k'\right)$ értékkel, ahol $\zeta_k' = \xi - \sum_{i=1}^{k-1} \mu_i + \eta_k'$ a $c_{k-1}' + \eta_k'$ mennyiségnek c-től való eltérése

$\left(c'_k = c' - \sum_{i=1}^{k} \mu_i\right)$, mikoris az n megfigyelés utáni teljes korrekció $m'_n = \sum_{i=1}^{n} \mu_i$.

Ha $\xi, \eta_1, \eta_2, \ldots, \eta_n$ együttes eloszlása megegyezik $\xi, \eta'_1, \eta'_2, \ldots, \eta'_n$ együttes eloszlásával, akkor a kétféle módon nyert korrigált érték egyforma eloszlású, azaz

$$\mathbf{P}(c'' < x) = \mathbf{P}(c'_n < x). \tag{6}$$

Bizonyítás: A (4) és (5) szerint

$$c'_n = c' - \sum_{i=1}^{n} \mu_i = c' - \sum_{i=1}^{n-1} \mu_i - \mathbf{M}\left(\xi - \sum_{i=1}^{n-1} \mu_i | \zeta'_1, \zeta'_2, \ldots, \zeta'_n\right) = \tag{7}$$
$$= c' - \mathbf{M}(\xi | \zeta'_1, \zeta'_2, \ldots, \zeta'_n),$$

kihasználva, hogy a μ_i mennyiségek rögzített $\zeta'_1, \zeta'_2, \ldots, \zeta'_n$ mellett konstansok, tehát

$$\mathbf{M}(\mu_i | \zeta'_1, \zeta'_2, \ldots, \zeta'_n) = \mu_i \qquad (i = 1, 2, \ldots, n-1). \tag{8}$$

Minthogy a $\zeta'_k = \xi + \eta'_k - \sum_{i=1}^{k-1} \mu_i$ $(k = 1, 2, \ldots n)$ értékek megadása ekvivalens a $\xi + \eta'_k$ $(k = 1, 2, \ldots, n)$ értékek megadásával,

$$\mathbf{M}(\xi | \zeta'_1, \zeta'_2, \ldots, \zeta'_n) = \mathbf{M}(\xi | \xi + \eta'_1, \xi + \eta'_2, \ldots, \xi + \eta'_n). \tag{9}$$

Feltételezésünk szerint $\xi, \eta_1, \eta_2, \ldots, \eta_n$ és $\xi, \eta'_1, \eta'_2, \ldots, \eta'_n$ együttes eloszlása megegyezik, így (9) szerint $\mathbf{M}(\xi | \zeta'_1, \zeta'_2, \ldots, \zeta'_n)$ és $\mathbf{M}(\xi | \xi + \eta_1, \xi + \eta_2, \ldots, \xi + \eta_n) = \mathbf{M}(\xi | \zeta_1, \zeta_2, \ldots, \zeta_n)$ azonos eloszlású.

Ebből (2) és (7) alapján a (6) már következik.

Megjegyzések: a) A bizonyításból következik, hogy ha $\eta'_k = \eta_k$ $(k = 1, 2, \ldots, n)$ akkor $c'' = c'_n$.

b) Ha $\eta_1, \eta_2, \ldots, \eta_n$; $\eta'_1, \eta'_2, \ldots, \eta'_n$ egyforma eloszlásúak és $\eta_1, \eta_2, \ldots, \eta_n$, illetőleg $\eta'_1, \eta'_2, \ldots, \eta'_n$ egymástól és ξ-től függetlenek, akkor a tétel állításainak feltétele biztosan teljesül. Gyakorlatilag főképp ez a speciális eset jelentős. Itt ahelyett, hogy $\eta_1, \eta_2, \ldots, \eta_n$; $\eta'_1, \eta'_2, \ldots, \eta'_n$ mind egyforma eloszlásúak, elegendő csak annyit feltenni, hogy az azonos indexű η-ák és η'-k eloszlása egyezik meg.

c) A bizonyított tételből az is következik, hogy az első korrigálást n_1 megfigyelés, a másodikat további n_2 megfigyelés stb. után végezzük $(n_1 + n_2 + \ldots + n_k = n)$, akkor az így kapott korrigált beállítási érték eloszlása is megegyezik c'' és c'_n eloszlásával.

d) A bebizonyított tétel szerint az n megfigyelés utáni korrigált beállítási érték nem függ attól, hogy a korrekciót minden megfigyelés után, bizonyos megfigyeléscsoportok után, ill. csak mind az összes megfigyelés elvégzése után végezzük-e. Jegyezzük meg azonban, hogy gyakorlatilag — hacsak egyéb okok nem szólnak ellene — mégis célszerűbb minden megfigyelés után korrigálni, mert ezáltal a beállítási érték már a teljes megfigyeléssorozat befejezése előtt közelebb kerül a helyes értékhez. Ez azt eredményezi, hogy pl. a bevezetésben említett a) esetben a használható munkadarabok aránya, ill. a c) esetben a találati valószínűség már a korrigálási eljárás befejezése előtt növekszik. Az a) esetben ennek persze elsősorban akkor van jelentősége, ha a korrigálás nem egymás után legyártott munkadarabok alapján történik, hanem pl. minden k-adik darab lemérése alapján.

2. §.

Az 1. §-ban mondottak szerint az n megfigyelés alapján végzett korrekció jósága nem függ attól, hogy az n megfigyelés után egyszerre korrigáltunk-e, vagy pedig minden megfigyelés, illetve bizonyos számú megfigyelések után. Ezért a továbbiakban mondottak bármelyik korrigálási eljárás esetén érvényesek.

Jegyezzük meg, hogy a $\mathbf{M}(\xi|\zeta_1, \zeta_2, \dots, \zeta_n)$ mennyiség mint ismeretes ξ-nek általában nem torzítatlan becslése, azaz

$$\mathbf{M}(\mathbf{M}(\xi|\zeta_1, \zeta_2, \dots, \zeta_n)|\xi) = \xi \tag{10}$$

általában nem teljesül. A gyakorlatilag fontos esetekben azonban ez a becslés legalábbis aszimptotikusan torzítatlan, vagyis

$$\lim_{n\to\infty} \mathbf{M}(\mathbf{M}(\xi|\zeta_1, \zeta_2, \dots, \zeta_n)|\xi) = \xi. \tag{11}$$

Sőt, ezen túlmenőleg

$$\lim_{n\to\infty} \mathbf{M}(\xi|\zeta_1, \zeta_2, \dots, \zeta_n) = \xi \tag{12}$$

is teljesül (1 valószínűséggel).

(11) azt mutatja, hogy elég sok megfigyelést végezve a $c' + \eta - c = \zeta = \xi + \eta$ hibára vonatkozólag a beállítási hibát tetszőleges pontossággal ki lehet küszöbölni. Az eddig mondottakat pontosabban az alábbi tétel fejezi ki:

2. TÉTEL: *Ha az $\eta_1, \eta_2, \dots, \eta_n$ valószínűségi változók sorozatára érvényes a nagy számok erős törvénye, akkor*

$$1^\circ \quad \lim_{n\to\infty} \mathbf{M}(\xi|\zeta_1, \zeta_2, \dots, \zeta_n) = \xi$$

$$2^\circ \quad \lim_{n\to\infty} \mathbf{D}^2(\xi|\zeta_1, \zeta_2, \dots, \zeta_n) = 0$$

1 *valószínűséggel.*

Bizonyítás: Mint ismeretes (l. pl. [2] 332. o.)

$$\lim_{n\to\infty} \mathbf{M}(\xi|\zeta_1, \zeta_2, \dots, \zeta_n) = \mathbf{M}(\xi|\zeta_1, \zeta_2, \dots, \zeta_n \dots). \tag{13}$$

(Bármely $\zeta_1, \zeta_2, \dots, \zeta_n \dots$ valószínűségi változók esetén, ha csak ξ-nek létezik a várható értéke.)

Esetünkben $\zeta_i = \xi + \eta_i$ $(i = 1, 2, \dots, n)$, tehát ha az η_i sorozatra érvényes a nagy számok erős törvénye, vagyis

$$\mathbf{P}\left(\lim_{n\to\infty} \frac{\eta_1 + \eta_2 + \dots + \eta_n}{n} = 0\right) = 1 \tag{14}$$

(η-ról feltettük, hogy 0 várható értékű), akkor

$$\mathbf{P}\left(\xi = \lim_{n\to\infty} \frac{\zeta_1 + \zeta_2 + \dots + \zeta_n}{n}\right) = 1. \tag{15}$$

Ez azt jelenti, hogy ξ függvénye a $\zeta_1, \zeta_2, \ldots, \zeta_n, \ldots$ sorozatnak, tehát

$$\mathbf{M}(\xi|\zeta_1, \zeta_2, \ldots, \zeta_n \ldots) = \xi \tag{16}$$

1 valószínűséggel. A (13) és (16)-ból 1° már következik. Ugyanúgy látható be, hogy

$$\lim_{n\to\infty} \mathbf{M}(\xi^2|\zeta_1, \zeta_2, \ldots, \zeta_n) = \xi^2 \tag{17}$$

1 valószínűséggel.

A (17) és 1° szerint

$$\lim_{n\to\infty} \mathbf{D}^2(\xi|\zeta_1, \zeta_2, \ldots, \zeta_n) = \lim_{n\to\infty} [\mathbf{M}(\xi^2|\zeta_1, \zeta_2, \ldots, \zeta_n) - (\mathbf{M}(\xi|\zeta_1, \zeta_2, \ldots, \zeta_n))^2] = \tag{18}$$
$$= \xi^2 - \xi^2 = 0$$

1 valószínűséggel.

3. §.

Az elmondottak gyakorlati alkalmazása szempontjából fontos esetekben feltehető, hogy a ξ hiba és az $\eta_1, \eta_2, \ldots, \eta_n$ „véletlen hibák" teljesen független valószínűségi változók (0 várható értékkel), ahol az η_i-k ($i = 1, 2, \ldots, n$) egyforma eloszlásúak $g(y)$ sűrűségfüggvénnyel és ξ-nek is létezik $f(x)$ sűrűségfüggvénye.

Ekkor a BAYES-tétel alapján ξ-nek a $\xi + \eta_1 = z_1, \ldots, \xi + \eta_n = z_n$ feltételek melletti feltételes sűrűségfüggvénye

$$f_n(x|z_1, \ldots, z_n) = \frac{f(x) \prod_{i=1}^{n} g(z_i - x)}{\int_{-\infty}^{+\infty} f(x) \prod_{i=1}^{n} g(z_i - x)\, dx} \tag{19}$$

s így

$$\mathbf{M}(\xi|\xi + \eta_1 = z_1, \ldots, \xi + \eta_n = z_n) = \frac{\int_{-\infty}^{+\infty} x f(x) \prod_{i=1}^{n} g(z_i - x)\, dx}{\int_{-\infty}^{+\infty} f(x) \prod_{i=1}^{n} g(z_i - x)\, dx}. \tag{20}$$

Ha tehát a ξ hiba kiküszöbölését n megfigyelés elvégzése után történő korrigálással végezzük, a szükséges m_n korrekciót a (20) formula adja meg.

Ha viszont minden egyes megfigyelés után korrigálunk, akkor (20) alapján a k-adik megfigyelés után szükséges korrekció kifejezése

$$\mu_k = \frac{\int_{-\infty}^{+\infty} x f(x) \prod_{i=1}^{k} g(z_i' + m_{i-1}' - x)\, dx}{\int_{-\infty}^{+\infty} f(x) \prod_{i=1}^{k} g(z_i' + m_{i-1}' - x)\, dx}, \qquad \left(m_i' = \sum_{j=1}^{i} \mu_j\right) \tag{21}$$

ahol z_i' a ζ_i' megfigyelt értékét jelenti ($i = 1, 2, \ldots, k$).

Példaképpen tekintsük a mondottaknak azt a speciális esetét, amikor a ξ hiba a priori eloszlása és az $\eta_1, \eta_2, \ldots, \eta_n$ valószínűségi változók közös eloszlása egyaránt normális, 0 várható értékkel és σ_1, ill. σ_2 szórással. Ez esetben (19)-ből egyszerű számolással adódik, hogy az n megfigyelés után ξ a posteriori eloszlása ugyancsak normális lesz, mégpedig a szórás független attól, hogy melyek voltak a megfigyelt értékek, továbbá, hogy a kísérletsorozatban az egyes kísérletek után végeztünk-e korrekciót, nevezetesen

$$\mathbf{D}^2(\xi|\zeta_1, \zeta_2, \ldots, \zeta_n) = \mathbf{D}^2(\xi|\zeta_1', \zeta_2', \ldots, \zeta_n') = \frac{\sigma_1^2\sigma_2^2}{n\sigma_1^2+\sigma_2^2}.$$

(Itt közvetlenül látható, hogy $n \to \infty$ esetén

$$\mathbf{D}^2(\xi|\zeta_1, \zeta_2, \ldots, \zeta_n) \to 0$$

minden $\zeta_1, \zeta_2, \ldots, \zeta_n$ sorozat esetén. A 2. tételből ennek a relációnak csak az 1 valószínűséggel való fennállása következik.) Az a posteriori eloszlás várható értéke pedig

$$m_n = \mathbf{M}(\xi|\zeta_1, \zeta_2, \ldots, \zeta_n) = \frac{\sigma_1^2}{\sigma_1^2 + \frac{1}{n}\sigma_2^2} \cdot \frac{z_1 + z_2 + \ldots + z_n}{n}, \tag{22}$$

ahol $z_1, z_2, \ldots, z_n$ a $\zeta_1, \zeta_2, \ldots, \zeta_n$ változók megfigyelt értékei (ha a kísérletsorozat közben nem végeztünk korrekciót). Ha pedig minden kísérlet után korrigáltunk, akkor az

$$m_k' = \frac{\sigma_1^2}{\sigma_1^2 + \frac{1}{k}\sigma_2^2} \cdot \frac{u_1 + u_2 + \ldots + u_k}{k}, \tag{23}$$

$$u_i = z_i' + m_{i-1}' \qquad (i = 1, 2, \ldots, k)$$

rekurzív összefüggés által definiált m_n' mennyiség adja ξ feltételes várható értékét, ahol z_k' a ζ_k' valószínűségi változó mért értékét jelenti. A közöltek alapján nyilvánvaló, hogy

$$\mathbf{P}\left(-\frac{\lambda\sigma_1\sigma_2}{\sqrt{n\sigma_1^2+\sigma_2^2}} < \xi - m_n < \frac{\lambda\sigma_1\sigma_2}{\sqrt{n\sigma_1^2+\sigma_2^2}} \,|\zeta_1 = z_1, \zeta_2 = z_2, \ldots, \zeta_n = z_n\right) = 2\Phi(\lambda) - 1,$$

illetve

$$\mathbf{P}\left(-\frac{\lambda\sigma_1\sigma_2}{\sqrt{n\sigma_1^2+\sigma_2^2}} < \xi - m_n' < \frac{\lambda\sigma_1\sigma_2}{\sqrt{n\sigma_1^2+\sigma_2^2}} \,|\zeta_1' = z_1', \zeta_2' = z_2', \ldots, \zeta_n' = z_n'\right) = 2\Phi(\lambda) - 1,$$

ahol

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} du.$$

Ha azt akarjuk, hogy a korrigált beállítási érték adott $2\Phi(\lambda)-1$ valószínűségi szint mellett előírt pontossággal megközelítse a helyes c értéket, vagyis ha megkívánjuk a

$$\mathbf{P}(|\xi - m_n| \leqq a) \geqq 2\Phi(\lambda) - 1,$$

illetve a

$$\mathbf{P}(|\xi - m_n'| \leqq a) \geqq 2\Phi(\lambda) - 1$$

egyenlőtlenség teljesülését, akkor a fentiekből következik, hogy ehhez $n \geqq \frac{\lambda^2}{a^2}\sigma_2^2 - \left(\frac{\sigma_2}{\sigma_1}\right)^2$ számú megfigyelést kell végezni.

Mint ismeretes, ha az $\eta_1, \eta_2, \ldots, \eta_n$ valószínűségi változók normális eloszlásúak σ_2 szórással és 0 várható értékkel, akkor n elég nagy értékei esetén a ξ hiba maga is közelítőleg normális a posteriori eloszlású $\frac{z_1 + z_2 + \ldots + z_n}{n}$ várható értékkel és $\frac{\sigma_2}{\sqrt{n}}$ szórással. (ξ a priori eloszlására tett igen általános feltételek mellett, pl. ha felteszszük, hogy ξ-nek folytonos sűrűségfüggvénye van.) Ez más szóval azt jelenti, hogy a ξ valószínűségi változó feltételes eloszlása n növekedésével egyre kevésbé függ ξ a priori eloszlásától. — Ha ugyanis ξ sűrűségfüggvényét $f(x)$-szel jelöljük, akkor

$$f_n(x|\zeta_1, \zeta_2, \ldots, \zeta_n) = \frac{e^{-\frac{n\left(\frac{z_1+z_2+\ldots+z_n}{n}-x\right)^2}{2\sigma_2^2}} f(x)}{\int\limits_{-\infty}^{+\infty} e^{-\frac{n\left(\frac{z_1+z_2+\ldots+z_n}{n}-x\right)^2}{2\sigma_2^2}} f(x)\,dx},$$

innen pedig a jól ismert „LAPLACE-módszer" alkalmazásával az állítás már következik (l. [1], 358. o.).

4. §.

A gyakorlati alkalmazások szempontjából fontosnak látszik megvizsgálni, hogy eredményeink hogyan módosulnak akkor, ha

a) a $\xi + \eta$ hibát nem tudjuk pontosan mérni

b) a számított korrekciót csak bizonyos hibával tudjuk végrehajtani.[1]

Az a) eset figyelembevétele nem jelent különösebb nehézséget, ugyanis ha a $\xi + \eta$ mérési hibáját ε-nal jelöljük, — tehát valójában $\zeta = \xi + \eta + \varepsilon$ értékét mérjük —, ahol ε ξ-től és η-tól független valószínűségi változó, akkor korábbi meggondolásainkat a $\zeta = \xi + \eta^*$ ($\eta^* = \eta + \varepsilon$) valószínűségi változóra alkalmazhatjuk, változatlan eredményekkel. (Ekkor η^* szórásnégyzete η és ε szórásnégyzeteinek összegével egyenlő.)

[1] Ez utóbbi kérdésre SARKADI KÁROLY hívta fel figyelmünket.

A b) esetben eredményeink már nem maradnak változatlanok. Az 1. Tétel ekkor érvényét veszti: a minden megfigyelés utáni és az egy egész megfigyeléssorozat alapján végzett korrekció ekkor már nem ekvivalens.

Jelöljük ugyanis τ_i-vel az i-edik korrekció végrehajtásakor fellépő hibát, (amelyekről feltesszük, hogy ξ-től és az η_i-ktől független, egyforma eloszlású valószínűségi változók). Ekkor a ξ korrigált értéke m_i korrekció esetén $\xi-(m_i+\tau_i)$ lesz. Egyszerűség kedvéért tegyük fel, hogy ξ, η_i és τ_i $(i=1, 2, \dots, n)$ normális eloszlásúak 0 várható értékkel és σ_1, σ_2, illetve σ_3 szórással. Ez esetben a $\xi+\eta_i$ érték első megfigyelése utáni korrekció (23) szerint

$$m_1' = \frac{\sigma_1^2}{\sigma_1^2+\sigma_2^2}(\xi+\eta_1')$$

és így a ξ korrigált értéke

$$\xi_1 = \xi - m_1' - \tau_1 = \frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}\xi - \frac{\sigma_1^2}{\sigma_1^2+\sigma_2^2}\eta_1' - \tau_1 .$$

Tehát ξ_1 is normális eloszlású 0 várható értékkel és

$$\mathbf{D}^2(\xi_1) = \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2} + \sigma_3^2 \tag{24}$$

szórásnégyzettel.

A meggondolás ismétlésével nyerjük, hogy a k-ik korrekció utáni ξ_k korrigált beállítási hiba is normális eloszlású és szórásnégyzete

$$\mathbf{D}^2(\xi_k) = \frac{\mathbf{D}^2(\xi_{k-1})\cdot\sigma_2^2}{\mathbf{D}^2(\xi_{k-1})+\sigma_2^2} + \sigma_3^2 = \frac{\sigma_2^2}{1+\dfrac{\sigma_2^2}{D^2(\xi_{k-1})}} + \sigma_3^2 \qquad (k=1, 2, \dots, n;\ \xi_0=\xi) \tag{25}$$

Ha viszont k megfigyelés után egyszerre korrigálunk, akkor a korrekció (22) szerint

$$m_k = \frac{\sigma_1^2}{\sigma_1^2+\dfrac{1}{k}\sigma_2^2}\left(\xi+\frac{\eta_1+\dots+\eta_k}{k}\right),$$

tehát ξ korrigált értéke

$$\xi_k^* = \xi-(m_k^*+\tau) = \frac{\sigma_2^2}{k\sigma_1^2+\sigma_2^2}\xi - \frac{\sigma_1^2}{k\sigma_1^2+\sigma_2^2}[\eta_1+\dots+\eta_k]-\tau,$$

azaz

$$\mathbf{D}^2(\xi_k^*) = \frac{\sigma_1^2\sigma_2^2}{k\sigma_1^2+\sigma_2^2} + \sigma_3^2 = \frac{\sigma_2^2}{1+\dfrac{\sigma_2^2}{D^2(\xi_{k-1}^*)-\sigma_3^2}} + \sigma_3^2 . \tag{26}$$

A (25) és (26) alapján teljes indukcióval azonnal látható, hogy

$$\mathbf{D}^2(\xi_k) \geqq \mathbf{D}^2(\xi_k^*), \tag{27}$$

ahol $k=1$-re az egyenlőség teljesül (24) miatt. Egyúttal az is látható, hogy (27)-ben $k \geqq 2$ esetben határozott egyenlőtlenség áll fenn.

Eredményünk azt jelenti, hogy a beállítási hiba adott korlát alá való csökkentéséhez (adott valószínűségi szint mellett) kevesebb megfigyelésre van szükség akkor, ha csak a teljes megfigyeléssorozat elvégzése után korrigálunk, nem pedig minden lépésben.

A most mondottakat az 1 § 1. Tétel d) megjegyzésével összevetve megállapíthatjuk, hogy bizonyos esetekben egy egész megfigyeléssorozat utáni, más esetekben viszont a lépésenkénti korrigálás a célszerűbb.

Hogy mikor melyiket előnyösebb használni, azt konkrét esetekben külön vizsgálat alapján kell eldönteni.

IRODALOM

[1] RÉNYI ALFRÉD: *Valószínűségszámítás,* Tankönyvkiadó, Budapest, 1954.
[2] J. L. DOOB: *Stochastic Processes,* Wiley, New York, 1953.

(Beérkezett: 1962. II. 15)