

A fizikai mérések hibája

A természettudományos megismerés egyik alapvető módszere a *kísérletezés, mérés*. Tapasztalatokat, adatokat szolgáltat a gyakran nem nyilvánvaló összefüggések felismeréséhez, és lehetőséget nyújt az elméleti eredmények ellenőrzésére. A mérés során a mérendő mennyiséget az egységül választott mennyiséggel (etalonnal) hasonlítjuk össze, melynek eredményét számokkal fejezzük ki.

A méréssel meghatározott fizikai mennyiségek szükségképpen véges pontosságúak. Az ismételt mérések valamilyen mértékben mindig eltérnek egymástól. A mért értékek bizonytalansága különböző forrásokból származhat, úm. a mérőberendezések tökéletlensége, ismeretlen környezeti hatások, ill. a vizsgált jelenségek sztochasztikus jellege. Ezért a mérés csak akkor tekinthető teljesnek, ha a legvalószínűbbnek tartott érték mellett megadjuk annak várható hibáját is. **A mérési hiba legalább közeli meghatározása nélkül a mérési eredmény információtartalma vitatható!**

Az alábbiakban összefoglaljuk a hibaszámítás alapjait a laboratóriumi gyakorlatokhoz szükséges szinten. Az ismertetésre kerülő fogalmak ill. eljárások szigorú megalapozására a valószínűségszámítás keretében kerül sor.

1. Hibatípusok

Eredetük szerint az alábbi hibafajtákat különböztetjük meg:

- rendszeres (szisztematikus) hiba;
- véletlen hiba;
- statisztikus hiba.

Rendszeres (szisztematikus) hiba: Meghatározott módon ható okok következményeként lép fel. Azonos körülmények között végzett méréseknél nagysága és előjele nem változik. Ilyenek a mérőeszköz tökéletlenségéből származó hibák (a működés ill. hitelesítés pontatlanságai), melyek nagyságát a műszerekre vonatkozó hitelesítési táblázatok, ill. műszerleírások tartalmazzák. [Pl. a mérlegsúlyoknál megadják, hogy a valódi érték mennyire térhet el a névleges (ráírt) értéktől.] Ezen adatokat valamennyi mérőeszközre ismernünk kell, hogy a hibaszámítás során figyelembe vehessük. A szisztematikus hibák bizonyos fajtái nem kezelhetők ilyen módon. Ide tartoznak a mérési módszerek specifikus hibái, vagy az elhanyagolt külső hatásokból (nyomás, hőmérséklet, páratartalom, elektromos és mágneses szórt terek, stb.) eredő bizonytalanság. Ezek általában csak akkor észlelhetők, ha a mérést több független módszerrel is elvégezzük.

Véletlen hiba: A mérést befolyásoló sokfajta külső ok együttes következményeként lép fel. Minden egyes mérésnél másképp jelentkezik. Előjele negatív és pozitív egyaránt lehet. Ilyen eredményre vezetnek a véletlenszerűen fellépő környezeti hatások (mechanikai rezgések, feszültség- ill. hőmérséklet-ingadozások, stb.), a mechanikus alkatrészek kotyogása, beállítási- és leolvasási pontatlanságok (az utóbbiak mérésautomatizálással nagymértékben csökkenthetők), stb. A véletlen hibák teljes kiküszöbölésére nincs mód, számításba venni csak átlagos hatásukat lehet.

Statisztikus hiba: Nagyszámú egymástól független esemény megfigyelésekor lép fel. Ilyen például a részecske-

számolásnál észlelt hiba. Ha t idő alatt n független esemény következett be, az nem jelenti azt, hogy t idő alatt minden esetben n az események bekövetkezésének száma. Egyszer több, másszor kevesebb esemény következik be. Ez a *statisztikus ingadozás*. A valószínűségszámítás szerint, ha az észlelt események száma meghatározott idő alatt középértékben n , akkor ez az érték $\pm\sqrt{n}$ statisztikus ingadozást mutat. Ez a *valószínű statisztikus hiba*. A statisztikusan ingadozó mennyiségek *relatív* hibája (ld. alább) úgy csökkenthető, hogy a megfigyelést lehetőleg hosszú időre ill. nagyszámú eseményre terjesztjük ki. A $\pm\sqrt{n}/n$ relatív hiba értéke ugyanis n növelésével csökken.

Jelölések: A mérési hiba jelölésére leggyakrabban a görög delta (Δ) betűt használjuk (így pl. az l hosszúság hibája Δl), de speciális hibatípusok esetén (mint pl. a szórás) más jelöléseket alkalmazunk (ld. alább). Az x mennyiség hibáját *abszolút* (Δx), *relatív* ($\Delta x/x$), vagy *százalékos* formában ($\Delta x/x \cdot 100\%$) adhatjuk meg: $\Delta l = \pm 0,002$ m, $\Delta l/l = \pm 0,0015$, vagy $\Delta l/l = \pm 0,15\%$. Az utóbbi kettő előnye, hogy önmagukban is jelzik a mérés pontosságát. A mérési eredményeket csak annyi számjegyig érdemes feltüntetni, ahány jegyig pontosságuk terjed.

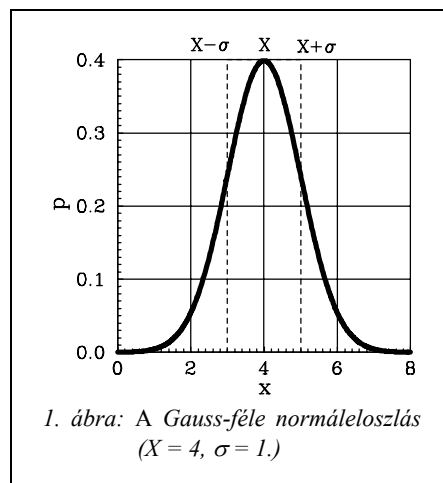
A laboratóriumi gyakorlat mérései során mindhárom hibatípus fellép. A meghatározandó mennyiségek vagy *közvetlenül* mérhetőek, vagy ismert függvénykapcsolat segítségével közvetlenül mért mennyiségekből *közvetve* határozhatók meg. Az alábbiakban a közvetlen és közvetett mérések hibáit tárgyaljuk.

2. Hibaszámítás közvetlen mérések esetén

A legkedvezőbb érték (átlagos érték): Tegyük fel, hogy egy x mennyiséget igen sokszor mérünk meg azonos körülmények között. A tapasztalat szerint az egyes mérésekből származó adatok egy értéktartományon belül szóródnak. A tartomány közepére eső értékek gyakoribbak, mint széleken levők. Ez a viselkedés rendszerint jól közelíthető a *Gauss-féle normáloszlással*:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-X)^2}{2\sigma^2}\right],$$

ahol $p(x)dx$ annak valószínűsége, hogy a mérés eredménye az $(x, x+dx)$ tartományba esik, σ az eloszlás szélessége.



gét meghatározó paraméter, és X a valódi érték. Belátható, hogy $n \rightarrow \infty$ esetén a mért értékek aritmetikai átlaga

$$x_k = \frac{1}{n} \sum_{i=1}^n x_i \rightarrow X.$$

A gyakorlatban a mérések száma korlátozott, így x_k csak X legkedvezőbb becslésének tekinthető. Az, hogy az x_i és x_k értékek mennyire közelítik meg X -et az egyes mérések hibáiból számítható, és többféle módon számszerűsíthető. Mivel a nagyszámú (normáloszlást követő) mérés hibáinak algebrai középértéke nullához tart, az egyszerű középértékképzés hiba meghatározásra nem alkalmas. Megadhatók ellenben az egyes mérések ill. a középérték *kvadratikusan* hibái. Az alábbiakban csak a leggyakrabban használt hiba definíciókra térünk ki.

Az egyes mérések kvadratikusan középhibája (várható v. standard eltérése): A mért x_i értékeknek a valódi X értéktől való átlagos eltérése. Szokásos jele σ . Gyakran használt további elnevezése az *egyes mérések várható vagy standard eltérése* (*standard deviation, Standardabweichung*). Nagyszámú mérés esetén az alábbi közelítő kifejezés alapján számítható:

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - x_k)^2} \quad (1)$$

- $x = x_k \pm \sigma$ a Gauss-görbe inflexió pontjaihoz tart. A mérési pontok 68,3 %-a esik ebbe a tartományba, míg 95,4 % ill. 99,7 % felel meg az átlag $\pm 2\sigma$ ill. 3σ nagyságú környezetének.
- Vegyük észre, hogy miközben a mérési pontok számának növelése az x_k átlagérték javulásához vezet, σ -t nem csökkenti. Ez azt tükrözi, hogy az egyes mérések pontosságán a mérés megismétlése nem változtat.

Az átlagérték kvadratikusan középhibája (várható v. standard eltérése): Az előbbinél (ti. az egyes mérésekből kapott eredmény bizonytalanságánál) lényegesebb fogalom az átlagérték közepes hibája, σ_k , mely az x_k átlagnak az X tényleges értéktől való közepes eltérését adja meg. Ennek megfelelően az *átlagérték várható vagy standard eltérése*nek is nevezik. σ_k az x_k átlagértéket körülfogó azon tartomány szélességének a fele, melyben a tényleges érték 68,3 % valószínűséggel található meg.

$$\sigma_k = \frac{\sigma}{\sqrt{n}} = \sqrt{\frac{1}{n(n-1)} \sum_{i=1}^n (x_i - x_k)^2} \quad (2)$$

- A tényleges érték 95,4 ill. 99,7 % valószínűséggel található a $x_k \pm 2\sigma_k$ és $x_k \pm 3\sigma_k$ tartományokban.
- Szemben σ -val, a mérések számának növekedésével σ_k csökken, ami azt tükrözi, hogy az átlag egyre jobban megközelíti a tényleges értéket.

A gyakran használt *szórás* (*scattering, Streuung*) és *variancia* (*variance, Varianz*) kifejezések a megfelelő *kvadratikusan középhibát* (σ , σ_k) jelentik. Szigorúan véve,

az (1) és (2) összefüggések csak igen nagyszámú mérés ($n > 100$) esetén érvényesek. A gyakorlatban sokszor meg kell elégednünk 3-10 méréssel. Ekkor a (2) egyenlettel számított σ_k -t az alábbi táblázatban megadott ún. *t-faktorral* kell megszorozni.

Általános megjegyzések: A fenti képleteket nem szabad formálisan alkalmazni. Tegyük fel, hogy egy távolságot többször megmérünk, de csak a cm-es osztásokat olvassuk le, s így mindig 100 cm-t kapunk. A (2) egyenlet formális alkalmazásával a mérési hiba nullának adódik. Valójában a pontatlan leolvasás következményével állunk szemben. A (2) képletből legfeljebb csak akkor kaphatunk reális eredményt, ha a műszer leolvasását a lehető legnagyobb pontossággal végezzük. Ekkor az utolsó jegy már becsült, tehát az egyes mérésekre különbözni fog. Ez az ingadozás nem egyszerűen a leolvasási hibából ered, hanem sok, esetleg teljesen ismeretlen tényező is szerepet játszhat benne. (Digitális műszereknél a leolvasási hiba lényegesen kisebb jelentőségű, de általában ott is található olyan jegy, melyhez a véletlenszerűen változó kijelzés miatt nehezebb értéket rendelni.) Ezért fontos, hogy a mérés ismétlésével a véletlen hibát lehetőség szerint csökkentsük. Ha a mérési adatok megegyeznek, hibaként az eredmény utolsó jegyében fél egységet veszünk. Amennyiben ismert a műszer hitelesítéssel megadott hibája is, a kettő közül a nagyobbat tekintjük a mérés hibájának. Ennél pontosabb felvilágosítás csak akkor kapható, ha a mérést több különböző módszerrel is elvégezzük.

Az (1) és (2) egyenletek származtatása: Az egyes mérések tényleges hibája $e_i = x_i - X$. Véges számú mérésből sem X , sem pedig e_i nem határozható meg. Az egyes mért értékek tényleges szórása definíció szerint $\sigma^2 = \frac{1}{n} \sum e_i^2$, ahol Σ az összes pontra való összegzést jelenti. Meghatározni ezt sem tudjuk, de a látszólagos hibákkal kifejezhető az alábbiak szerint. A tényleges hiba definíciójából

$$X = \frac{1}{n} \sum (x_i - e_i) = \frac{1}{n} \sum x_i - \frac{1}{n} \sum e_i.$$

Mint hogy $x_k = \frac{1}{n} \sum x_i$, ezért $X = x_k - \frac{1}{n} \sum e_i$, azaz a valódi érték, az egyes mérések átlagértékétől, a tényleges hibák középértékével különbözik.

A tényleges hibát definiáló egyenletből e_i -t kifejezve

$$e_i = x_i - X = x_i - x_k + \frac{1}{n} \sum e_i = v_i + e_k,$$

ahol $v_i = x_i - x_k$ a látszólagos hiba, $e_k = \frac{1}{n} \sum e_i$ a tényleges hibák átlaga. Ebből az egyenletből adódik, hogy $v_i = e_i - e_k$ ill. $v_i^2 = e_i^2 - 2e_i e_k + e_k^2$. Képezzük most az egyes mérések látszólagos hibanégyzeteinek átlagát:

$$\begin{aligned} \frac{1}{n} \sum v_i^2 &= \frac{1}{n} \sum (e_i^2 - 2e_i e_k + e_k^2) = \frac{1}{n} \sum e_i^2 - 2e_k \frac{1}{n} \sum e_i + e_k^2 = \\ &= \frac{1}{n} \sum e_i^2 - e_k^2. \end{aligned}$$

Vegyük észre, hogy

	$n = 3$	4	5	6	8	10	20
68,3 % bizonyosság	$t = 1,32$	1,20	1,15	1,11	1,08	1,06	1,03
99,7 % bizonyosság	$t = 19,2$	9,2	6,6	5,5	4,5	4,1	3,4

$$e_k^2 = \left(\frac{1}{n} \sum e_i^2 \right)^2 = \left(\frac{1}{n} \right)^2 (e_1^2 + e_2^2 + \dots + e_n^2 + 2e_1e_2 + \dots) \cong \left(\frac{1}{n} \right)^2 \sum e_i^2,$$

mivel nagyszámú normáloszlású adat esetén hasonló nagyságú vegyesszorzatok azonos valószínűséggel fordulnak elő pozitív és negatív előjellel, s így a vegyesszorzatok összege elhanyagolható. Az utóbbi két kifejezést az $\frac{1}{n} \sum v_i^2 = \frac{1}{n} \sum e_i^2 - e_k^2$

egyenletbe helyettesítve, és azt $\sum v_i^2$ -re megoldva megkapjuk a tényleges és a látszólagos hibák közti összefüggést:

$$\sum v_i^2 = \frac{n-1}{n} \sum e_i^2.$$

Ezt a szórás $\sigma^2 = \frac{1}{n} \sum e_i^2$ kifejezésébe helyettesítve az egyes mérések *tényleges* kvadratikus hibájára az (1), míg az átlagérték *tényleges* kvadratikus hibájára a (2) kifejezések adódnak.

3. Hibaszámítás közvetett mérések esetén

Méréseinknél a keresett mennyiséget legtöbbször nem közvetlenül észleljük, hanem közvetlenül mérhető adatokból ismert függvénykapcsolat alapján számítjuk. A kiinduló adatokat ismételt mérésekkel határozzuk meg. Az egyes mérési eredményekből kiszámítjuk a legvalószínűbb értéket, melynek a képletbe történő helyettesítésével kapjuk a keresett mennyiség legvalószínűbb értékét. Ebben az esetben a mérés hibája az alábbiak szerint számítható.

Tegyük fel, hogy meghatározandó mennyiségünk y , az a, b, \dots mérhető mennyiségek függvénye: $y = y(a, b, \dots)$. Mivel általában a hibája $\Delta a \ll a$, $\Delta b \ll b$, stb. felhasználhatjuk az $y = y(a, b, \dots)$ függvény Taylor-sorát, melyből a magasabbrendű tagok elhagyásával y **maximális abszolút hibáját** a következő kifejezés adja:

$$\Delta y = \left| \frac{\partial y}{\partial a} \right| \Delta a + \left| \frac{\partial y}{\partial b} \right| \Delta b + \dots \quad (3)$$

Mint hogy ebben a kifejezésben a parciális differenciálhányadosok abszolút értéke szerepel, a legrosszabb esetet tételeztük fel, amikor valamennyi hiba egyszerre és azonos irányban lép fel. (A $\Delta a, \Delta b, \dots$ hibák is pozitívnak tekintendők.) A differenciálást természetesen csak a számottevő relatív hibával rendelkező mennyiségekre kell elvégezni.

A Gauss-féle hibaterjedési törvény figyelembe veszi, hogy az egyes mennyiségek hibái részben kompenzálják egymást. A **kvadratikus abszolút hiba** az alábbi módon fejezhető ki:

$$\Delta y = \sqrt{\left(\frac{\partial y}{\partial a} \Delta a \right)^2 + \left(\frac{\partial y}{\partial b} \Delta b \right)^2 + \dots} \quad (4)$$

Mivel azonos abszolút hiba különböző nagyságrendű mennyiségekhez tartozhat, azt a mérést ítéljük pontosabbnak, amelynél a relatív hiba kisebb. Közvetett mérések esetén a **relatív hiba**

$$\frac{\Delta y}{y} = \frac{\left| \frac{\partial y}{\partial a} \right| \Delta a + \left| \frac{\partial y}{\partial b} \right| \Delta b + \dots}{y(a, b, \dots)} \quad (\text{maximális}), \quad (5a)$$

vagy

$$\frac{\Delta y}{y} = \sqrt{\left(\frac{\partial y}{\partial a} \Delta a \right)^2 + \left(\frac{\partial y}{\partial b} \Delta b \right)^2 + \dots} \quad (\text{kvadratikus}). \quad (5b)$$

Ha eredményünket tényezőkből álló kifejezés adja, akkor a tényezők relatív hibái ill. hibanégyzetei összegződnek. Ha eredményeinket hányados fejezi ki, akkor a számláló és nevező relatív hibái ill. hibanégyzetei összegződnek.

Példa: Tekintsük a nehézségi gyorsulás inga segítségével történő meghatározását. Az inga l hosszát és T lengésidejét Δl ill. ΔT hibával mérjük, majd a $g = 4\pi^2 \frac{l}{T^2}$ összefüggésből számoljuk a nehézségi gyorsulást. Ezen eljárás kvadratikus hibája a (4) egyenlet szerint

$$\Delta g = \pm \sqrt{\left(\frac{\partial g}{\partial l} \Delta l \right)^2 + \left(\frac{\partial g}{\partial T} \Delta T \right)^2},$$

vagy a differenciálás elvégzése után

$$\Delta g = \pm \sqrt{\left(\frac{4\pi^2}{T^2} \Delta l \right)^2 + \left(\frac{8\pi^2 l}{T^3} \Delta T \right)^2}.$$

Ebből a kvadratikus relatív hiba (5b) alapján

$$\frac{\Delta g}{g} = \pm \sqrt{\left(\frac{\Delta l}{l} \right)^2 + \left(2 \frac{\Delta T}{T} \right)^2}.$$

Hasonlóképp (3) és (5a) felhasználásával a maximális relatív hibára

$$\frac{\Delta g}{g} = \pm \left(\frac{\Delta l}{l} + 2 \frac{\Delta T}{T} \right) \quad \text{adódik.}$$

4. Egyenes illesztés a legkisebb négyzetek módszerével

Gyakori feladat, hogy a mért $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ adatokat összekapcsoló $y = a + bx$ lineáris egyenlet a és b együtthatóit kell meghatároznunk, ahol az y_i -t meghatározó egyes mérések hibája σ_i , és az egyszerűség kedvéért feltesszük, hogy az x értékek pontosak. A feladat a **legkisebb négyzetek módszerével** oldható meg. Az összes mérési adat figyelembe vételével olyan a és b értékeket keresünk, melyekre az $y = a + bx$ egyenestől való eltérés a lehető legkisebb. Ezen követelmény teljesülését a

$$\chi^2 = \sum_{i=1}^n \frac{[y_i - (a + bx_i)]^2}{\sigma_i^2} \quad (6)$$

kifejezés minimalizálásával biztosítjuk, melynek feltétele, hogy a és b kielégítse a

$$0 = \frac{\partial \chi^2}{\partial a} = -2 \sum_{i=1}^n \frac{y_i - a - bx_i}{\sigma_i^2} \quad \text{és}$$

$$0 = \frac{\partial \chi^2}{\partial b} = -2 \sum_{i=1}^n x_i \frac{y_i - a - bx_i}{\sigma_i^2}$$

egyenleteket.

Vezessük be az $S = \sum \frac{1}{\sigma_i^2}$, $S_x = \sum \frac{x_i}{\sigma_i^2}$, $S_y = \sum \frac{y_i}{\sigma_i^2}$,

$S_{xx} = \sum \frac{x_i^2}{\sigma_i^2}$, $S_{xy} = \sum \frac{x_i y_i}{\sigma_i^2}$ és $\Delta = SS_{xx} - (S_x)^2$ jelöl-

léseket, és fejezzük ki segítségükkel a -t és b -t. A keresett megoldás:

$$a = \frac{S_{xx}S_y - S_x S_{xy}}{\Delta} \quad \text{és} \quad b = \frac{SS_{xy} - S_x S_y}{\Delta} \quad (7)$$

Meg kell még határozni a és b bizonytalanságát, melyhez minden pontpár hozzájárul. A hibaterjedésnek megfelelően a és b szórása

$$\sigma_a^2 = \sum \left(\frac{\partial a}{\partial y_i} \right)^2 \sigma_i^2 \quad \text{ill.} \quad \sigma_b^2 = \sum \left(\frac{\partial b}{\partial y_i} \right)^2 \sigma_i^2.$$

$$\text{Mint} \quad \frac{\partial a}{\partial y_i} = \frac{S_{xx} - S_x x_i}{\sigma_i^2 \Delta} \quad \text{ill.} \quad \frac{\partial b}{\partial y_i} = \frac{S x_i - S_x}{\sigma_i^2 \Delta}, \quad a$$

mérési pontokra való összegzés után a és b szórása (varianciája)

$$\sigma_a^2 = \frac{S_{xx}}{\Delta} \quad \text{és} \quad \sigma_b^2 = \frac{S}{\Delta}. \quad (8)$$

Az illesztés jóságát a $Q\left(\frac{n-2}{2}, \frac{\chi^2}{2}\right)$ kifejezés adja

meg (ahol Q a *nem-teljes gamma függvény*). Ez annak a valószínűsége, hogy a fenti módon számított χ^2 érték (véletlen folyamatok során) megvalósul. Ha $Q > 0,1$, akkor az illesztés eredménye hihető. Amennyiben $Q > 0,001$, az illesztés elfogadható lehet, de csak akkor, ha a hibák nem követik a normáloszlást, vagy kissé alulbecsültek. Végül, $Q < 0,001$ esetben a lineáris viselkedés és/vagy az illesztési eljárás jogosan megkérdőjelezhető.

- Nem-teljes gamma függvény: $Q(a, x) = \Gamma(a)^{-1} \int_x^\infty t^{a-1} e^{-t} dt$,

ahol $\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt$ a teljes gamma függvény.

- Ha az egyes mért y_i értékek σ_i hibája nem ismert, megpróbálkozhatunk becslésével. Ennek feltétele, hogy minden pontra vonatkozóan a standard eltérés azonos legyen ($\sigma_i = \sigma$), továbbá hogy az x és y közt feltételezhető lineáris kapcsolat jól illeszkedjen a mérési adatokra. Ekkor a $\sigma_i = 1$ feltevéssel számíthatjuk a és b értékét, ill. a χ^2 -et, majd ezek ismeretében az egyes mérések szórását $\sigma^2 = \frac{\chi^2}{n-2}$ -vel közelítjük. A paraméter értékek szórását pedig úgy kapjuk, hogy a $\sigma_i = 1$ feltételezéssel a (8) egyenletek alapján számított σ_a^2 ill. σ_b^2 értékeket megszorozzuk $\frac{\chi^2}{n-2}$ -vel. Ez az eljárás nem teszi lehetővé az illesztés jóságának független meghatározását.
- Vegyük észre, hogy számos látszólag különböző probléma lefordítható a fenti egyenes illesztésre. Ilyenek pl. az $\eta^j = a + b\xi^m$ ill. $\eta = \alpha \exp(bx)$ függvénykapcsolatok, melyeknél az $y = \eta^j$ és $x = \xi^m$, ill. az $y = \ln(\eta)$ és $a = \ln(\alpha)$ helyettesítéseket alkalmazzuk.

Irodalom:

- A hibaszámítás elméleti kérdéseiről ld. Jánossy Lajos *A valószínűségelmélet alapjai és néhány alkalmazása* című könyvét (Tankönyvkiadó, Budapest, 1967).
- A hibaszámítással kapcsolatos problémák korszerű, gyakorlati megközelítése (összefoglalók, alkalmas numerikus módszerek, kész programok, szubrutinok) W. H. Press, B. P. Flannery, S. A. Teukolsky és W. T. Vetterling, *Numerical Recipes, The Art of Scientific Computing* című könyvében található (Cambridge University Press, 1986).