

HIBASZÁMÍTÁS

A mérési eredmény hibája

Hiba: A kísérlet jól meghatározott (reprodukálható) körülmények között játszódik le, lefolyását azonban sok apró, külön-külön nehezen figyelembe vehető tényező is befolyásolja. A zavaró tényezőket összefoglaló néven zajnak is nevezik. Ezek az egyes kísérletek során más-más összhatást eredményeznek. Ez gyakorlatilag a beavatkozásra adott válaszban, illetve a mért eredmény ingadozásában nyilvánul meg. A mért eredmény értékében bekövetkező ezen ingadozásokat nevezik **hibának**.

A hiba lehet



Szisztematikus (rendszeres)



Véletlen

A **szisztematikus (rendszeres) hibák** olyan okok következményei, amelyek rendszeresen, meghatározott irányban fejtik ki hatásukat. Ezek lehetnek a „legveszélyesebbek”, mert nehezen ismerhetők fel. Ha már felismertük a szisztematikus hibát, annak mennyiségi figyelembevétele az adott mérőműszer ismételt kalibrációjával lehetséges. Amennyiben szisztematikus hibára gyanakszunk, célszerű a szóban forgó mennyiséget más kísérleti módszerrel is meghatározni.) A rendszeres hibát torzításnak is nevezik.

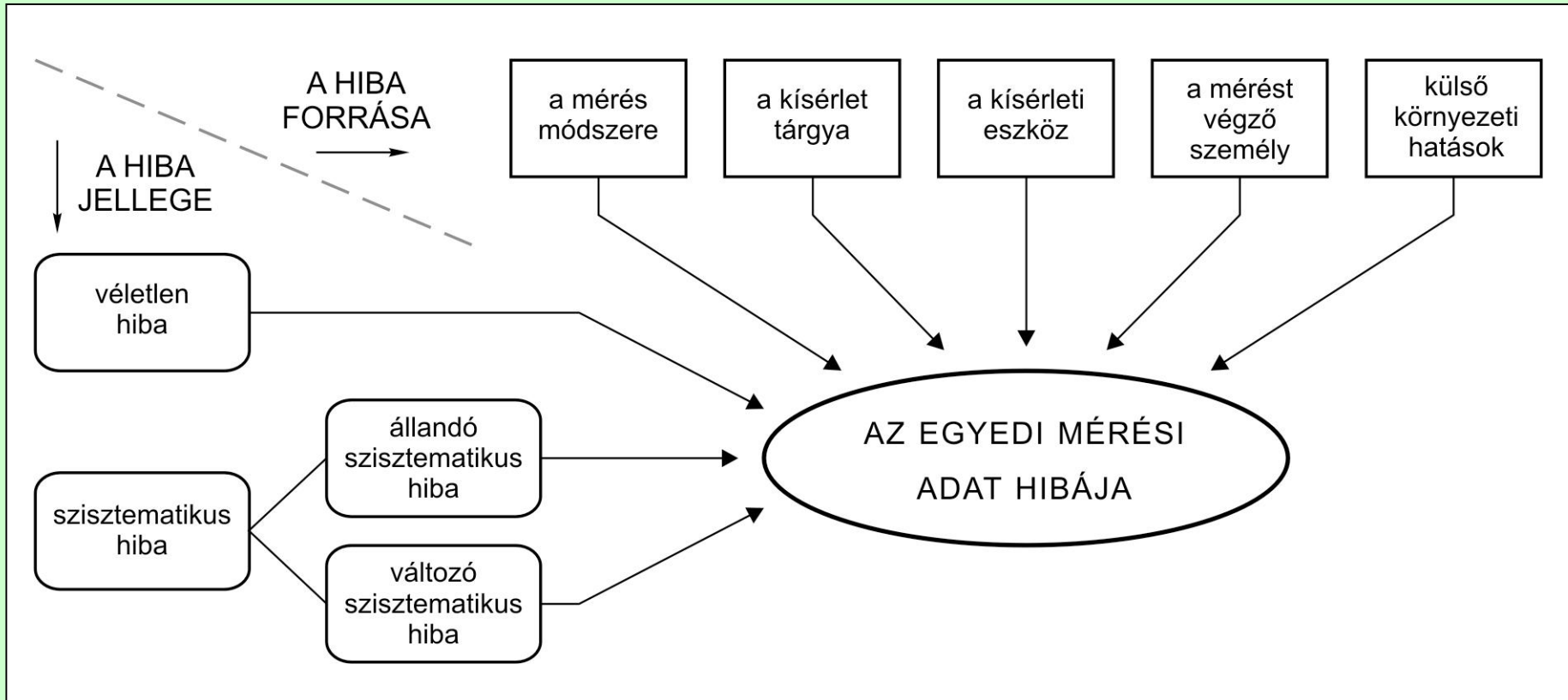
Véletlen hibáknak azokat a hibákat nevezzük, amelyek nem rendszeresen és nem meghatározott irányban lépnek fel, keletkezésük oka ismeretlen, illetve ezeket az okokat nem tudjuk közvetlenül figyelembe venni.

A szisztematikus és a véletlen hibák nagyszámú ún. elemi hibából állhatnak.

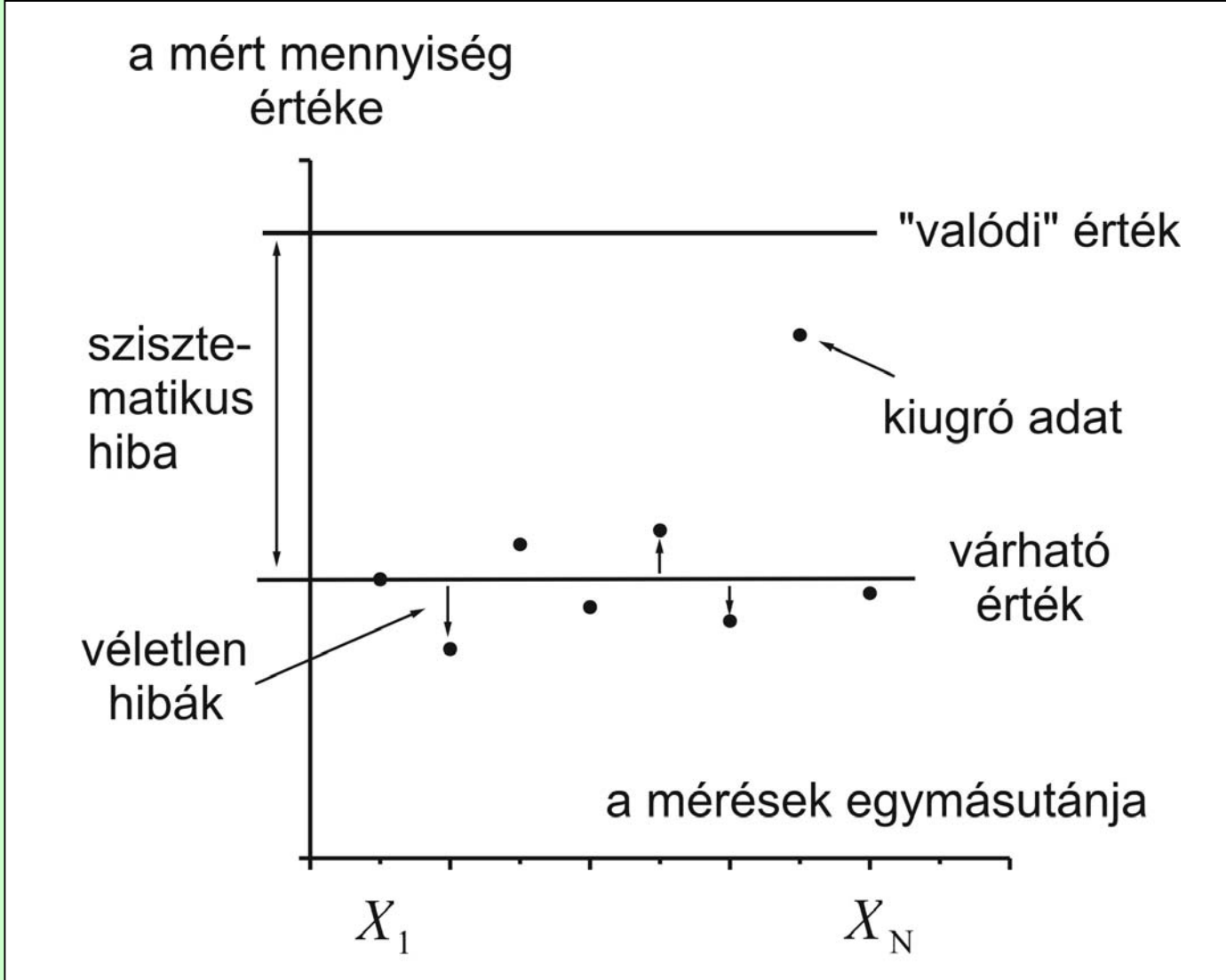
A mérési eredmények kiértékelésekor mindig világosan különbséget kell tennünk a szóban forgó mérés eredménye és a mérendő mennyiségek valódi értékei között. A kapcsolatot e két mennyiség között az eredmény **várható értéke** biztosítja.

(Valamely valószínűségi változó vagy valamely eloszlás (elméleti) középértékét nevezzük várható értéknek.)

Az esetek többségében – mintegy munkahipotézisként – kénytelenek vagyunk a várható értéket a meghatározandó mennyiség valódi értékének elfogadni.



A mérési hiba lehetséges forrásai és a hibák jellege



A valódi és a várható érték, valamint a szisztematikus (rendszeres) és véletlen hiba kapcsolatának szemléltetése

Abszolút hiba és hibakorlát

Az abszolút hiba (megállapodásszerűen) a mérendő mennyiség valódi értékének (X) és a méréssel meghatározott értékének (x) az eltérése egymástól:

$$\Delta = |X - x|$$

vagy

$$X = x \pm \Delta$$

Δ dimenziója megegyezik X és x dimenziójával.

A hibát meghatározni az esetek többségében teljesen lehetetlen, mivel a mérendő mennyiség valódi értékének pontos nagyságát nem ismerjük. Ilyenkor nem tehetünk mást, mint megbecsüljük, hogy ez az abszolút hiba milyen értéknél nem lehet nagyobb, vagyis megadunk egy ún. **abszolút hibakorlátot**, amire

$$\Delta = |X - x| \leq h$$

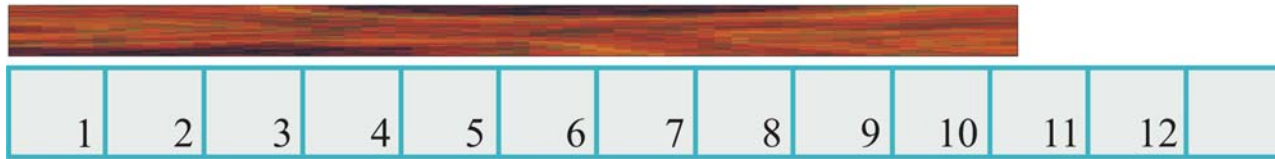
Adott mérőeszköz vagy -berendezés abszolút hibakorlátja (h) előre ismert lehet.

Legkisebb osztás (least count) : az eszközön jelölt legkisebb osztásrész.

Adott mérőeszköz vagy -berendezés abszolút hibakorlátja (instrument limit of error, ILE) : a „pontosság” amennyire az eszköztől az értékeket le lehet olvasni.

Értéke egyenlő a legkisebb osztással vagy kisebb annál.

a)



b)



c)



Az alábbi, centiméterben skálázott mérőeszközökre adjuk meg a legkisebb osztásrészt, a mérőeszköz abszolút hibakorlátját, illetve olvassuk le a rudak hosszát!

Relatív hiba:

$$\Delta\% = \frac{\Delta}{x} \cdot 100$$

A relatív hiba megadásához az abszolút hibát kellene ismernünk, tehát itt is fennállnak azok a nehézségek, amelyeket az abszolút hiba meghatározásával kapcsolatban már elmondtunk. Ezért ebben az esetben is meg kell elégednünk egy korlát megadásával ($h\%$), amelyre vonatkozóan szavatolhatjuk, hogy ezt a relatív hiba nagysága nem haladja meg:

$$\Delta\% \leq h\%$$

Amikor a gyakorlatban abszolút, illetve relatív hibát mondunk, ezen mindig a megfelelő hibakorlátot kell érteni.

Ld. még.: „Pontossági osztályok”

Hibaterjedés

A közvetlenül mért mennyiségeket általában különböző összefüggések alapján újabb mennyiség kiszámítására használjuk. Fontos annak ismerete, hogy a méréskor jelentkező hibák hogyan hatnak a számítással kapott mennyiségek pontosságára, vagyis hogyan „terjednek” a hibák.

Ha a meghatározandó mennyiség (y) az x_1, x_2, x_3, \dots közvetlenül mért mennyiségekből számítható az

$$y = f(x_1, x_2, x_3, \dots)$$

összefüggés alapján.

Akkor az egyes mennyiségek abszolút hibájából az eredményben várható hibát az alábbi módon kapjuk meg:

$$\Delta y = |\Delta y_1| + |\Delta y_2| + |\Delta y_3| + \dots$$

Ahol $\Delta y_1 = \frac{\partial y}{\partial x_1} \Delta x_1$; $\Delta y_2 = \frac{\partial y}{\partial x_2} \Delta x_2$; ... stb.

Használható a közepes hibákkal megadott összefüggés:

$$\Delta y = \left| \frac{\partial y}{\partial x_1} \cdot \Delta x_1 \right| + \left| \frac{\partial y}{\partial x_2} \cdot \Delta x_2 \right| + \left| \frac{\partial y}{\partial x_3} \cdot \Delta x_3 \right| + \dots$$

Vagy a Gauss féle hibaterjedési törvény:

$$\Delta y = \sqrt{\left(\frac{\partial y}{\partial x_1} \cdot \Delta x_1 \right)^2 + \left(\frac{\partial y}{\partial x_2} \cdot \Delta x_2 \right)^2 + \left(\frac{\partial y}{\partial x_3} \cdot \Delta x_3 \right)^2 + \dots}$$

Az értékes (helyes) jegyek száma

A tizedes tört alakjában felírt szám külső képe alapján megítélhetjük annak pontosságát az ún. **értékes (v. helyes) jegyek** megszámlálása útján. A közvetlenül mért adat (leolvasott érték) jegyeit a következőképpen számoljuk össze: balról az első nem nulla jegynél kezdjük a számolást, és addig folytatjuk, amíg számjegyük van. Ennél a módszernél mindegy, hogy tizedes tört vagy normálalak formájában írjuk fel az adatot.

Digitális műszernél egyértelmű a jegyek száma, hiszen azt a kijelzés szabja meg.

Analóg műszernél az utolsó jegyet mindig becsüljük, és ezt a becsült jegyet is értékesként kezeljük.

A számítások során az alábbi gyakorlati szabályok alkalmazhatóak:

1. Úgy járunk el a számítások során, hogy a részeredményeket (legalább) két jeggyel többre adjuk meg, mint a kiindulási adatok értékes jegyeinek száma, a végeredményt pedig annyi jegyre, mint ahány értékes jeggyel a kiindulási adat bír (legfeljebb eggyel többre). Részeredményt sohasem kerekítünk, a végeredmény megadásánál pedig a kerekítés szabályai alkalmazhatók.

2. A másik lehetőség, ha a mérési eljárásról tudjuk (pl. irodalmi adatok alapján), hogy milyen pontosságot tesz lehetővé, vagy a meghatározandó fizikai mennyiségről tudjuk, hogy az általában milyen pontossággal határozható meg, akkor az ismert relatív hibakorlátból indulunk ki. Pl. ha 1%-os a relatív hibakorlát, akkor az eredményben ezt az 1%-ot képviselő helyi értéken lévő jegyet még elfogadjuk, az ez alatt lévő helyi értéket elhagyjuk a kerekítés szabályainak figyelembe vételével.

Feladat: Mennyi az értékes jegyek száma az alábbi
(mérő)számokban?

- (i) 0.00035 (ii) 0.12700 (iii) 3.8×10^6 (iv) -124.090×10^{-17}
(v) 00035.47

A mérés megbízhatósága

A jelenségek megfigyelésének, illetve a kísérletek végzésének végző célja, hogy a kapott eredményeket általánosítsuk minden olyan esetre, amely a vizsgált jelenség körébe tartozik. A jelenségeknek ezt a teljes körét a matematikai statisztika **alapsokaságnak** nevezi (a biometriában a populáció elnevezés használatos).

Az alapsokaságot nem tudjuk teljesen megismerni, mivel az összes lehetséges eseteknek csak korlátozott részét vagyunk képesek megfigyelni. Ezt a megfigyelhető részt nevezzük **statisztikai mintának**, és ebből próbálunk az alapsokaság jellemzőire következtetni.

Minden kísérlet vagy mérés matematikai statisztikai értelemben mintavételt jelent.

A mért eredmény értéke általában véletlen eseményektől is függ.

Hogy ezek hatását kiküszöbölhessük, illetve figyelembe vehessük, **a mérés eredményét valószínűségi változónak kell tekintenünk.**

Valószínűségi változónak az olyan mennyiséget nevezzük, amelynek értéke véletlen eseményektől függ, pl. a gáz egyes molekuláinak sebessége; az egyes radioaktív atomok élettartama stb.

Minden **folytonos valószínűségi változóval** kapcsolatban nagyon lényeges kérdés, mi annak a valószínűsége, hogy az egy bizonyos $x + dx$ intervallumba eső értéket vegyen fel. Ezt a **valószínűségi változó sűrűségfüggvénye** írja le. Aszerint, hogy milyen ez a függvény, beszélünk normális (Gauss-féle), Poisson, egyenletes, exponenciális, F , t , stb. eloszlásokról.

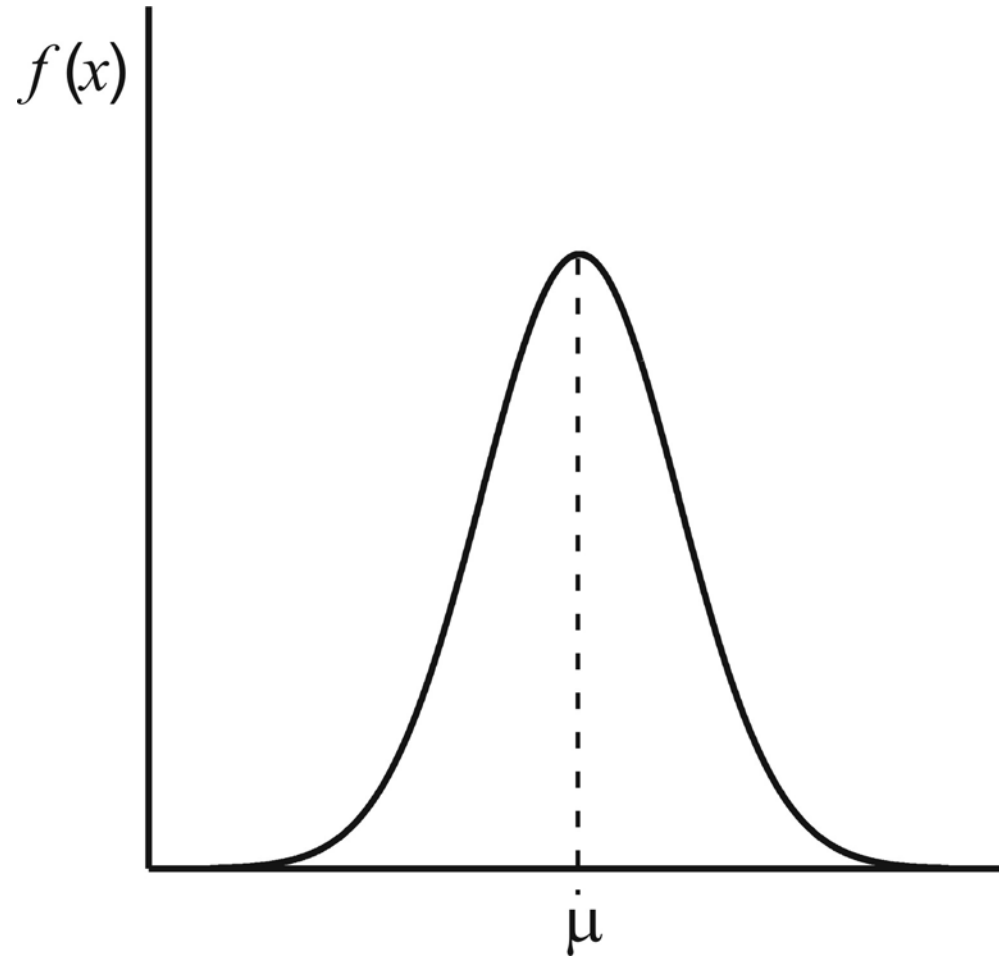
A természetben akkor találkozunk normális eloszlással, ha sok, egymástól független valószínűségi változó hatása összegeződik, feltéve, hogy az összeg minden egyes tagjának ingadozása kicsi az egész összeg ingadozásához képest.

A mérési hibák zöme normális eloszlású. A normális eloszlás nagy előnye, hogy matematikailag jól kezelhető. A statisztikai becslések és próbák nagy része is ezen az eloszláson alapszik. A normális eloszlású valószínűségi változó sűrűségfüggvénye:

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

A normális eloszlású valószínűségi változó ugyanolyan valószínűséggel veszi fel a μ -nél nagyobb és kisebb értékeket, a μ -től lényegesen eltérőket pedig jóval kisebb valószínűséggel, mint a μ -höz közeliakat. μ és σ tetszőleges számok ($\sigma > 0$), az eloszlás két paramétere. μ a valószínűségi változó várható értéke, σ a szórása. Ez a két paraméter a normális eloszlást teljesen meghatározza. Ezért, ha olyan természeti jelenséget akarunk leírni, amely normális eloszlást követ, elég a várható értéket és a szórást meghatározni.

Mivel egy mennyiség valódi értékét nem ismerjük, ezért a mért mennyiség „valódi” értékének a várható értéket (μ) tekintjük



A normális eloszlás sűrűségfüggvényének alakja

Normális eloszlású valószínűségi változó várható értékére legjobb becslés a számtani középérték:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

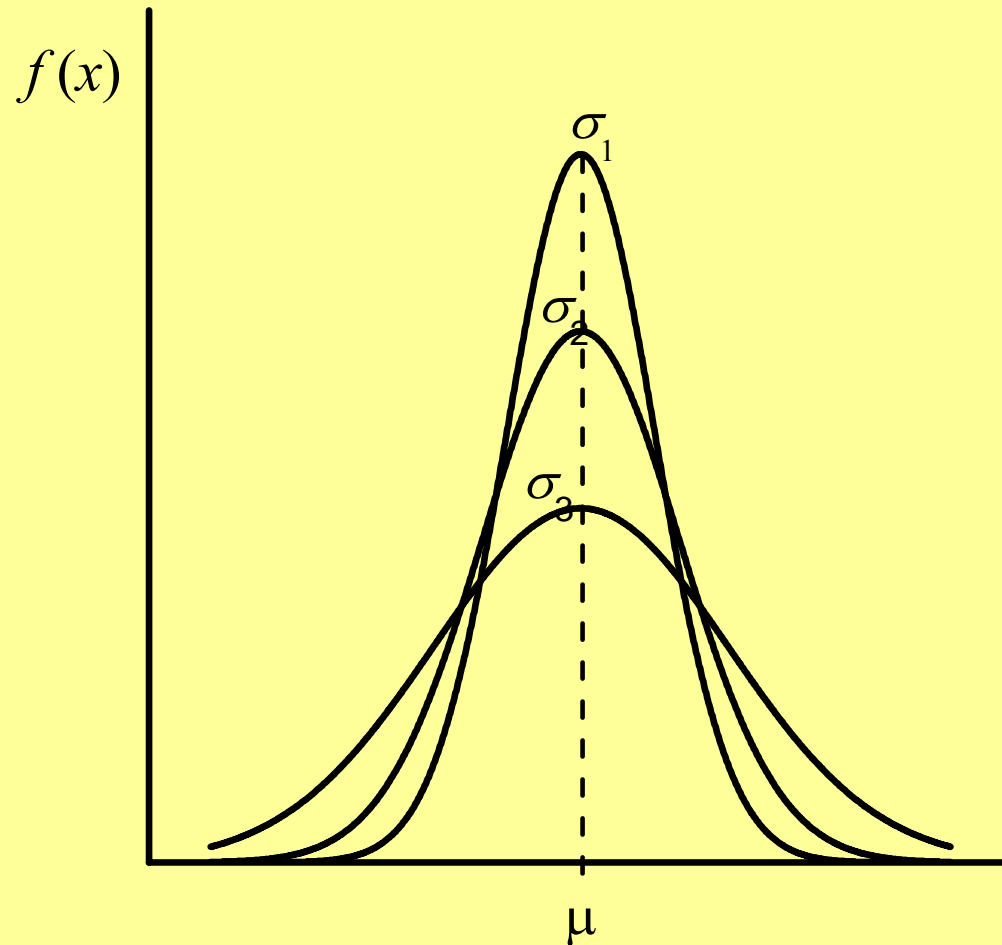
A szórás (σ)

A szórás a mérés reprodukálhatóságára és arra vonatkozóan ad felvilágosítást, hogy a mérési eredmények átlagosan mennyire térnek el a középértéktől. Minél kisebb a szórás, annál meredekebb a haranggörbe, az egyes mérési eredmények annál kevésbé ingadoznak a középérték körül, és így annál megbízhatóbb az észlelési sorozat. Ha a szórás nagy, akkor az észlelt adatok többsége a középértéktől jelentősen eltér.

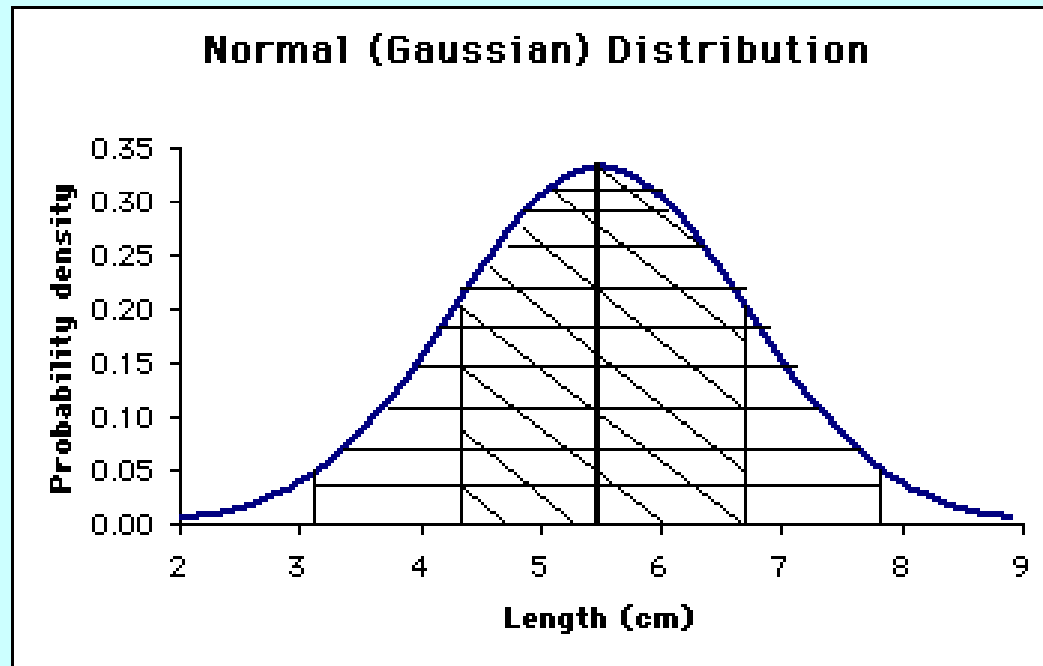
σ pontos meghatározásához szintén ismerni kellene a sűrűségfüggvényt a $(-\infty; \infty)$ intervallumban, valamint μ értékét.

Mivel μ -t nem ismerjük, (hiszen csak becsültük), σ -t is csak becsüljük a korrigált tapasztalati (empirikus) szórásnégyzettel:

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



Ugyanakkora várható értékű, de különböző szórású normális eloszlások



A Gauss-féle sűrűségfüggvényt fel lehet használni annak megállapítására, milyen valószínűséggel esik egy mérési eredmény valamely adott intervallumba. Ehhez integrálni kell a sűrűségfüggvényt az adott határok között.

Fontos lehet annak ismerete, hogy milyen intervallumba esik a mért mennyiség várható értéke pl. 95 vagy 99 %-os valószínűséggel, vagyis milyen határok között lesz a fenti integrál értéke 0,95, illetve 0,99. (Ezt az adatot statisztikus biztonságnak nevezik). Ez az intervallum a konfidencia (megbízhatósági) intervallum, amelyet σ segítségével is meg lehet adni

$$\bar{x} \pm k\sigma$$

alakban.

a szórás együtthatója	statisztikai biztonság
k	%
1,00	68,26
1,64	90,00
1,96	95,00
2,58	99,00
3,00	99,73
3,29	99,90
3,50	99,95

A szórás k együtthatója és a konfidenciaszint kapcsolata nagy mintaelemszám ($n \geq 30$) esetén

Megjegyzés: 30 alatti mintaelemszámnál a Student-féle t -eloszlás írja le a viszonyokat.

Statisztikai próba

Gyakran fordul elő, hogy a becslések „jóságát” (biztonságát), kell ellenőriznünk, vagy két mérési sorozat jellemzőit akarjuk összehasonlítani, esetleg két valószínűségi változó egymástól való függetlenségét vagy függetlenségét kívánjuk bizonyítani. Ilyenkor statisztikai feltevés (hipotézis-) vizsgálatot alkalmazunk. Ez abból indul ki, hogy egy bizonyos állítást (az alapsokaságra vonatkozóan) érvényesnek tételez fel. Ezt a feltevést nullhipotézisnek nevezzük és H_0 -val jelöljük. Ennek helyességét kell az alapsokaságból vett minta (megfigyelési értékek) adatai alapján bizonyítanunk.

A statisztikai hipotézisek ellenőrzésénél kétféle hibát követhetünk el:

- a) elvetünk egy feltevést, jóllehet az igaz (elsőfajú hiba),
- b) megtartunk egy hipotézist, bár az hamis (másodfajú hiba).

A leggyakrabban alkalmazott próbák a t -, F - és a χ -próba. A t -próbát pl. a mérések középértéke és a valódi érték vagy két mérési sorozat eredményeinek összehasonlítása és a mérési eredmény hibahatárainak megadása során alkalmazzuk. Az F -próba pl. két mérési sorozat szórásának összehasonlítására alkalmas. A χ -próba segítségével dönthetjük el, hogy mérési adataink normális eloszlásúak-e.

A t -próba esetén a próbastatisztika:

$$t = \frac{\bar{x} - \mu}{S^*} \sqrt{n}$$

Az eredmény hibahatárainak megadása (konfidencia intervallum számítása kis mintaelemszám esetén)

Általában nem ismerjük a mérés várható értékét, ennek ellenére a t -próba segítségével tudunk valamit mondani eredményeink megbízhatóságáról. A t -próba-statisztika képletéből a várható érték:

$$\mu = \bar{x} \pm t_{\alpha} \frac{S^*}{\sqrt{n}}$$

($1-\alpha$) %-os biztonsággal állíthatjuk, hogy a várható érték \bar{x} -tól legfeljebb $t_{\alpha} \frac{S^*}{\sqrt{n}}$ értékkel tér el.

A kiugró érték(ek) ellenőrzése

A kiugró (extrém), adatok a számított középértéket és a szórást torzíthatják, ezért szükségessé válhat azok kizárása a további vizsgálatokból. Erre alkalmas pl. a Gauss-féle g statisztika:

$$\frac{|x_{\text{extr}} - \bar{x}|}{S^*} = g$$

ahol x_{extr} a „gyanús” kísérleti eredmény, \bar{x} a többi adat átlaga a kiugró értéket figyelmen kívül hagyva, S^* pedig ezekből a mintaelemekből számított korrigált tapasztalati szórás.

