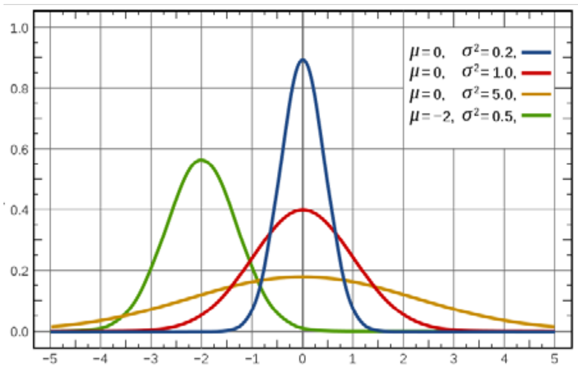
1. Gauss-eloszlás, természetes szórás

A Gauss-eloszlásnak megfelelő függvény:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right),$$

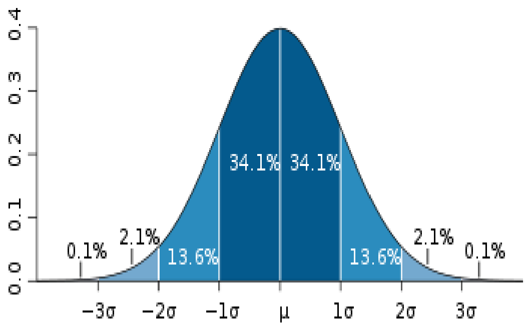
amely egy σ szélességű, μ középpontú, 1-re normált (azaz a teljes görbe alatti terület 1) görbét ír le.



A természetben a centrális határeloszlás tétel miatt általában a mérhető mennyiségeknek Gauss-eloszlása van, azaz:

Ha egy mérhető mennyiség várható/elméleti értéke μ, akkor a mért értékek eloszlása egy μ körüli Gauss-görbe lesz, melynek szélessége arányos a mérés hibájával.

Ha például egy 5 m magas fa magasságát megmérjük 100-szor, minden esetben 10 cm mérési bizonytalanságban, akkor a legtöbbször 5 m körüli értékeket kapunk természetesen, de az esetek jó részében a hibával összemérhető, esetenként annál nagyobb eltérést tapasztalunk. A jobb oldali ábrán látható, hogy hány σ távolságnál nagyobb eltérésnek mekkora a valószínűsége. Eszerint például az esetek 31.8%-ban (100%-2×34,1%) kapunk 1σ-nál nagyobb eltérést, míg 3σ-nál nagyobb eltérést csak az esetek 0.1%-ban.



2. Mérések hibája, szórása

Legyen mérések egy sorozata az $\{x_i\}$ rendezett N-elemű számsor (amelynek minden eleme a mérés egy-egy eredménye). Ekkor a mérési eredmények átlaga:

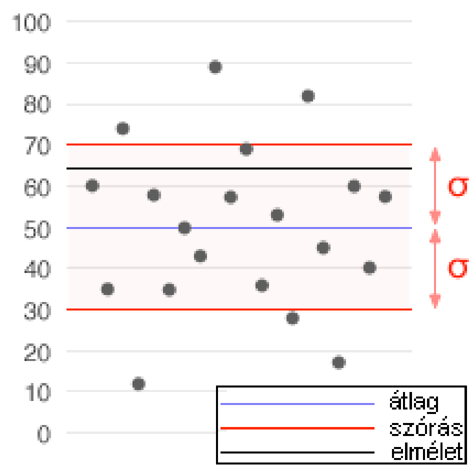
$$\overline{x} = \frac{x_1 + x_2 + \cdots + x_N}{N} = \frac{1}{N}\sum_{i=1}^{N} x_i\,.$$

A mérési eredmények szórása (ez tulajdonképpen a mérés pontosságát jelenti) pedig:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})^2},$$

Vegyük észre, hogy ha a fenti képletben nem lenne négyzet, az eredmény egzaktul nulla lenne (lásd az átlag definícióját).

Látható továbbá, hogy a szórás fordítottan arányos a mérések számának gyökével, azaz ha egy mérést N-szer annyiszor végzünk el, a szórás a $\sqrt{N}$-ed részére csökken.



Illusztrációul lásd a jobbra lévő ábrát, amely 10 és 90 közötti mérési eredmények szórását mutatja. Az átlag itt 50, a szórás 20.

Fontos megérteni, hogy a Gauss-eloszlás itt azt jelenti, hogy ha megvizsgálnánk, egy adott érték körül egy kis tartományban hány mérési eredmény volt, és ezt a számot (az ezt az eredményt adó mérések számát) ábrázolnánk az adott érték függvényében, akkor kapnánk Gauss-eloszlást. Szemléletesen, ha a fenti ábrát a függőleges tengelyre „vetítenénk”, akkor a kapott alak egy Gauss-görbe lenne. Látható az is, hogy a 20 mérésből 6 esetben kaptunk 1σ-nál nagyobb eltérést, ami körülbelül a fenti 31.8%-ot jelenti.

3. Mérések elméleti leírása, χ^2-próba

Amennyiben egy mérés-sorozattal egy elméleti eredményt szeretnénk megcáfolni vagy alátámasztani, a legjobb eszközünk az úgynevezett χ^2-próba. Ez úgy végezhető el, hogy kiszámítjuk a mérési eredményekhez ($\{x_i\}$) rendelt χ^2 értéket:

$$\chi^2 = \sum_{i=1}^{N} \left(\frac{x_i - x_{\text{elmélet}}}{\Delta x_i} \right)^2$$

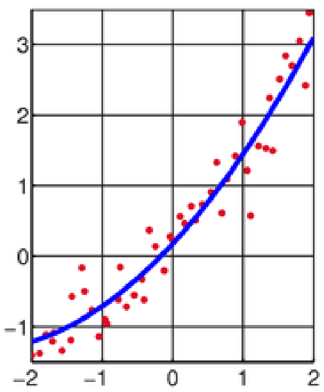
ahol Δx_i az egyes mérések hibája. Amennyiben $\chi^2<N$, akkor az elméleti érték az átlagos hibahatáron belül van. A fenti ábrán az átlagos χ^2 érték láthatólag a hibánál kisebb, tehát az említett feltétel itt teljesül. Ez azt jelenti, hogy a mérés nem zárja ki az elméletet.

4. Illesztések

Amennyiben nem ugyanazt a mérést végezzük el többször egymás után, hanem a mérési pontjaink adat-párok, avagy egy mennyiséget mérünk a másik *függvényében* (azaz például a mágneses teret a forrástól való távolság függvényében, vagy egy minta radon-tartalmát az idő függvényében), akkor egy kicsit módosítva alkalmazzuk a χ^2-próbát (khi-négyzet próbát). Ekkor egy elméleti *függvényünk* van, legyen ez $f(x)$. Az adatpontjaink legyenek az $\{f_i, \Delta f_i, x_i\}$ értékek, ahol f_i a mért mennyiség (mágneses tér, radon-tartalom stb.), Δf_i a hibája, és x_i a változó, amelynek függvényében mérünk (idő, távolság stb.). Ekkor:

$$\chi^2 = \sum_{i=1}^{N} \left(\frac{f_i - f(x_i)}{\Delta f_i} \right)^2$$

Itt $f(x_i)$ az elméleti függvény x_i-ben vett értéke. Ekkor a χ^2 értéke alapján megmondhatjuk, hogy **az adott elmélet mellett mekkora a valószínűsége, hogy az adott mérési pontok jöjjenek ki**. Ezt a valószínűséget az Excel „KHI.ELOSZLÁS” függvényével számíthatjuk ki, melynek első paramétere a χ^2 , a második pedig a szabadsági fokok száma, azaz a mérési pontok száma. Amennyiben ez **a valószínűség 0.1% alatt van**, azt mondjuk, hogy **a mérés alapján az elméletet ki lehet zárni**. Egyéb esetben a mérés megerősíti az elméletet.



További módosítást jelent, ha az elméleti függvénynek van *paramétere*, azaz például nem az az elméleti jóslatunk, hogy radonkoncentráció a 0.5·exp(-32 Hz·*t*) függvény szerint fog változni, hanem hogy az *A*·exp(*-B·t)* függvény szerint. Ekkor a χ^2 próba alapján meghatározhatjuk a modell paramétereit illesztéssel. Ekkor a χ^2 minimális értékét keressük meg a paraméterek változtatásával. Azok a paraméterek, ahol a χ^2 **minimális**, az **optimális paraméterek**. Ezeket lehet a mérésből meghatározni. Ha például az $f(x)$ függvénynek van egy a paramétere, azaz $f(x,a)$ valójában, akkor az alábbi függvény minimumát keressük:

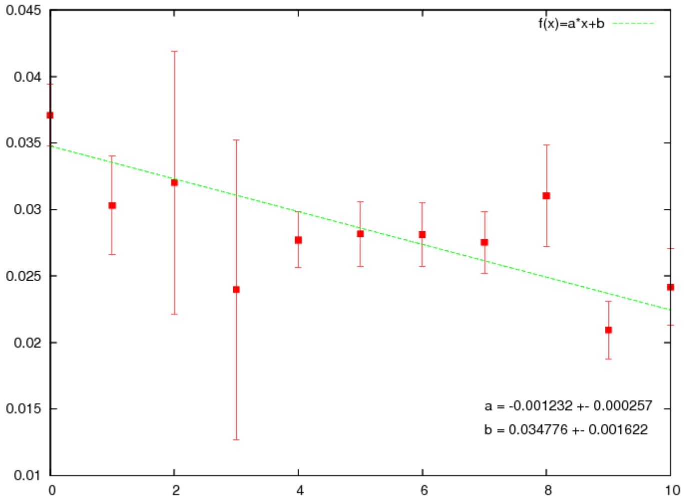
$$\chi^2(a) = \sum_{i=1}^{N} \left(\frac{f_i - f(x_i, a)}{\Delta f_i} \right)^2$$

Ezt nevezzük **illesztésnek**. Amennyiben a minimális χ^2-ből és a pontok számából számolt valószínűség nagyobb 0.1%-nál, akkor az elmélet leírja a mérési adatokat az optimális *a* paraméter mellett. Fontos megemlíteni, hogy ekkor a „KHI.ELOSZLÁS" Excel függvénybe az adatpontok számát a paraméterek számával csökkentve kell beírni. Ugyanis két mérési pontra értelmetlen egy kétparaméteres függvényt illeszteni – ez mindig egzaktul lehetséges (két ponton pontosan egy egyenes megy át).
A szabadsági fokok száma tehát a mérési pontok száma mínusz az illesztési paraméterek száma.

Ekkor *a*-t is a mérés eredményének tekintjük, azt mondjuk, hogy az adatokkal megmértük az adott *f(x,a)* elméleti görbe *a* paraméterét. Alább látható egy adathalmaz és az őt leíró minimalizált χ^2-tel rendelkező görbe. Ez a gnuplot segítségével készült, a

```
f(x) = a*x+b
fit f(x) 'adatok.txt' using 1:2:3 via a,b
plot f(x), 'adatok.txt' using 1:2:3 with yerrorbars
```

parancsok segítségével



5. Egyenes illesztése

Általánosságban a fenti χ^2-minimalizáció rendkívüli probléma, bonyolult számítógépes algoritmusokkal végezhető csak el, azaz csak így találhatóak meg az optimális paraméterek. A fenti módszert azonban egyszerűen alkalmazhatjuk egyenes illesztésére.

Ekkor az illeszteni kívánt függvény (ahol *x* a változó, *y* a mérési eredmény):

$$f(x) = ax + b$$

azaz a minimalizálni kívánt mennyiség:

$$\chi^2(a,b) = \sum_{i=1}^{N} \left(\frac{f_i - ax_i - b}{\Delta f_i} \right)^2$$

Egy mennyiségnek akkor lehet minimuma egy adott paraméter-érték mellett, ha az aszerinti deriváltja nulla. Ez alapján a fenti χ^2-ből az optimális *a*-t és *b*-t úgy kaphatjuk meg, hogy az alábbi egyenleteket megoldjuk:

$$\frac{\partial}{\partial b}\chi^2(a,b) = 0 \qquad\qquad \frac{\partial}{\partial a}\chi^2(a,b) = 0$$

Mivel a fenti χ^2 polinom alakú, ezért a megoldást egyszerűen megkapjuk a lineáris egyenletrendszer megoldásából:

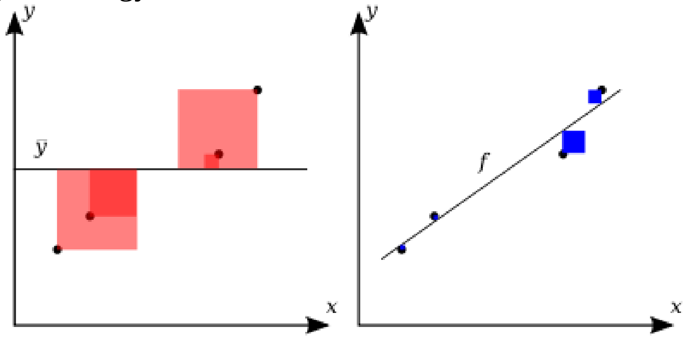
$$\boxed{\begin{aligned} a &= \frac{\sum_i \frac{x_i f_i}{\Delta f_i^2} \sum_i \frac{1}{\Delta f_i^2} - \sum_i \frac{x_i}{\Delta f_i^2} \sum_i \frac{f_i}{\Delta f_i^2}}{\sum_i \frac{x_i^2}{\Delta f_i^2} \sum_i \frac{1}{\Delta f_i^2} - \sum_i \frac{x_i}{\Delta f_i^2} \sum_i \frac{x_i}{\Delta f_i^2}} \\ b &= \frac{\sum_i \frac{x_i}{\Delta f_i^2} \sum_i \frac{x_i f_i}{\Delta f_i^2} - \sum_i \frac{x_i^2}{\Delta f_i^2} \sum_i \frac{f_i}{\Delta f_i^2}}{\sum_i \frac{x_i}{\Delta f_i^2} \sum_i \frac{x_i}{\Delta f_i^2} - \sum_i \frac{x_i^2}{\Delta f_i^2} \sum_i \frac{1}{\Delta f_i^2}} \end{aligned}}$$

6. Egyenes illesztés Excellel

Excellel való illesztés esetén nem a fent leírt χ^2 mennyiséget minimalizáljuk, hanem az úgynevezett R-négyzet (R^2) értékét maximalizáljuk. Ennek definíciója:

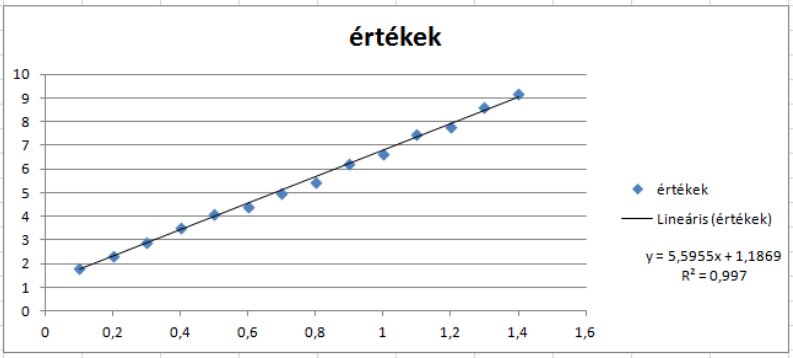
$$R^2 \equiv 1 - \frac{SS_{\text{err}}}{SS_{\text{tot}}}.$$

ahol $SS_{err} = -\Sigma(f_i - f(x_i))^2$ és $SS_{tot} = /\Sigma(f_i - f_{átlag})^2$, tehát az egyik a függvénytől való négyzetes eltérés, a másik az átlagtól való négyzetes eltérés. Tehát R értéke nulla, ha a függvény csak annyira illeszti jól az adatokat, mint az átlag. Ha az R=1, akkor az összes adatpont éppen az illesztett függvényen van. Ezt illusztrálja az alábbi ábra, ahol a piros négyzetek területének összege az SS_{tot}, míg a kék négyzeteké az SS_{err}.



Ennek az illesztésnek az előnye, hogy nincs hozzá szükség a mérési adatok bizonytalanságára, ugyanakkor a tudományos értéke is kisebb az előző fejezetben említett tesztnél. Egy Excelben elvégzett illesztést és annak R^2-értékét láthatunk alább:

változó	értékek
0,1	1,803103
0,2	2,342432
0,3	2,909427
0,4	3,526526
0,5	4,090796
0,6	4,400196
0,7	4,981401
0,8	5,452286
0,9	6,238171
1	6,608888
1,1	7,470231
1,2	7,775663
1,3	8,589584
1,4	9,179963



7. Hibaterjedés

Általában nagyon fontos ismernünk a mérési eredményeink hibáját. Többnyire a mérési eredményt nem közvetlenül, hanem egy számítás eredményeképpen kapjuk meg. Például ha egy asztal felületét akarjuk megmérni, akkor a szélességét és hosszúságát mérjük meg, majd a kettő szorzata lesz a terület. A kérdés az, hogy ha a szélesség és a hossz hibája egyaránt 10 cm, mekkora lesz a terület hibája? A kérdésre általánosságban érvényes választ kaphatunk a következő szakaszban.

Ha van egy *X* mennyiségünk (például az asztal felülete):

$$X = f(A, B, C, \ldots)$$

amely függ az *A*, *B*, *C* mennyiségektől (az asztal oldalszélességei, tehát ekkor X=AB), akkor az *A*, *B* és *C* hibájából megkaphatjuk *X* hibáját:

$$\Delta X = \left|\frac{\partial f}{\partial A}\right| \cdot \Delta A + \left|\frac{\partial f}{\partial B}\right| \cdot \Delta B + \left|\frac{\partial f}{\partial C}\right| \cdot \Delta C + \cdots$$

Ez általában felülbecsli a hibát, néha szokás a

$$\Delta X^2 = \left|\frac{\partial f}{\partial A}\right|^2 \cdot \Delta A^2 + \left|\frac{\partial f}{\partial B}\right|^2 \cdot \Delta B^2 + \left|\frac{\partial f}{\partial C}\right|^2 \cdot \Delta C^2 + \ldots$$

képlettel számolni. Mi az elsőt részesítjük előnyben. Az asztalos példa esetében $\Delta X = B\Delta A + A\Delta B$ lesz, vagy $\Delta X^2 = B^2\Delta A^2 + A^2\Delta B^2$.

Jelöljük mostantól egy adott mennyiség hibáját úgy, hogy: $\delta(X)$

relatív hibáját pedig így: $\mathcal{R}(X) = \delta(X)/X$

Ekkor az alapműveletekre könnyen kiszámíthatjuk az eredmény hibáját:

$$\begin{aligned}
\delta(a \cdot n) &= n\delta(a) \\
\delta(a \pm b) &= \delta(a) + \delta(b) \\
\delta(a \cdot b) &= |b|\delta(a) + |a|\delta(b) \\
\delta\left(\frac{a}{b}\right) &= \frac{1}{|b|}\delta(a) + \frac{|a|}{b^2}\delta(b) \\
\delta(a^n) &= na^{n-1}\delta(a)
\end{aligned}$$

$$\begin{aligned}
\mathcal{R}(a \cdot n) &= \mathcal{R}(a) \\
\mathcal{R}(a \pm b) &= \frac{\mathcal{R}(a)|a| + \mathcal{R}(b)|b|}{|a \pm b|} \\
\mathcal{R}(a \cdot b) &= \mathcal{R}(a) + \mathcal{R}(b) \\
\mathcal{R}\left(\frac{a}{b}\right) &= \mathcal{R}(a) + \mathcal{R}(b) \\
\mathcal{R}(a^n) &= n\mathcal{R}(a)
\end{aligned}$$

Vegyük észre, hogy összeadásnál és kivonásnál a hiba adódik össze, osztásnál és szorzásnál pedig a relatív hiba!

8. Záró megjegyzések

Egy mérési adat sosem egy számot jelent, hanem minimum két számot: az értéket és a hibáját! Hiba nélkül a mérési eredmény értelmetlen. Pl. az $x = 5.2$ mm ± *0.1* mm mérési eredmény értelmes, az ~~$x = 5.2$ mm~~ nem.

Figyeljünk oda továbbá arra, hogy a mérés hibájának mindig egy értékes számjegye legyen, a mért adat utolsó számjegye pedig azon a helyiértéken álljon, ahol a hiba egyetlen értékes számjegye. Például az $x = 5.2$ mm ± *0.1* mm mérési eredmény értelmes, az ~~$x = 5.2032$ mm ± *0.1* mm~~ illetve az ~~$x = 5$ mm ± *0.01* mm~~ eredmények azonban helytelenül megadottak.

9. Ellenőrző kérdések

1. Milyen eloszlása van egy adott mérés eredményeinek általában?
2. Hogyan függ egy mérés eredményeinek eloszlása a mérés megismétléseinek számától?
3. Mi egy mérés-sorozat átlaga?
4. Mi egy mérés-sorozat szórása?
5. Miért van négyzet a szórás definíciójában?
6. Hogyan függ a szórás a mérések számától?
7. Mi a khi-négyzet próba?
8. Khi-négyzet próba esetén a khi-négyzet milyen értékénél van az elméleti jóslat az átlagos hibahatáron (szóráson) belül?
9. Hogyan lehet egy elméletet cáfolni khi-négyzet próba segítségével?
10. Mi a minimalizáció szerepe egy elmélet optimális paramétereinek megtalálásában?
11. Mit jelent az, hogy „illesztés”?
12. Egy illesztés után hogyan döntjük el, hogy a mérés cáfolja vagy megerősíti az elméletet?
13. Milyen függvény illesztése végezhető el egyszerűen, és miért?
14. Hogyan függ egy mennyiség hibája azon mennyiségek hibájától, amelyektől függ?
15. Mi a relatív hiba?
16. Milyen származtatott mennyiség esetén adódik össze a hiba illetve a relatív hiba?
17. Mekkora egy különbség illetve egy hányados hibája?
18. Hány számjegyig adjuk meg a mérési eredményt és annak hibáját?
19. Mi a „szabadsági fokok száma” egy illesztés esetén, és miért fontos?