MÉRÉS ÉS MÉRÉSKIÉRTÉKELÉS ALAPFOGALMAI

Statisztika, mérések, megfigyelések feldolgozása

Összeállította: Dr. Katona Tamás János

A téma és a megválaszolandó kérdések:

1. *Mi a mérés? Mi a mérési hiba?*
2. *Mi a méréskiértékelés (a statisztika) tárgya?*
3. *A mért, megfigyelt változók jellemzői. A várható érték, a szórás becslése?*
4. *Hogyan adjuk meg a mérés eredményét?*
5. *Mi az empirikus valószínűségi eloszlás, illetve sűrűségfüggvény?*
6. *Trendek, korrelációk, illesztés – alapfogalmak, Excel-példák.*
7. *Mi a hipotézis-vizsgálat lényege?*

1. A mérés

Mérés: Az a folyamat, amely magába foglalja a mérendő mennyiségről szerzett információ felvételét, az információnak a mérőberendezés segítségével történő átalakítását abból a célból, hogy eredményül a mérendő mennyiség kvalitatív értékét kapjuk. A mérés energia- és információátvitel, illetve feldolgozás.

A mérés lehet:

- Közvetlen: etalonnal összehasonlítjuk.
- Közvetett: olyan paramétert mérünk közvetlenül, amely függvénykapcsolatban van a vizsgálni kívánt.
- Analóg- és digitális.

A mért mennyiség tartománya:

$$T=(x_{max}-x_{min})$$

Legyen a mérőműszer érzékenysége Δx, amin belül nem tud az értékek között különbséget tenni, akkor

$$k=(x_{max}-x_{min})/\Delta x$$

értéket tudunk megkülönböztetni egymástól.

Az 1/k %-ban kifejezve a műszer osztálypontossága.

2. A mérési hiba

Minden mérés tartalmaz hibát.

Hiba keletkezhet, előállhat:

- a mérési módszer,

- a mérőberendezés,
- a mérendő mennyiség,
- a mérést végző,
- a környezet

miatt.

Legyen x mért mennyiség valódi értéke x_0,

Az abszolút eltérés $|x-x_0| <a$

A mérési hiba az *i*-ik mérésben

$$H_i=x_i-x_0$$

x_i a mért érték *i*-ik mérés;

x_0 a „helyes” érték;

A hiba lehet rendszeres, amelyekkel a mérési eredményt korrigálni lehet (pl. a műszer skálahibái, kitérést befolyásoló elemek pontossága). Skálaeltolódás, nullpont-eltolódás. Kitéréssel arányos hiba. Kalibrálás.

Véletlen hiba: nagyságuk, előjelük, időbeliségük ismeretlen:

$$x_h=x_i\pm\gamma$$

ahol γ véletlen változó.

Rendszeres hiba a mérési eredményt torzítja, a véletlen bizonytalanná teszi.

3. A statisztika tárgya

A folyamatok, azok állapot-jellemzői vagy természetüknél fogva valószínűségi változók, mint pl. a turbulens sebesség-, hőmérsékletmező, vagy a megismerés módja-eszközei, azaz a mérés következtében a megfigyelt, vagy mért jellemzők véletlenszerű hibával terheltek.

A valószínűségi változókat eloszlás- és sűrűségfüggvényük, és azok paraméterei jellemzik.

A leggyakoribb eloszlás-típus a normáleloszlás: Annak valószínűsége, hogy a ξ normál eloszlású, m várható értékű és σ szórású valószínűségi változó megfigyelt értéke az [a,b) intervallumba esik:

$$P(a \le \xi < b) = \frac{1}{\sigma\sqrt{2\pi}} \int_a^b e^{-\frac{(x-m)^2}{2\sigma^2}}$$

Ez a feladat egy egyértelmű számítás a valószínűségelméletben, de a valóságban soha sem áll rendelkezésünkre végtelen adathalmaz, eloszlás- illetve sűrűségfüggvény, illetve ezek paramétereinek pontos értéke a vizsgálni kívánt jelenségről, vagy objektumról.

A valóságban a megfigyelések véges számúak, s vagy az eloszlás ismert, s a paramétereket keressük, vagy az eloszlás is ismeretlen.

A statisztika

1. becslésekkel
2. hipotézisek vizsgálatával

foglalkozik, s a megfigyelések, mérések alapján, a valószínűségi változók ismeretlen valószínűségeire, valószínűségfüggvényeire, és azok paramétereire következtet.

A becslés lehet pontbecslés vagy intervallumbecslést.

Statisztikai minta: véges számú kísérlet vagy megfigyelés. A mintaelemek függetlenek egymástól. A minta reprezentatív, ha a mintaelemek egyenként is reprezentálják a valószínűségi változót.

Statisztikának hívjuk a ξ valószínűségi változó bármely függvényét.

4. Becslés

Igen általános becslési feladat a következő: Legyen ξ, normál eloszlású valószínűségi változó, *m* várható értékkel és *σ* szórással. Méréssel határozzuk meg a várható értéket és a szórást.

4.1. Mérési sorozat átlaga

A sorozat álljon *n* mérésből, ennek átlaga, vagy számtani közepe az alábbi:

$$x_1, x_2, \dots x_n$$

$$\bar{x} = \frac{1}{n}\sum_1^n x_i$$

A középérték maga is valószínűségi változó. Ha *n* mérési sorozatot végzünk, akkor

$$M(\bar{x}) = [M(x_1)+M(x_2)+M(x_3)+\dots+M(x_n)]/n$$

Az *n* független mérési eredmény középértékének szórása egyenlő az egyes értékek (mérési sorozatok) szórásának *1/√n*-szeresével, azaz a mérések számának növelésével csökken, jó közelítéssel adja az *m* várható értéket, ha *n* elegendően nagy.

4.2. A szórás közelítése

Az empirikus szórásnégyzet, s^2:

$$s^2 = \frac{1}{n}\sum_1^n \left(x_i - \bar{x}\right)^2$$

ami maga is valószínűségi változó, amelynek várható értéke

$$M\left(\sigma^2\right) = \frac{n-1}{n}\sigma^2$$

Ezért célszerű bevezetni a korrigált empirikus szórásnégyzetet, ami:

$$\overset{*}{s}^{2} = \frac{n}{n-1} s^2$$

Ennek várható értéke pontosan a szórásnégyzet, azaz

$$M(\overset{*}{s}^{2}) = \sigma^2$$

5. Mérési eredmény megadása

A véletlen hibák értéke nem határozható meg pontosan, ezért a véletlen hibákkal terhelt mérések, vagy véletlen változó mintái esetében csak azt tudjuk megmondani, hogy a változó milyen valószínűséggel marad belül egy megadott intervallumon.

Legyen ξ, normál eloszlású valószínűségi változó, *m* várható értékkel és *σ* szórással.

Ebben az esetben az

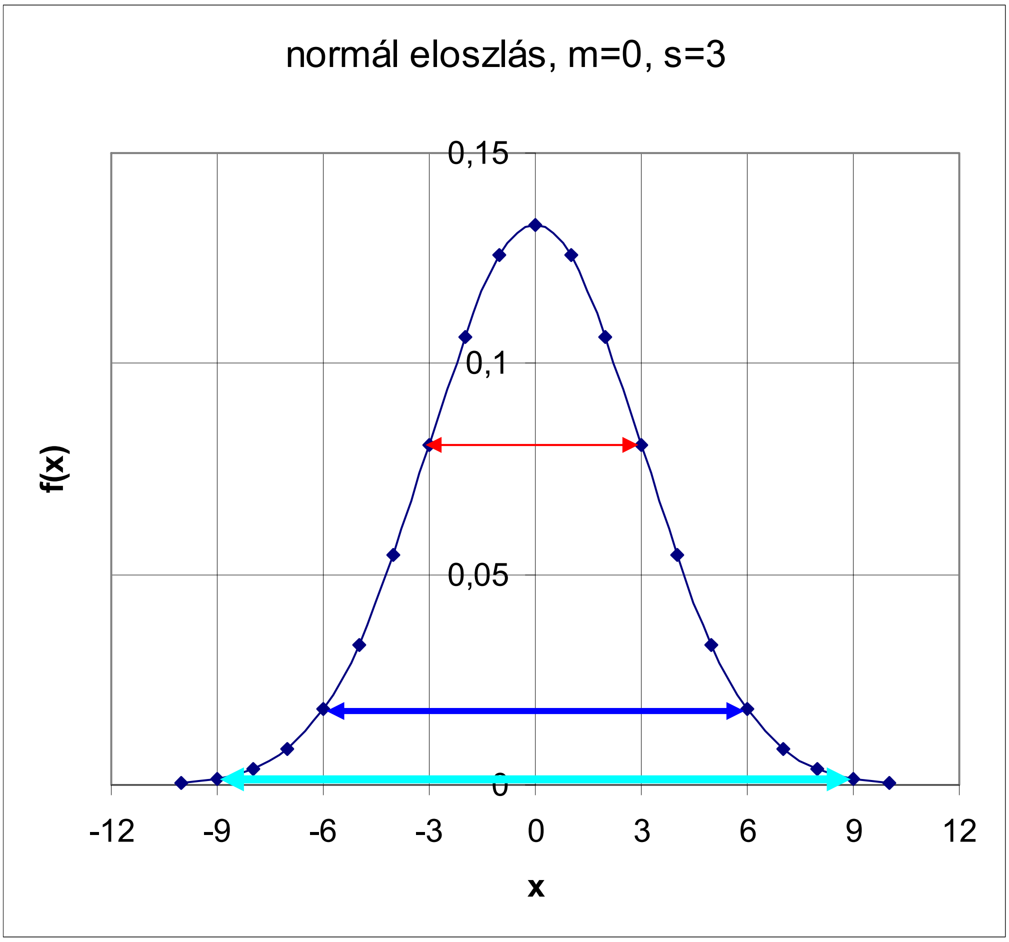
$$m \pm k^*\sigma$$

intervallumban

az értékek 68,27%-a esik, ha k=1,

95,45%-a, ha k=2,

99,73%-a, ha k=3.



1. ábra. Normál eloszlás és a konfidencia-intervallumok

A mérési eredmény, ha egy mérés történt:

$$X = x - H_a \pm t*s$$

formában adják meg, ahol

x, az egyetlen mért érték,

H_a , a mérési eljárás ismert rendszeres hibáinak összege,

s, a mérési eljárás ismert szórása.

A *n* mérési sorozat eredményét pedig az alábbiak szerint adják meg:

$$X = \bar{x} - H_a \pm \frac{t * s}{\sqrt{n}}$$

Ezek az úgynevezett megbízhatósági szintek.

6. Gyakorisági és a sűrűséghisztogram

Osszuk fel az [a,b] intervallumot:

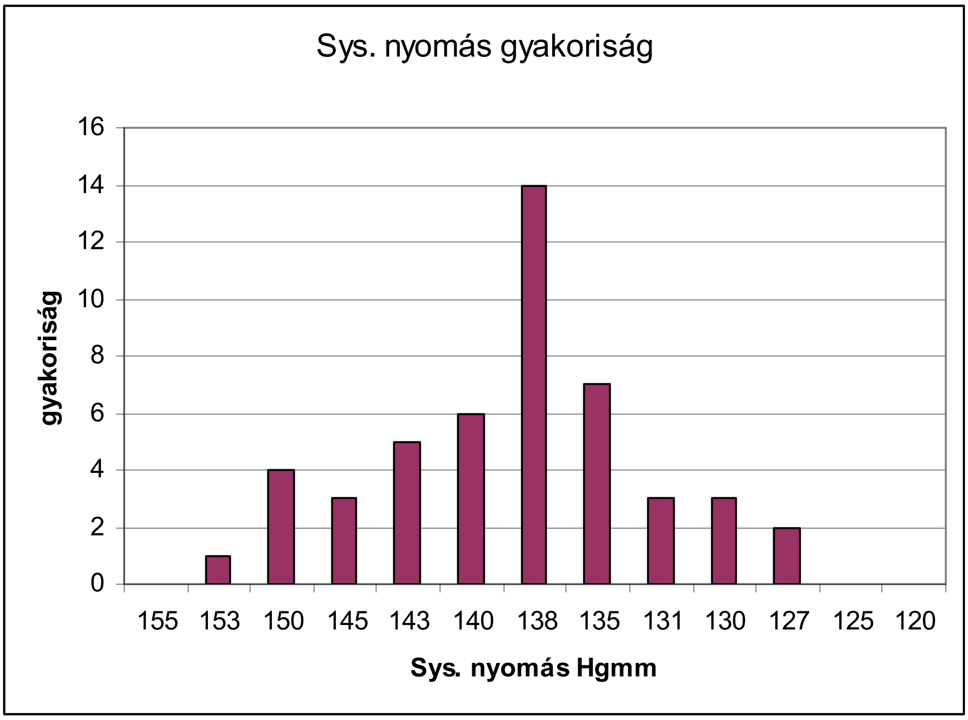
$$a=d_0<d_1< d_2< d_3< \ldots < d_r=b$$

Ábrázoljuk az egyes az egyes részintervallumokba eső minták, k_i, számát a rárajzolt téglalap magasságával arányosan:

$$f_n(x)=\frac{\frac{k_i}{n}}{d_i-d_{i-1}}=\frac{\Delta F_n(x)}{\Delta_i}$$

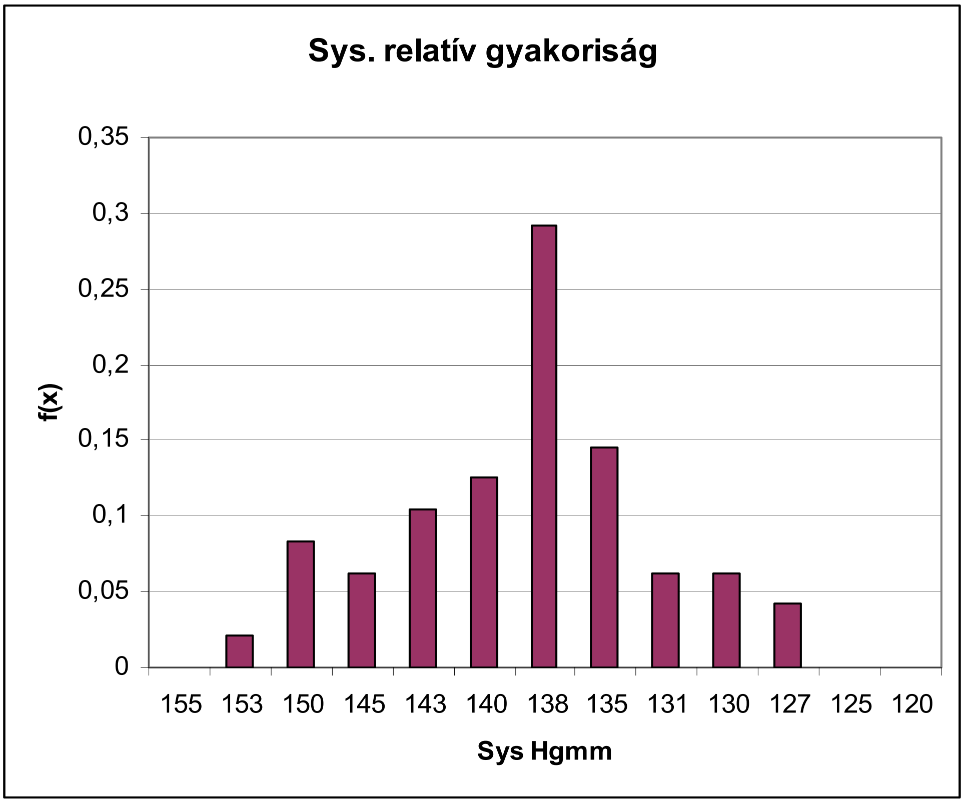
$$d_{i-1}\le x<d_i$$
$$i=1,2,\ldots,r$$

Az így kapott ábrát az empirikus sűrűségfüggvény grafikonjának is szokták nevezni:

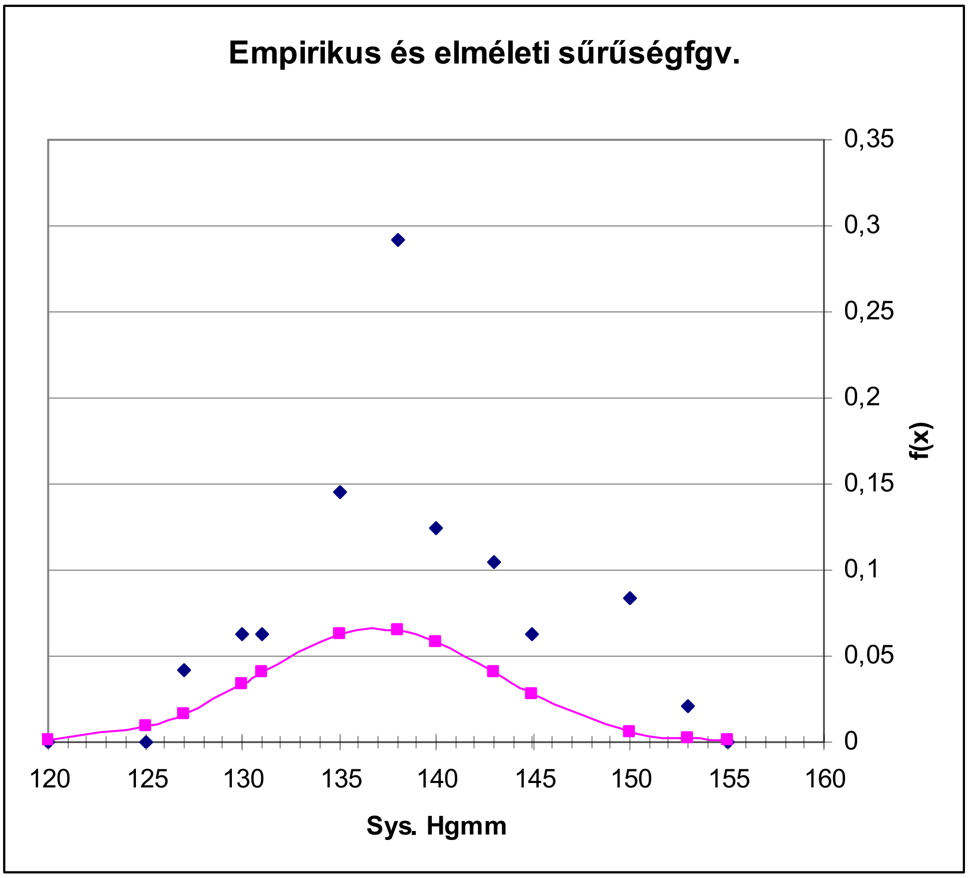


2. ábra. Eloszlási hisztogram

Lehet relatív értékeket ábrázolni:



3. ábra. Relatív eloszlási hisztogram



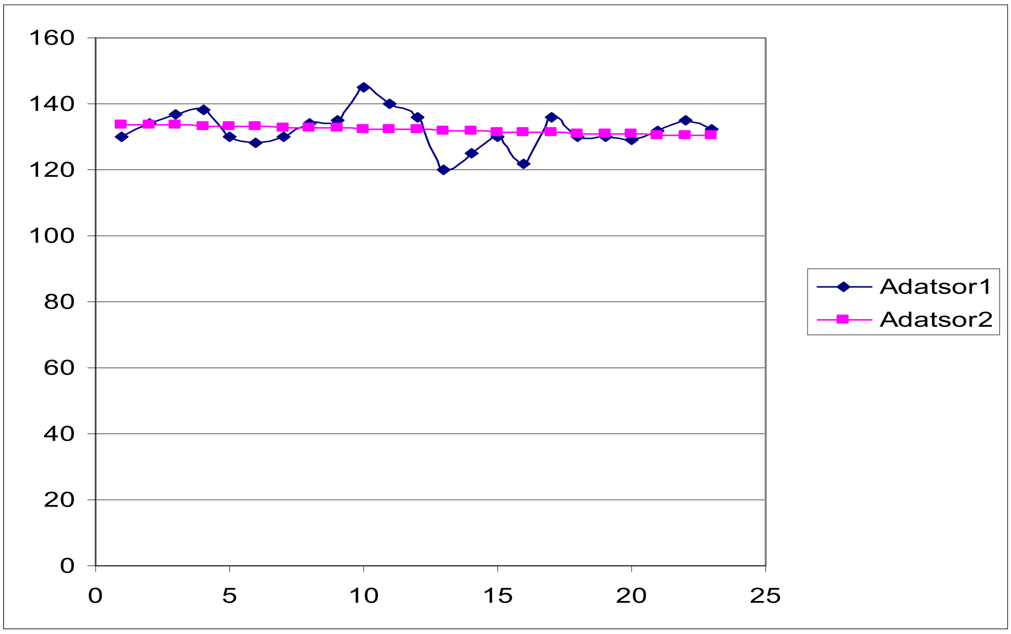
4. ábra. Elméleti és empirikus sűrűségfüggvény

7. Illesztés, trend

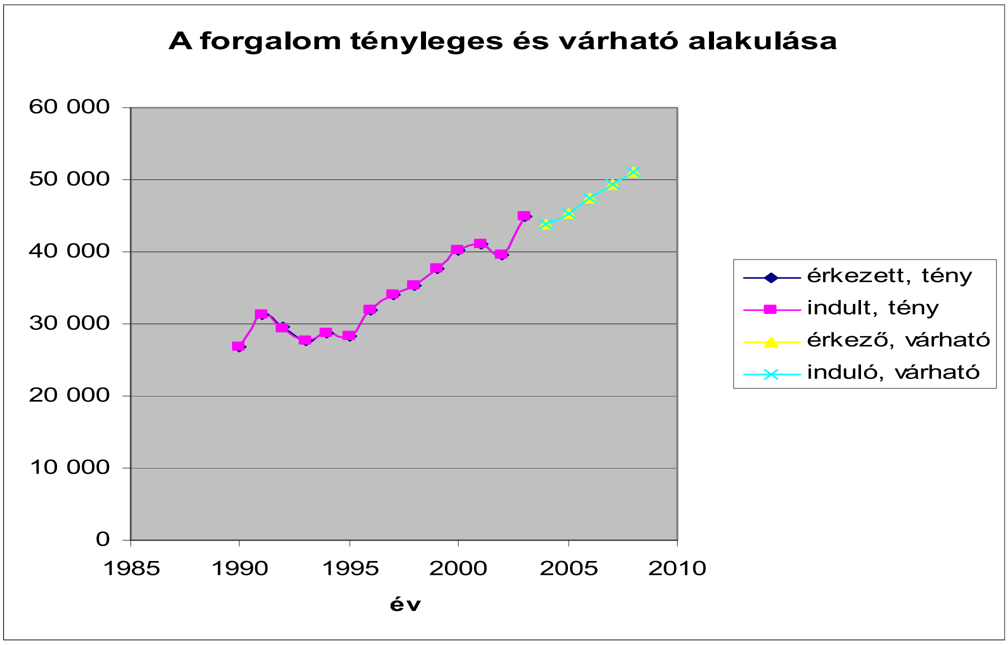
Illesztés legkisebb négyzetek módszerével:

Lineáris közelítés esetén a legkisebb négyzetek módszere szerint ***a*** és ***b*** értékét úgy kell meghatározni, hogy az eltérés négyzetének várható értéke a lehető legkisebb legyen, azaz:

$$M\left[\eta-(a\xi+b)\right]^2 = minimum$$



5. ábra. Illesztés mérési adatokra



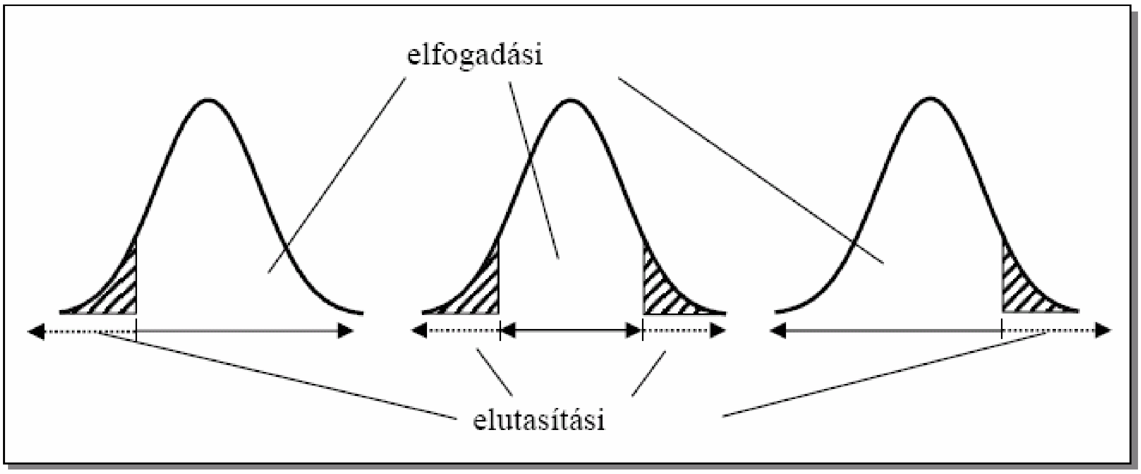
6. ábra. Trend-elemzés

8. Hipotézisvizsgálat

Statisztikai hipotézis: a valószínűség-eloszlásra vonatkozó feltevés.

	Ha a H_0 hipotézist	
	elfogadjuk	elvetjük
H_0 fennáll	helyes döntés	első fajú hiba
H_0 nem áll fenn	másodfajú hiba	helyes döntés

Ezt szemlélteti a sűrűségfüggvények példáján az 7. ábra:



7. ábra: Hipotézis-vizsgálat

Az olyan statisztikai próbát, amelynek alapján arról döntünk, hogy valamely ξ valószínűségi változó eloszlása lehet-e adott F(x) eloszlásfüggvénnyel jellemzett eloszlás, illeszkedés-vizsgálatnak nevezzük. Ilyenkor a H_0 nullhipotézis:

$$H_0\text{: } F(x)$$

Ha a nullhipotézis az eloszlás paramétereinek ismeretét is feltételezi, akkor tiszta illeszkedésvizsgálatról beszélünk. Ha a hipotézisünk csak az eloszlás jellegét (normál, exponenciális, stb.) tételezi fel, és a paramétereket a mintából kell becsülnünk, akkor becsléses illeszkedésvizsgálatot végzünk.

Az illeszkedésvizsgálatra szolgáló próbák alkotják a nem-paraméteres próbák egyik nagy csoportját. E próbák közül legelterjedtebb a χ^2-próba és a Kolmogorov-próba. A χ^2-próba mind diszkrét, mind folytonos eloszlások esetében alkalmazható, de nagy minta-elemszámot igényel. A próba segítségével azt tudjuk eldönteni, hogy adott szignifikancia szinten a tapasztalati gyakoriságok szignifikánsan eltérnek-e a feltételezett elméleti gyakoriságoktól, avagy az eltérés csupán a véletlen következménye.

χ^2-próbával történő illeszkedésvizsgálatnál az un. próbastatisztikát (a számított értéket) az alábbi eljárás szolgáltatja:

Tekintsük a

$$u = \sum_{i}^{k} \frac{(r_i - m_i)^2}{m_i} \quad \text{valószínűségi változót}$$

amely χ^2 eloszlást követ.

$$DF = r - l - 1$$

ahol:

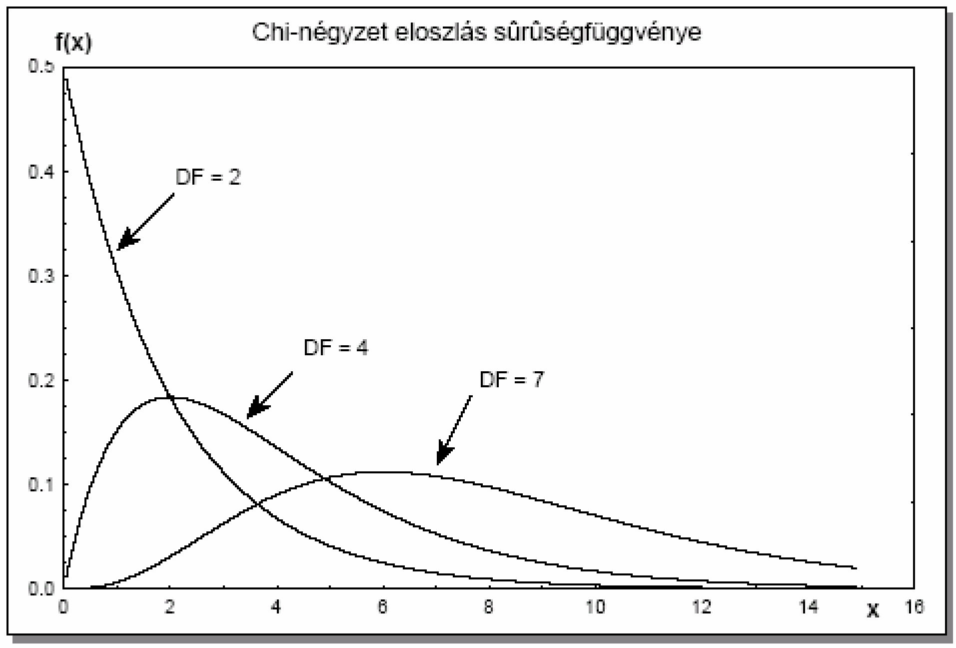
DF: a szabadságfok, az eloszlás paramétere

r_i : a tapasztalati gyakoriság;

m_i: az elméleti gyakoriság, a hipotézis alapján kiszámolva;

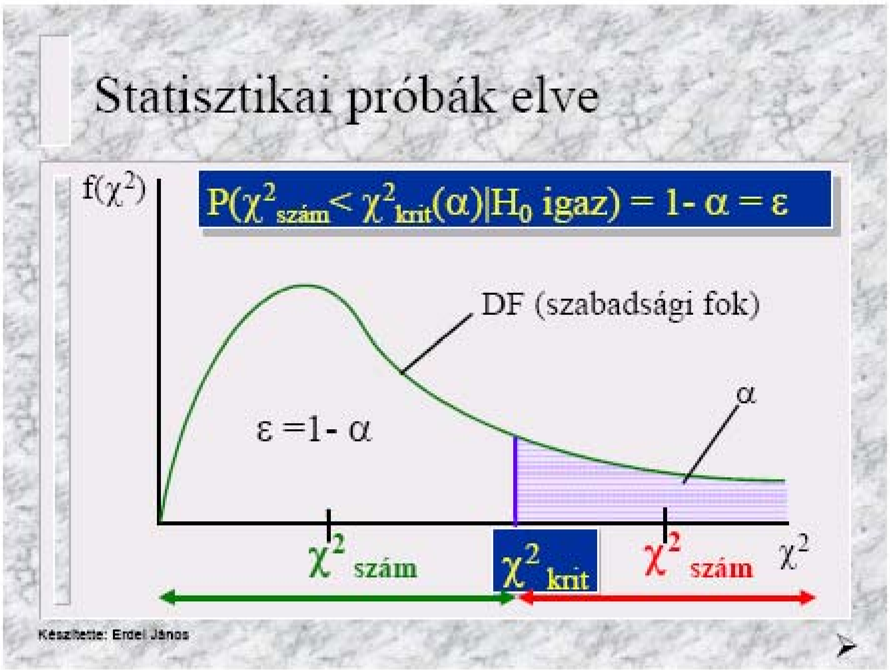
l : a becsült paraméterek száma

Az eloszlásfüggvényt a 8. ábrán láthatjuk.



8. ábra. Véletlen események Poisson-eloszlással való leírásának χ^2 próbája

Az eljárást a 9. ábra szemlélteti:



9. ábra: A hipotézisvizsgálat elve

<u>Példa: a Poisson-eloszlás hipotézis-vizsgálatára χ^2-próbával</u>

Hipotézis: Kövessék a meghibásodási időpontok Poisson-eloszlást:

$$F(n) = (\lambda\ T)^{\ n}\ \exp\ (\ -\lambda\ T\ )\ /\ n!$$

λ: empirikus gyakoriság;

T: a vizsgált időtartam;

n: az esetek száma.

$$p_i=\lambda^i\ e^{-\lambda}/i!$$

$$m_i=np_i$$

χ^2 –próba, 1. eseménysor

1-es kategória, gyakoriság 1/év	1,27			
x	0	1	2	3 vagy több
r_x	3	6	5	1
m_x	4,226539322	5,353616489	3,390623785	1,431596713
$u_x=(r_x-m_x)^2/m_x$	0,355941018	0,078042879	0,763898316	0,130117456
			$\Sigma\ u_x$	1,327999669
			χ^2érték	5,99

A hipotézis helyes!

9. Ajánlott irodalom

Igen sok jó szakkönyv, tankönyv létezik, s vannak mérés, és méréskiértékelés tárgyában kötelező szabványok is.

Reimann József, Tóth Julianna: Valószínűségszámítás és matematikai statisztika, Nemzeti Tankönyvkiadó, Bp., 1985

Dr. Kaposvári Zoltán: Méréselmélet, Tankönyvkiadó, Bp. 1980.

Fred Caswell, Success in Statistics, John Murray, 1989.

Derek J. Hudson, Statistics, Geneva, 1964