Valószínűségszámítás
5. RÉSZ
Kétdimenziós folytonos valószínűségi változók (B rész)
—
FOGALMAK ÉS KIDOLGOZOTT PÉLDÁK

Vetier András

2018. december 5.

Tartalomjegyzék

1. Kovariancia

1.1. Kovariancia adatrendszerekre

1.1.1. Kétdimenziós adatrendszer fogalma

Számpároknak egy sorozatát *kétdimenziós adatrendszer*nek hívjuk. Egy számpár egy vektort is jelent:

- ha a számpár elemeit egymás mellé, azaz "egy sorba" írjuk, akkor a *sorvektor*t kapunk
- ha a számpár elemeit egymás alá, azaz "egy oszlopba" írjuk, akkor a *oszlopvektor*t kapunk

Izlés dolga, és igazából egyremegy, hogy sor- vagy oszlopvektorokkal dolgozunk, de választani kell a két lehetőség közül, és a a két lehetőséget nem szabad összekeverni. Mi itt ebben a könyvben oszlopvektorokat fogunk használni.

Vektorok összege, átlaga jól ismert fogalmak, ezért egy kétdimenziós adatrendszer *átlag*ának fogalma sem szorul sok magyarázatra. Nyilvánvaló, hogy egy kétdimenziós adatrendszer átlagának koordinátái a koordináták külön-külön vett átlagaival egyenlő.

1. Példa: Kétdimenziós adatrendszer 5 elemből. Az alábbi adatrendszer 5 számpárból áll:

x	54	55	44	44	60
y	85	83	92	65	75

Vegyük most minden (x, y) adat esetén az $x \cdot y$ szorzatot, és tekintsük ezeknek a szorzatoknak az átlagát. Ezt az átlagot a kétdimenziós adatrendszer *vegyes (második) momentum*ának nevezünk:

						átlag
$x \cdot y$	4 590	4 565	4 048	2 860	4 500	4 112.6
						vegyes momentum

A vegyes momentum fogalma akkor jut szerephez, amikor egy kétdimenziós adatrendszerből – például – az

x	54	55	44	44	60
y	85	83	92	65	75

adatrendszerből a megfelelő x és y értékek összeadásával egy új egydimenziós adatrendszert állítunk elő:

$x + y$	139	138	136	109	135

Ennek az egydimenziós adatrendszernek a második momentumát a szokásos módon kiszámíthatjuk. A második momentum értéke $17\,393.4$.

A középiskolai tanulmányainkból jól ismert $(a+b)^2 = a^2 + b^2 + 2ab$ azonosság mintájára egyszerűen adódik, hogy *az összegzéssel kapott adatrendszer második momentuma egyenlő a külön-külön vett egydimenziós adatrendszerek* *második momentumainak az összegével plusz 2-szer a vegyes momentum.* A mi numerikus példánkra ez a szabály teljesül:

$$17\,393.4 = 2\,682.6 + 6\,485.6 + 2 \cdot 4\,112.6$$

Az általános formula megfogalmazása és annak algebrai úton való bizonyítása legyen az Olvasó feladata.

Ha a kétdimenziós adatrendszer elemeit oszlopvektorként kezeljük, és az adatokat egymás mellé írjuk, akkor az egymás mellé írt oszlopvektorok egy mátrixot alkotnak. A mátrixnak két sora van, és annyi oszlopa, ahány számpárból áll az adatrendszer.

Ha a kétdimenziós adatrendszer által definiált mátrixot megszorozzuk jobbról az ő transzponáltjával, és osztunk még az adatok számával (ami itt 5), akkor egy olyan 2×2-es mátrixot kapunk, melynek

- bal felső eleme az x-koordináták adatrendszerének második momentuma
- jobb alsó eleme az y-koordináták adatrendszerének második momentuma
- a bal alsó és a jobb felső eleme pedig a vegyes momentum

Az állítás igazságát a kétdimenziós példánkon ellenőrizhetjük:

$$\begin{array}{|ccccc|} \hline 54 & 55 & 44 & 44 & 60 \\ 85 & 83 & 92 & 65 & 75 \\ \hline \end{array} \cdot \begin{array}{|cc|} \hline 54 & 85 \\ 55 & 83 \\ 44 & 92 \\ 44 & 65 \\ 50 & 75 \\ \hline \end{array} \cdot \frac{1}{5} = \begin{array}{|cc|} \hline 2\,682.6 & 4\,112{,}6 \\ 4\,112.6 & 6\,485{,}6 \\ \hline \end{array}$$

Ennek a mátrixnak a neve: a kétdimenziós adatrendszer *második momentum mátrix*a.

1.1.2. Kovariancia és kovariancia mátrix

Ha egy kétdimenziós adatrendszernek vesszük egy elemét, és ebből az elemből kivonjuk az adatrendszer átlagát, (vagyis az első koordinátájából kivonjuk az első koordináták átlagát, és a második koordinátájából kivonjuk a második koordináták átlagát), akkor a szóbanforgó elem *cenralizáltjá*hoz jutunk. A centralizálással kapott számpár megadja a szóbanforó elemnek az átlaghoz viszonyított helyzetét. Ha az adatrendszer minden elemének vesszük a centralizáltját, akkor a kapott kétdimenziós adatrendszert az eredeti *adatrendszer centralizáltjá*nak nevezzük. Az adatrendszer centralizáltja az adatrendszernek az átlaghoz viszonyított helyzetét adja meg.

Példa: A

						átlag
x	54	55	44	44	60	51.4
y	85	83	92	65	75	80

kétdimenziós adatrendszer centralizáltja:

$x - x$ átlag	2.6	3.6	-7.4	-7.4	8.6
$y - y$ átlag	5	3	12	-15	-5

A centralizált adatrendszer vegyes momentumát az eredeti adatrendszer *kovarianciá*jának nevezzük. Egyszerűen igazolható tény, hogy

$$\text{kovariancia} \;=\; \text{vegyes momentum} \;-\; \text{első momentumok szorzata}$$

A centralizált adatrendszer momentum mátrixát az eredeti adatrendszer *kovariancia mátrixá*nak nevezzük. Nem nehéz belátni, hogy

$$\text{kovariancia mátirx} \;=\; \text{második momentum mátrix} \;-\; \text{várható értékek oszlopvektora} \cdot \text{várható értékek sorvektora}$$

A kovarinacia mátrix

- bal felső eleme az x-koordináták adatrendszerének varianciája
- jobb alsó eleme az y-koordináták adatrendszerének varianciája
- a bal alsó és a jobb felső eleme pedig a kovariancia

Az állítás igazságát a kétdimenziós példánkon ellenőrizhetjük:

$$\boxed{\begin{array}{ccccc} 2.6 & 3.6 & -7.4 & -7.4 & 8.6 \\ 5 & 3 & 12 & -15 & -5 \end{array}} \cdot \boxed{\begin{array}{cc} 2.6 & 5 \\ 3.6 & 3 \\ -7.4 & 12 \\ -7.4 & -15 \\ 8.6 & -5 \end{array}} \cdot \frac{1}{5} = \boxed{\begin{array}{cc} 40.6 & 0.6 \\ 0.6 & 85.6 \end{array}}$$

Mint fentebb tisztáztuk: *az összegzéssel kapott adatrendszer második momentuma egyenlő a külön-külön vett egydimenziós adatrendszerek második momentumainak az összege plusz 2-szer a vegyes momentum.* Ezt a alkalmazva centralizált kétdimenziós adatrendszerre azt kapjuk, hogy *az összegzéssel kapott adatrendszer varianciája egyenlő a külön-külön vett egydimenziós adatrendszerek varianciáinak az összege 2-szer a kovariancia*.

1.1.3. Kovariancia mátrix transzformálódása

Ha egy kétdimenziós oszlopvektort megszorzunk balról egy 2×2 -es ún. "transzformációs" mátrixszal, akkor egy másik oszlopvektort kapunk. Ha egy ilyen szorzást megcsiálunk egy kétdimenziós adatrenszer minden oszlpvektorával, akkor egy új kétdimenziós adatrendszert kapunk. Ha az új adatrendszerre előállítjuk a előző részben leírt módon a kovariancia mátrixot, akkor az egész procedúrából balról kiemelhető a 2×2 -es transzformációs mátrix, jobbról pedig a transzformációs mátrix transzponáltja, és kiadódik a következő szabály: *az új kétdimenziós adatrendszer kovariancia mátrixa egyenlő az eredeti adatrendszer kovariancia mátrix balról megszorozva a transzformációs mátrixszal, jobbról pedig a transzformációs mátrix transzponáltjával*.

Ha a transzformációs mátrix egy olyan sorvektor, mely egy egységvektor koordinátáiból áll, akkor transzformációs mátrixszal balról szorozva egy oszlopvektort a szorzat értéke az oszlopvektornak az egységvektor irányára vonatkozó vetületét adja. Ezért *ha egy egységvektornak megfelelő sorvektorral balról, és az ő transzponáltjával jobbról szorzunk egy kovariancia mátrixot, akkor az eredmény annak az adatrendszernek a varianciáját adja, melyet akkor kapunk, ha a kétdimenziós adatrendszer elemeit a sorvektor irányára vetítjük.*

Ezért a kovariancia mátrix segítségével a kétdimenziós adatrendszer mindenféle vetületeinek a varianciáját ki lehet számolni. Ilyen értelemben a kovariancia mátrix a variancia fogalmának többdimenzióra vonatkozó kiterjesztésének tekinthető.

1.2. Kovariancia valószínűségi változókra

1.2.1. A kovariancia fogalma

Az X és az Y valószínűségi változók szorzatának $\mathrm{E}\,(X\;Y)$ várható értékét **vegyes (második) momentumnak**nak nevezzük. Az X és az Y valószínűségi változók centralizáltjai szorzatának a várható értékét **kovarianciá**nak nevezzük:

$$\mathrm{COV}(X,Y) = \mathrm{E}\,(\;[X - \mathrm{E}(X)\,]\,[Y - \mathrm{E}(Y)\,]\;)$$

Könnyű belátni, hogy a kovarianciát úgy is megkapjuk, hogy a vegyes momentumból kivonjuk a várható értékek szorzatát:

$$\mathrm{COV}(X,Y) = \mathrm{E}(X\;Y) - \mathrm{E}(X)\;\mathrm{E}(Y)$$

1.2.2. A kovariancia tulajdonságai

1. *Egy valószínűségi változó önmagával vett kovarianciája a valószínűségi változó varianciájával egyenlő:*

$$\mathrm{COV}(X,X) = \mathrm{VAR}(X)$$

2. *A kovariancia szimmetria tulajdonsága:*

$$\mathrm{COV}(Y,X) = \mathrm{COV}(X,Y)$$

3. *A kovariancia linearitási tulajdonsága:*

   - Két tagra:

$$\mathrm{COV}(\;a\;X + b\;Y\;,\;Z) = a\;\mathrm{COV}(X,\;Z) + b\;\mathrm{COV}(Y,\;Z)$$

   - Több tagra:

$$\mathrm{COV}(a_1X_1 + a_2X_2 + \ldots + a_nX_n, Y) = a_1\;\mathrm{COV}(X_1,Y) + a_2\;\mathrm{COV}(X_2,Y) + \ldots + a_n\;\mathrm{COV}(X_n,Y)$$

4. *A kovariancia abszolút értéke mindig kisebb vagy egyenlő, mint a szórások szorzata:*

$$|\mathrm{COV}(X,Y)| \le \mathrm{SD}(X)\;\mathrm{SD}(Y)$$

5. *Független valószínűségi változók között a kovariancia 0:*

$$\mathrm{COV}(X,Y) = 0$$

6. *Összeg varianciája:*

$$\mathrm{VAR}(X+Y) = \mathrm{VAR}(X) + \mathrm{VAR}(Y) + 2\;\mathrm{COV}(X,Y)$$

7. *Egymástól független vektorváltozók összegének koordinátái közötti kovariancia egyenlő a tagok koordinátái közötti kovarianciák összegével:*

   - Két tagra:
     Ha az $(X_1,Y_1),\;(X_2,Y_2)$ véletlen vektorok függetlenek egymástól, és

$$(X,Y) = (X_1,Y_1) + (X_2,Y_2)$$

     akkor

$$\mathrm{COV}(X,Y) = \mathrm{COV}(X_1,Y_1) + \mathrm{COV}(X_2,Y_2)$$

- Több tagra:
  Ha az $(X_1, Y_1), \ldots \ (X_n, Y_n)$ véletlen vektorok függetlenek egymástól, és
$$(X, Y) = (X_1, Y_1) + \ldots + (X_n, Y_n)$$
  akkor
$$\text{COV}(X, Y) = \text{COV}(X_1, Y_1) + \ldots + \text{COV}(X_n, Y_n)$$

Megjegyzés: A feltételben nem az X_1 független Y_1-tól, illetve X_2 független Y_2-től, hanem az (X_1, Y_1) vektor független az (X_2, Y_2) vektortól.

8. *Független vektor valószínűségi változók összegének koordinátái közötti kovariancia, amikor a tagok koordinátái közötti kovarianciák azonosak:* Ha az $(X_1, Y_1), \ldots \ (X_n, Y_n)$ véletlen vektorok függetlenek egymástól, és mindegyik tag esetében a koordiánáták közötti kovariancia azonos, azaz
$$\text{COV}(X_i, Y_i) = c \quad \text{minden } i\text{-re}$$
   és
$$(X, Y) = (X_1, Y_1) + \ldots + (X_n, Y_n)$$
   akkor az összeg koordiánáti közötti kovariancia $c \cdot n$ -nel egyenlő:
$$\text{COV}(X, Y) = c \cdot n$$

1.2.3. A korrelációs együttható fogalma és tulajdonságai

1. Az X és az Y valószínűségi változók standardizáltjai szorzatának a várható értékét **korrelációs együtthatójának**nak nevezzük:
$$\text{CORR}(X, Y) \ = \ \text{E}\left(\frac{X - \text{E}(X)}{\text{SD}(X)} \cdot \frac{Y - \text{E}(Y)}{\text{SD}(Y)} \right) \ =$$
   A korrelációs együttható nyilván így is írható:
$$= \ \frac{\text{E}\left(\left(X - \text{E}(X) \right) \cdot \left(Y - \text{E}(Y) \right) \right)}{\text{SD}(X) \cdot \text{SD}(Y)} \ = \ \frac{\text{COV}(X, Y)}{\text{SD}(X) \cdot \text{SD}(Y)}$$
   Ezek szerint a korrelációs együttható nem más, mint a **kovariancia osztva a szórások szorzatával**.

2. *A korrelációs együttható abszlút értéke mindig kisebb vagy egyenlő, mint 1:*
$$|\ \text{CORR}(X, Y)\ | \ \le \ 1$$
   vagyis a korreláció értéke mindig -1 és 1 közé esik:
$$-1 \le \ \text{CORR}(X, Y) \ \le \ 1$$

3. *Egy valószínűségi változónak bármely lineáris transzormáltjával vett korrelációs együtthatója -1 -gyel vagy 1 -gyel egyenlő:*
$$\text{CORR}(\ X\ ,\ aX + b\ ) \ = \ 1 \quad \text{ha} \quad a > 0$$
$$\text{CORR}(\ X\ ,\ aX + b\ ) \ = \ -1 \quad \text{ha} \quad a < 0$$

4. *Ha X és Y között a korreláció -1 vagy $+1$, akkor vannak olyan a és b konstansok, hogy*
$$Y \ = \ aX + b \quad \text{fennáll 1 valószínűséggel, azaz} \quad \text{P}(\ Y = aX + b\ ) \ = \ 1$$
   vagyis az X és Y véletlen értékek között determinisztikus lineáris kapcsolat áll fenn.

5. *Független valószínűségi változók között a korreláció 0 :*
$$\text{CORR}(X, Y) \ = \ 0$$

1.2.4. A kovariancia mátrix fogalma és tulajdonságai

1. Az (X, Y) kétdimenziós valószínűségi változó **kovariancia mátrix**ának definíciója:

$$C = \begin{pmatrix} \mathrm{COV}(X, X) & \mathrm{COV}(X, Y) \\ \mathrm{COV}(Y, X) & \mathrm{COV}(Y, Y) \end{pmatrix}$$

   azaz

$$C = \begin{pmatrix} \mathrm{VAR}(X) & \mathrm{COV}(X, Y) \\ \mathrm{COV}(X, Y) & \mathrm{VAR}(Y) \end{pmatrix}$$

   A kovariancia mátrix a variancia fogalmának általánosítása kétdimenzióra.

2. *Várható érték vektor és kovariancia mátrix transzformálódása lineáris transzofrmáció esetén.*

   Ha az **x**-síkról egy kétdimenziós eloszlást az

$$\mathbf{u} = \mathbf{A}\ \mathbf{x}\ +\ \mathbf{b}$$

   lineáris transzformációval az **u**-síkra képezünk, akkor a kapott új eloszlás várható érték vektora a régi $\mathbf{m}_{\text{régi}}$ várható érték vektor lineáris transzformáltja:

$$\mathbf{m}_{\text{új}} = \mathbf{A}\ \mathbf{m}_{\text{régi}}\ +\ \mathbf{b}$$

   Ha a régi eloszlás kovariancia mátrixa $\mathbf{C}_{\text{régi}}$, akkor a transzformációval kapott új eloszlás kovariancia mátrixa

$$\mathbf{C}_{\text{új}} = \mathbf{A}\ \mathbf{C}_{\text{régi}}\ \mathbf{A}^T$$

3. *Független vektor valószínűségi változók összegének kovariancia mátrixa.* Független vektor valószínűségi változók összegének kovariancia mátrixa egyenlő a tagok kovariancia mátrixainak összegével.

4. *Független vektor valószínűségi változók összegének kovariancia mátrixa, amikor a tagok kovariancia mátrixai azonosak.* Független vektor valószínűségi változók összegének kovariancia mátrixa, amikor a tagok kovariancia mátrixai azonosak, egyenlő a tagok közös kovariancia mátrixa szorozva a tagok számával.

2. NSZT a vegyes momentumra, kovarianciára, korrelációs együtthatóra

Heurisztikus megfogalmazás: Tekintsünk egy (X, Y) kétdimenziós valószínűségi változót. A kétdimenziós valószínűségi változó eloszlásából ki lehet számolni a vegyes momentumot, a kovarianciát, a korrelációs együtthatót. Ha a kétdimenziós valószínűségi változóra N kísérletet végzünk, és a kísérleti eredmények

$$(X_1, Y_1), (X_2, Y_2), \ldots, (X_N, Y_N)$$

akkor a kísérleti eredményekből adódó kétdimenziós adatrendszer vegyes momentuma, kovarianciája, korrelációs együtthatója szintén kiszámolható. Fontos – matematikailag precízen is igazolható – tény, hogy *nagy kísérletszám esetén a kísérleti eredményekből adódó kétdimenziós adatrendszer vegyes momentuma, kovarianciája, korrelációs együtthatója körülbelül egyenlő az* (X, Y) *kétdimenziós valószínűségi változó eloszlásából kiszámolt vegyes momentumával, kovarianciájával, korrelációs együtthatójával.*

3. Kétdimenziós normális eloszlások alakja *(Extra tananyag)*

3.1. A korrelációs együttható szemléletes jelentése

Ez itt csak rövid emlékeztetője az alőadáson elhangzottaknak. A készülő jegyzetben alaposabban lesz kifejtve. Az előadás gondolatmenetének fontos részét képezi az idevonatkozó Excel fájl, ami a tárgy honlapján elérhető.

A kétdimenziós normális eloszlásokkal kapcsolatban a korrelációs együttható és az eloszlás alakja között kapcsolatot áll fenn. A kapcsolat leírásának céljából tekintsük a

$$(\mu_1 - \sigma_1 \ ; \ \mu_1 + \sigma_1) \quad \text{és a} \quad (\mu_2 - \sigma_2 \ ; \ \mu_2 + \sigma_2)$$

intervallumok direktszorzataként adódó téglalapot, vagy általánosabban a

$$(\mu_1 - s \cdot \sigma_1 \ ; \ \mu_1 + s \cdot \sigma_1) \quad \text{és a} \quad (\mu_2 - s \cdot \sigma_2 \ ; \ \mu_2 + s \cdot \sigma_2)$$

intervallumok direktszorzataként adódó téglalapot, ahol $s = 1$ vagy $s = 2$ vagy $s = 3$. Ezeket a téglalapokat a **szórások alapján felvett téglalapok**nak nevezzük. Képzeljünk el egy-egy skálát -1 -től 1 -ig a téglalapok minegyik oldalán úgy, hogy

- a vízszintes oldalak bal végében a -1, a jobb végében pedig a 1 helyezkedik el
- a függőleges oldalak alsó végében a -1, a felső végében pedig a 1 helyezkedik el

Mindegyik oldalon tekintsük most az r értékhez tartozó pontot, ahol r a korrelációs együttható. Így négy pontot jelöltünk ki a síkon. Ezen a négy ponton át nem nehéz egy olyan ellipszist rajzolni, mely belülről érinti a téglalapot. **Ez az ellipszis a normális eloszlás sűrűségfüggvényének egy színvonala**. Ez az ellipszis meghaározza a normális eloszlás alakját, mert a többi színtvonal ennek az ellipszisnek a kicsinyítsével, vagy nagyításával keletkezik.

3.2. A kovariancia mátrix sajátvektorai, sajátértékei

Egy kétdimenziós normális eloszlás sűrűséfüggvényének színtvonalai olyan ellipszisek, melyek tengelyinek irányát a kovariancia mátrix sajátvektorai adják meg. Az ellipszisek tengelyeinek hosszai pedig arányosak a sajátértékek négyzetgyökeivel.

4. Polinomiális eloszlás közelítése kétdimenziós normális eloszlással *(Extra tananyag)*

Ha az n paraméter elég nagy, és a p_1, p_2 paraméterek se a 0-hoz se az 1-hez nincsenek túl közel, akkor a síkon vett n-ed rendű, (p_1, p_2) paraméterű polinomiális eloszlás közelíthető kétdimenziós normális eloszlással. A normális eloszlás paramétereit a polinomiális eloszlás paramétereihez úgy kell hozzáigazítani, hogy a megfelelő paraméterek (várható értékek, szórások, korrelációs együttható) megegyezzenek, vagyis

$$\mu_1 = np_1 \qquad \mu_2 = np_2 \qquad \sigma_1 = \sqrt{np_1(1-p_1)} \qquad \sigma_2 = \sqrt{np_2(1-p_2)} \qquad r = -\sqrt{\frac{p_1}{(1-p_1)}\,\frac{p_2}{(1-p_2)}}$$

5. Többdimenziós normális eloszlások transzformációi *(Extra tananyag)*

5.1. A standard normális eloszlás lineáris transzformációi

Jön majd ide

5.2. Tetszőleges normális eloszlás lineáris transzformációi

Jön majd ide

5.3. Khí-négyzet eloszlás

Ez itt csak rövid emlékeztetője az alőadáson elhangzottaknak. A készülő jegyzetben alaposabban lesz kifejtve. Az előadás gondolatmenetének fontos részét képezi az idevonatkozó Excel fájl, ami a tárgy honlapján elérhető.

Ha n darab független standard normális eloszlású valószínűségi változó négyzetét összeadjuk, akkor egy valószínűségi változóhoz jutunk. Ennek a valószínűségi változónak az eloszlását n **szabadságfokú khí-négyzet eloszlás**nak hívjuk.

6. Lineáris regresszió

6.1. Általános probléma

Tekintsünk egy (X, Y) kétdimenziós valószínűségi változót. Célunk hogy az, hogy megtaláljuk azokat az A és B konstansokat, melyekre az

$$\mathrm{E}\left((Y - (AX + B))^2 \right) = \iint\limits_{R^2} (y - (Ax + B))^2 \; f(x, y)\, dx\, dy$$

várható érték miniális.

Megoldás: Ha a fenti integrálban elvégezzük a négyzetre emelést, és az integrált tagokra bontjuk, akkor hat tagot kapunk:

$$\mathrm{E}\left((Y - (AX + B))^2 \right) =$$

$$= \iint\limits_{R^2} y^2 \; f(x, y)\, dx\, dy + A^2 \iint\limits_{R^2} x^2 \; f(x, y)\, dx\, dy + B^2 \iint\limits_{R^2} f(x, y)\, dx\, dy -$$

$$-2A \iint\limits_{R^2} xy \; f(x, y)\, dx\, dy - 2B \iint\limits_{R^2} y \; f(x, y)\, dx\, dy + 2AB \iint\limits_{R^2} x \; f(x, y)\, dx\, dy$$

Az itt szereplő integrálok értékei konstansokat adnak, ezért A -ra, B -re nézve egy kvadratikus formulát kaptunk. Azokat az A, B értékeket, melyreke ez a kvadratikus formula minimális, megkaphatjuk például úgy, hogy parciálisan deeriválunk A és B szerint, majd a parciális deriváltakat nullával tesszük egyenlővé, és az így kiadódó egyenletrendszert A -ra, B -re megoldjuk. A számolás részleteitől eltekintünk, csak az eredményt közöljük:

$$A_{\text{opt}} = r \; \frac{\sigma_2}{\sigma_1}$$

$$B_{\text{opt}} = \mu_2 \; - \; r \frac{\sigma_2}{\sigma_1} \; \mu_1$$

Az optimumot nyújtó lineáris függvény tehát így fest:

$$y = \mu_2 + r \; \frac{\sigma_2}{\sigma_1} \; (x - \mu_1)$$

Az egyenletet átírhatjuk erre a – könnyebben megjegyezhető – formára is:

$$\frac{y - \mu_2}{\sigma_2} = r \; \frac{x - \mu_1}{\sigma_1}$$

A kapott egyenes neve: **regressziós egyenes**.

6.2. Számolós példa

Legyen $X = \mathrm{RND}_1 \cdot \mathrm{RND}_2$ és $Y = \mathrm{RND}_1$. Határozzuk meg

- X, Y, XY várható értékét,
- X, Y varianciáját, szórását,
- az X és közötti Y kovarianciáját és a korrelációs együtthatót!
- Írjuk fel a regressziós egyenes egyenletét, ha Y -ból akarunk X -re tippelni!
- Írjuk fel a regressziós egyenes egyenletét, ha X -ből akarunk Y -ra tippelni!

Megoldás:

$$\mathrm{E}(X) = \mathrm{E}(\mathrm{RND}_1 \cdot \mathrm{RND}_2) = \mathrm{E}(\mathrm{RND}_1) \cdot \mathrm{E}(\mathrm{RND}_2) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

$$\mathrm{E}\left(X^2\right) = \mathrm{E}\left(\mathrm{RND}_1^2 \cdot \mathrm{RND}_2^2\right) = \mathrm{E}\left(\mathrm{RND}_1^2\right) \cdot \mathrm{E}\left(\mathrm{RND}_2^2\right) = \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}$$

$$\mathrm{VAR}(X) = \mathrm{E}\left(X^2\right) - (\mathrm{E}(X))^2 = \frac{1}{9} - \left(\frac{1}{4}\right)^2 = \frac{7}{144}$$

$$\mathrm{SD}(X) = \sqrt{\mathrm{VAR}(X)} = \sqrt{\frac{7}{144}}$$

$$\mathrm{E}(Y) = \mathrm{E}(\mathrm{RND}_1) = \frac{1}{2}$$

$$\mathrm{E}\left(Y^2\right) = \mathrm{E}\left(\mathrm{RND}_1^2\right) = \frac{1}{3}$$

$$\mathrm{VAR}(Y) = \mathrm{E}\left(Y^2\right) - (\mathrm{E}(Y))^2 = \frac{1}{3} - \left(\frac{1}{2}\right)^2 = \frac{1}{12}$$

$$\mathrm{SD}(Y) = \sqrt{\mathrm{VAR}(Y)} = \sqrt{\frac{1}{12}}$$

$$\mathrm{E}(X \cdot Y) = \mathrm{E}(\mathrm{RND}_1\mathrm{RND}_2 \cdot \mathrm{RND}_1) = \mathrm{E}\left(\mathrm{RND}_1^2 \cdot \mathrm{RND}_2\right) = \mathrm{E}\left(\mathrm{RND}_1^2\right) \cdot \mathrm{E}(\mathrm{RND}_2) = \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$$

$$\mathrm{COV}(X;Y) = \mathrm{E}(X \cdot Y) - \mathrm{E}(X) \cdot \mathrm{E}(Y) = \frac{1}{6} - \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{24}$$

$$\mathrm{CORR}(X;Y) = \frac{\mathrm{COV}(X;Y)}{\mathrm{SD}(X)\,\mathrm{SD}(Y)} = \frac{\frac{1}{24}}{\sqrt{\frac{7}{144}}\sqrt{\frac{1}{12}}} = \sqrt{\frac{3}{7}}$$

$$y = \mu_2 + r\,\frac{\sigma_2}{\sigma_1}\,(x - \mu_1)$$

$$y = \frac{1}{2} + \sqrt{\frac{3}{7}}\,\frac{\sqrt{\frac{1}{12}}}{\sqrt{\frac{7}{144}}}\,\left(x - \frac{1}{4}\right)$$

$$y = \frac{6}{7}x + \frac{2}{7}$$

7. Statisztika *(Extra tananyag)*

7.1. Khí-négyzet próba

Tekintsünk egy r elemű teljes eseményrendszert. Az egyes események valószínűségei legyenek

$$p_1 \ , \ p_2 \ , \ \ldots \ , \ p_r$$

Ha N kísérletet végzünk, akkor nézhetjük, hogy a teljes eseményrendszer egyes elemei hányszor következnek be. Legyenek ezek a gyakoriság értékek:

$$X_1 \ , \ X_2 \ , \ \ldots \ , \ X_r$$

Legyenek

$$p_1^{\text{hip}}, p_2^{\text{hip}}, \ldots, p_r^{\text{hip}}$$

valamilyen **hipotetikus valószínűségek**, vagyis olyan valószínűségek, melyeket valaki csak úgy gondol, képzel, és épp a kísérleti eredményekből akarja eldönteni, hogy az események igazi valószínűségei mind megegyeznek-e az általa gondolt hipotetikus valószínűségekkel:

$$p_1 = p_1^{\text{hip}} \ , \quad p_2 = p_2^{\text{hip}} \ , \ \ldots \ , \ p_r = p_r^{\text{hip}} \qquad ?$$

Az

$$X_1 \ , \ X_2 \ , \ \ldots \ , \ X_r$$

valószínűségi változók mindegyike binomiális eloszlást követ, ezért a várhatató értékeik:

$$Np_1 \ , \quad Np_2 \ , \ \ldots \ , \ Np_r$$

Ha a tényleges valószínűségek mind megegyeznek a hipotetikus valószínűségekkel, akkor az

$$X_1 \ , \ X_2 \ , \ \ldots \ , \ X_r$$

gyakoriságok várható értékei:

$$Np_1^{\text{hip}} \ , \quad Np_2^{\text{hip}} \ , \ \ldots \ , \ Np_r^{\text{hip}}$$

Ezért ezeket a szorzatokat **hipotetikus várható értékek**nek nevezzük.

Tekintsük a gyakoriságoknak a hipotetikus várható értékektől való eltérésének a négyzetét:

$$\left(X_1 - N\,p_1^{\text{hip}}\right)^2 \ , \ \left(X_2 - N\,p_2^{\text{hip}}\right)^2 \ , \ \ldots \ , \ \left(X_r - N\,p_r^{\text{hip}}\right)^2$$

aztán az alábbi hányadosokat:

$$\frac{\left(X_1 - N\,p_1^{\text{hip}}\right)^2}{Np_1^{\text{hip}}} \ , \ \frac{\left(X_2 - N\,p_2^{\text{hip}}\right)^2}{Np_2^{\text{hip}}} \ , \ \ldots \ , \ \frac{\left(X_r - N\,p_r^{\text{hip}}\right)^2}{Np_r^{\text{hip}}}$$

s végül ezeknek az összegét, amit $\chi^2_{\text{megfigyelt}}$ -tel jelölünk, és **megfigyelt khí-négyzet érték**nek nevezünk:

$$\chi^2_{\text{megfigyelt}} \ = \ \frac{\left(X_1 - N\,p_1^{\text{hip}}\right)^2}{Np_1^{\text{hip}}} + \frac{\left(X_2 - N\,p_2^{\text{hip}}\right)^2}{Np_2^{\text{hip}}} + \ldots + \frac{\left(X_r - N\,p_r^{\text{hip}}\right)^2}{Np_r^{\text{hip}}}$$

Tegyük fel, hogy az N kísérletszám annyira nagy, hogy hipotetikus várható értékek mindegyike nagyobb 10 -nél. (Ha ez a feltétel nem teljesül, akkor az alábbi eljárás – pontatlansága miatt – nem használható.)

Be lehet látni, hogy az alábbiak igazak:

- *Ha a tényleges valószínűségek mind megegyeznek a hipotetikus valószínűségekkel, akkor a megfigyelt khí-négyzet érték közelítőleg $r-1$ szabadságfokú khí-négyzet eloszlást követ.*
- *Ha pedig a tényleges valószínűségek nem mind egyeznek meg a hipotetikus valószínűségekkel, akkor a megfigyelt khí-négyzet érték eloszlása jelentősen eltér az $r-1$ szabadságfokú khí-négyzet eloszlástól, és a megfigyelt khí-négyzet érték általában nagy.*

Ezeket a tényeket érezhetővé tehetjük, ha a megfigyelt khí-négyzet értéket az alábbi alakra hozzuk:

$$\chi^2_{\text{megfigyelt}} = N \cdot \left[\frac{\left(\frac{X_1}{N} - p_1^{\text{hip}}\right)^2}{p_1^{\text{hip}}} + \frac{\left(\frac{X_2}{N} - p_2^{\text{hip}}\right)^2}{p_2^{\text{hip}}} + \ldots + \frac{\left(\frac{X_r}{N} - p_r^{\text{hip}}\right)^2}{p_r^{\text{hip}}} \right]$$

Ugyanis ebből az alakból láthatóak, hihetőek az alábbiakiak:

- Ha a tényleges valószínűségek mind megegyeznek a hipotetikus valószínűségekkel, akkor minden i -re az $\frac{X_i}{N}$ relatív gyakoriság – nagy N esetén – közel van a van a p_i^{hip} hipotetikus valószínűséghez, és így a most felírt kifejezés szögletes zárójelében álló tagok mind kicsik. N növekedtével a tagok még kisebbek, hiszen a relatív gyakoriságok még közelebb kerülnek a valószínűséghez. Emiatt elhihető, elképzelhető, hogy a szögletes zárójelben lévő összeg értéke N -nel megszorozva (ahogy a megfigyelt khí-négyzet értéket kapjuk) olyan véletlen érték, ami közelítőleg egy jól meghatározott eloszlást követ. Ez az eloszlás történetesen az $r-1$ szabadságfokú khí-négyzet eloszlás. Tehát ilyenkor a megfigyelt khí-négyzet érték közelítőleg $r-1$ szabadságfokú khí-négyzet eloszlást követ.
- Ha pedig a tényleges valószínűségek nem mind egyeznek meg a hipotetikus valószínűségekkel, akkor van olyan i, hogy az $\frac{X_i}{N}$ relatív gyakoriság – nagy N esetén – a p_i^{hip} hipotetikus valószínűségtől különböző p_i -hez lesz közel, és így az
$$\frac{\left(\frac{X_i}{N} - p_i^{\text{hip}}\right)^2}{p_i^{\text{hip}}}$$
kifejezés értéke a pozitív
$$\frac{\left(p_i - p_i^{\text{hip}}\right)^2}{p_i^{\text{hip}}}$$
szám közelében lesz. Mivel a szögeletes zárójelben álló kifejezés a $\frac{\left(\frac{X_i}{N} - p_i^{\text{hip}}\right)^2}{p_i^{\text{hip}}}$ tagnál nagyobb, a szögletes zárójel előtt álló N -nel való szorzás – nagy N esetén – a megfigyelt khí-négyzet értéket naggyá teszi. Ha N nagyobb, akkor még nagyobbá. Tehát ilyenkor – nagy N esetén – a megfigyelt khí-négyzet érték szeret nagy lenni.

Az $r-1$ szabadságfokú khí-négyzet eloszlás segítségével tudunk választani egy χ^2_{kritikus} -sal jelölt ún. **kritikus khí-négyzet érték**et, mely arról nevezetes, hogy az $r-1$ szabadságfokú khí-négyzet eloszlást követő valószínűségi változó 0.95 valószínűséggel kisebb, 0.05 valószínűséggel nagyobb ennél a χ^2_{kritikus} -sal jelölt értéknél. (Arról, hogy itt miért éppen a 0.05 valószínűség érték áll, és nem valami más, később ejtünk szót.) A χ^2_{kritikus} értéket az $F(x)$ baloldali- vagy a $T(x)$ jobboldali eloszlásfüggvény inverzével kapjuk a

$$\chi^2_{\text{kritikus}} = F^{-1}(0.95) \qquad \chi^2_{\text{kritikus}} = T^{-1}(0.05)$$

összefüggések valamelyikéből.

Ennek segítségével így dönthetünk afelől, hogy a tényleges valószínűségek mind megegyeznek-e a hipotetikus valószínűségekkel:

- **Ha a megfigyelt khí-négyzet érték kisebb, mint a kritikus khí-négyzet érték, $\chi^2_{\text{megfigyelt}} < \chi^2_{\text{kritikus}}$, akkor arra tippelünk, hogy a tényleges valószínűségek mind megegyeznek a hipotetikus valószínűségekkel.**
- **Ha a megfigyelt khí-négyzet érték nagyobb mint a kritikus khí-négyzet érték, $\chi^2_{\text{megfigyelt}} > \chi^2_{\text{kritikus}}$, akkor arra tippelünk, hogy a tényleges valószínűségek nem mind egyeznek meg a hipotetikus valószínűségekkel.**

Ez az eljárás – a fentebb mondottak miatt – ilyen:

- *Ha a tényleges valószínűségek mind megegyeznek-e a hipotetikus valószínűségekkel, akkor kb.* 0.95 *valószínűséggel a megfigyelt khí-négyzet érték kisebb, mint a kritikus khí-négyzet érték, és így az eljárás helyesen dönt, és csak kb.* 0.05 *valószínűséggel lesz a megfigyelt khí-négyzet érték nagyobb, mint a kritikus khí-négyzet érték, ami miatt az eljárás hibázik.*
- *Ha pedig a tényleges valószínűségek nem mind egyeznek meg a hipotetikus valószínűségekkel, akkor nagy valószínűséggel a megfigyelt khí-négyzet érték nagyobb, mint a kritikus khí-négyzet érték, és így az eljárás helyesen dönt, és csak kis valószínűséggel lesz a megfigyelt khí-négyzet érték kisebb, mint a kritikus khí-négyzet érték, ami miatt az eljárás hibázik.* (Azt, hogy a "nagy valószínűséggel", "kis valószínűséggel" kifejezések pontosan mit jelentenek, itt nem vizsgáljuk.)

Eléggé elterjedt az alkalmazók körében, hogy a kritikus értéket úgy választják meg, hogy a hipotézis fennállása esetén a hipotézis elfogadásának (a helyes döntésnek) a valószínűsége 0.95, az elutasítás (a hibás döntésnek) a valószínűsége 0.05 legyen. A 0.95, 0.05 értékek kitüntetett szerepét igazából semmi sem indokolja. Egyes esetekben más **elfogadási-** és **elutasítási valószínűség**ek használata is indokolt lehet. Az elfogadási- és elutasítási valószínűségek megválasztása nem matematikai kérdés, az alkalmazók szubjektív választásán múlik.

- Ha valakinek az a fontos, hogy az eljárás alkalmazásakor a hipotézis fennállása esetén ritkábban tévedjen, akkor az eljárást a 0.95 helyett nagyobb – mondjuk – 0.99 elfogadási- és 0.01 elutasítási valószínűség értékkel – és ennek megfelelően – nagyobb kritikus értékkel kell használnia.
- Ha pedig az a fontos, hogy lehetőleg ne tévedjünk, amikor a hipotézis nem áll fenn, akkor 0.95 helyett kisebb – mondjuk – 0.90 elfogadási és 0.10 elutasítási valószínűség értékkel és így kisebb kritikus értékkel kell dolgoznunk.

A két igény, hogy se a hipotézis fennállása, se a hipotézis nem fennállása esetén ne tévedjünk, egymásnak ellentmond. Az egyik igény javítása csak a másik kárára történhet.

7.1.1. Prüntyőke fiókák száma

Már rég megfigyelték, hogy a prüntyőke madarak fészkében tavaszi költéskor a tojások száma 2, 3 vagy 4. Mostanában olvatam egy újságban, hogy valaki – elméleti kutatómunkája eredményeként – azt hozta ki, hogy

P(2 tojás) = 0.2   P(3 tojás) = 0.5   P(2 tojás) = 0.3

Én – gyakorlati oldalról – kissé utánanéztem a dolognak. Nem kis munkával 90 fészket kutattam fel. 12 esetben 2 tojást, 43 esetben 3 tojást, 35 esetben 4 tojást láttam. Tehát a korábbi jelölésekkel:

$$r = 3\,, \quad p_1^{\text{hip}} = 0.2\,, \quad p_2^{\text{hip}} = 0.5\,, \quad p_3^{\text{hip}} = 0.3\,, \quad N = 90\,, \quad X_1 = 12\,, \quad X_2 = 43\,, \quad X_3 = 35$$

A megfigyelt khí-négyzet értéket könnyűszerrel kiszámoltam: $\chi^2_{\text{megfigyelt}} = 4.46$. Ezt összevetve a 0.95 -ös elfogadási valószínűséghez tartozó $\chi^2_{\text{kritikus}} = 5.99$ -es kritikus értékkel, levontam a következtetést. Mivel $\chi^2_{\text{megfigyelt}} < \chi^2_{\text{kritikus}}$, a megfigyelésem nem mond ellent a tudós állításának, ezért elfogadtam a tudós által kijelentett valószínűség értékeket.

1. Megjegyzés: Fontos, hogy a 0.95 -ös elfogadási valószínűség jelentését ne értsük félre. A 0.95 nem azt jelenti, hogy a tudós állítása 0.95 valószínűséggel igaz. Azt sem jelenti, hogy az én ítéletem 0.95 valószínűséggel helyes. A 0.95 -ös elfogadási valószínűség az általam használt döntési eljárásnak a jellemzője: a döntési eljárás, amit én itt most egyszer használtam olyan, hogy ha igazak a

$$\text{P(2 tojás) = 0.2} \qquad \text{P(3 tojás) = 0.5} \qquad \text{P(2 tojás) = 0.3}$$

valószínűségek, akkor

- 0.95 a valószínűsége annak, hogy a megfigyelt khí-négyzet érték az 5.99 -es kritikus értéknél kisebbnek adódik
- 0.05 a valószínűsége annak, hogy a megfigyelt khí-négyzet érték az 5.99 -es kritikus értéknél nagyobbnak adódik

Másszóval

- 0.95 a valószínűsége annak, hogy az eljárás helyesen dönt
- 0.05 a valószínűsége annak, hogy az eljárás tévesen dönt

2. Megjegyzés: Megjegyezzük, hogy ha a 90 fészek közül 8 -ban 2 tojást, 47 -ben 3 tojást, 35 -ben 4 tojást találtam volna, akkor a megfigyelt khí-négyzet érték 8.01 -nek adódott volna, ami – 0.95 -ös elfogadási valószínűség mellett – már a tudós állításának az elutasítást indikálná.

3. Megjegyzés: Ámde a 0.95 -nél nagyobb 0.99 -es elfogadási valószínűség mellett ugyanez a 8.01 -es megfigyelt khí-négyzet érték – mivel a kritikus érték 9.21 – a tudós állításának az elfogadást indikálná. Nem meglepő, hogy ha növeljük az elfogadási valószínűséget, akkor elfogadóbbá válunk.

7.1.2. Madárfiókák helyett színes golyók

Lehet, hogy van, akinek segít a fenti probléma megemésztésében az alábbi könnyen elképzelhető modell, mely ekvivalens az előző pontban tálalt problémával.

Tegyük fel, hogy egy dobozban 10 színes golyó van: pirosak, fehérek és zöldek. Valaki azt állítja, hogy

$$2 \text{ piros, } 5 \text{ fehér és } 3 \text{ zöld golyó van a dobozban}$$

Én nem nézhetek bele a dobozba, viszont csukott szemmel húzhatok a dobozból – mondjuk – 90 -szer, és – mielőtt visszateszem a golyókat, mindig megnézhetem a színüket.

Tegyük fel, hogy a 90 húzásból 12 esetben pirosat, 43 esetben fehéret, 35 esetben zöldet húztam. Elfogadjam-e, hogy 2 piros, 5 fehér és 3 zöld golyó van a dobozban?

Az okoskodás pontosan ugyanaz lenne, mint a prüntyőke fiókák számával kapcsolat fentebb volt. A

$$r = 3\,, \quad p_1^{\text{hip}} = 0.2\,, \quad p_2^{\text{hip}} = 0.5\,, \quad p_3^{\text{hip}} = 0.3\,, \quad N = 90\,, \quad X_1 = 12\,, \quad X_2 = 43\,, \quad X_3 = 35$$

adatokból kiszámolt megfigyelt khí-négyzet érték: $\chi^2_{\text{megfigyelt}} = 4.46$, amit összevetve a 0.95 -ös elfogadási valószínűséghez tartozó $\chi^2_{\text{kritikus}} = 5.99$ -es kritikus értékkel, elfogadtam a doboz állított összetételét.

7.1.3. Egyforma esélyűek a lottószámok?

1956 óta 3120 húzás volt az ötöslottón. Az alábbi táblázat azt mutatja, hogy az eltelt 60 év során az egyes számokat hányszor húzták ki:

1	188	16	166	31	156	46	171	61	166	76	185
2	159	17	158	32	172	47	186	62	162	77	195
3	202	18	187	33	171	48	153	63	145	78	179
4	176	19	178	34	174	49	177	64	180	79	166
5	147	20	178	35	178	50	169	65	174	80	154
6	163	21	163	36	176	51	185	66	195	81	179
7	180	22	185	37	172	52	177	67	173	82	157
8	156	23	183	38	176	53	164	68	159	83	179
9	160	24	185	39	157	54	178	69	191	84	176
10	193	25	177	40	158	55	181	70	158	85	176
11	178	26	168	41	171	56	192	71	181	86	194
12	185	27	174	42	194	57	175	72	183	87	148
13	191	28	162	43	178	58	155	73	185	88	135
14	163	29	201	44	161	59	177	74	164	89	154
15	190	30	157	45	175	60	177	75	197	90	171

Látjuk, hogy a 3 -as számot 202 -szer, a 88 -at pedig csak 135 -ször. Mivel a 202 a 135 -nek másfélszerese, felmerülhet bennünk a gyanu, hogy a lottószámok nem egyformán valószínűek. Pedig úgy illene, hogy a 90 lottó szám egyformán valószínű legyen!

Ha a hipotézisünk az, hogy a lottószámok húzásánál minden szám valószínűsége $\frac{1}{90}$, akkor a fenti táblázatban olvasható gyakoriságokkal számolva (aki nem hiszi, számoljon utána) a megfigyelt khí-négyzet érték:

$$\chi^2_{\text{megfigyelt}} = 95.91$$

A 89 szabadságfogú khí-négyzet eloszlásból a 0.95 elfogadási szinthez tartozó kritikus érték

$$\chi^2_{\text{kritikus}} = F^{-1}(0.95) = 112.02$$

Mivel a megfigyelt khí-négyzet érték kisebb, mint a kritikus érték, a táblázat adatai nem mondanak ellent a hipotézisnek, elfogadhatjuk, hogy a lottószámok egyformán valószínűek.

7.2. U-próba

7.2.1. Döntés az átlag alapján

Tegyük fel, hogy a férfiak testmagassága a Föld minden országában normális eloszlást követ 10 cm szórással és valamilyen – az országra jellemző – várható értékkel. A várható érték más-más lehet: pl. a svédek magasabbak, a pigmeusok alacsonyabbak, de a szórásról azt feltételezzük, hogy minden országban $= 10$ cm.

Egy marslakó leszáll a Földre, kinéz az űrhajóból, és látja a nagy plakátot: "Ebben az országban a férfiak tesmagasságának a várható értéke 180 cm". A marslakó el akarja dönteni, hogy a plakát igazat beszél-e. Tanult valószínűségszámítást, tudja, hogy a várható érték az országos átlagot jelenti. Tisztában van vele, hogy az ország minden férfi emberének magasságát nem tudja megmérni, ezért 4 véletlenül odaverődő, bámészkodó férfi testmagasságát megméri. Tegyük fel, hogy ezeket az adatokat kapja: 177.3, 174.8, 184.2, 178.5 cm. Kiszámolja az átlagot, ami 178.7.

Vakarja is ám a fejét a marslakó: mert a 178.7 nem 180, ahogy a plakáton látja. Először – elhamarkodottan – azt gondolja: lám, hazudnak! De hamar rájön, hogy nagy butaság lenne, ha csak akkor adna igazat a plakáton látottaknak, ha a 4 megmért ember magasságának az átlaga pontosan 180 lenne! Hiszen annak a valószínűsége, hogy egy folytonos valószínűségi változó, mint amilyen a négy mérési eredmény átlaga, pontosan egyenlő egy adott értékkel – mint azt jól megtanulta – nullával egyenlő.

Gondolkodni kezd. A 178.7 -nek és a 180 -nak az eltérése nem túl nagy, ezért elképzelhető, hogy a várható érték mégis 180. Vagy mégse? Hiszen a 178.7 -nek és a 180 -nak az eltérése azért mégse olyan kicsike, hiszen 1 -nél azért csak nagyobb.

Segítsünk a nagy hezitálásban szenvedő marslakónak! Szeretné elkerülni, hogy hibázzon.

- Hiba lenne a 180 -as országos átlag teljesülése esetén elutasítani a 180 -as átlagot.
- De hiba lenne az is, hogy ha az országos átlag nem 180, és ő mégis elfogadja a plakáton reklámozott 180 -as átlagot.

Mivel a marslakó döntési eljárásába a véletlen is beleszól (véletlentől függ, hogy az ország sok lakosa közül éppen ki az a 4 odaverődő, bámészkodó ember, akiket megmér), nyilvánvaló, hogy bele kell törödnie abba, hogy esetleg hibázik.

Naív álmai szerint azt szeretné, hogy a döntési eljárás olyan legyen, hogy

- ha az országos átlag 180, akkor a döntési eljárás 1 valószínűséggel a 180 mellett döntsön,
- ha az országos átlag különbözik 180 -tól, akkor a döntési eljárás 0 valószínűséggel döntsön a 180 mellett.

Hamar ráébred, hogy meg kell elégednie egy olyan döntési eljárással, ami olyan, hogy

1. ha az országos átlag 180, akkor a döntési eljárás 1-hez közeli p_0 valószínűséggel a 180 mellett dönt,
2. ha az országos átlag különbözik 180 -tól, akkor a döntési eljárás p_0 -nál kisebb valószínűséggel dönt a 180 mellett.
3. ha az országos átlag nagyon különbözik 180 -tól, akkor a döntési eljárás csak kis valószínűséggel dönt a 180 mellett.

Mindebből kirajzolódik, hogy választani kell egy alkalmas $\Delta\mu_0$ értéket (hogy hogyan, arra lejjebb rávilágítunk), és

- ha az N kísérleti eredmény $overlineX_N$ átlagának és a hipotetikus $\mu_0 = 180$ értéknek az egymástól való távolsága kisebb $\Delta\mu_0$ -nél, azaz
$$\left| \, \overline{X}_N - \mu_0 \, \right| < \Delta\mu_0$$
akkor elfogadjuk a hipotézist, vagyis azt, hogy a plakát igazat mond,
- ha pedig
$$\left| \, \overline{X}_N - \mu_0 \, \right| \; > \; \Delta\mu$$
akkor elutasítjuk a hipotézist, vagyis azt mondjuk, hogy a plakát nem mond igazat.

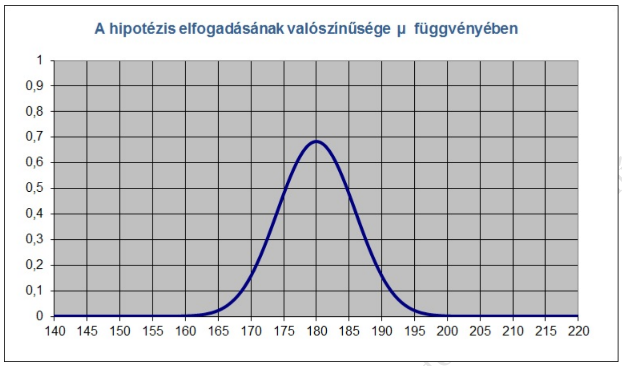
A hipotézis elfogadásának a
$$\mathrm{P}\left(\left| \, \overline{X}_N - \mu_0 \, \right| \; < \; \Delta\mu_0 \right)$$

valószínűsége nyilván a μ paraméter függvénye, hiszen az $\overline{X}_N$ átlag olyan valószínűségi változók átlaga, melyek μ várható értékű normális eloszlást követnek! A függvény képletének felírása céljából említjük, hogy $\overline{X}_N$ normális eloszlást követ μ várható értékkel és $\frac{\sigma}{\sqrt{N}}$ szórással, ezért a hipotézis elfogadásának a valószínűsége:

$$\mathrm{P}\left(\left| \overline{X}_N - \mu_0 \right| < \Delta\mu_0 \right) = \mathrm{P}\left(\mu_0 - \Delta\mu_0 < \overline{X}_N < \mu_0 + \Delta\mu_0 \right) =$$

$$= \Phi\left(\frac{(\mu_0 + \Delta\mu_0) - \mu}{\frac{\sigma}{\sqrt{N}}} \right) - \Phi\left(\frac{(\mu_0 - \Delta\mu_0) - \mu}{\frac{\sigma}{\sqrt{N}}} \right)$$

Ennek a függvénynek a grafikonját a mellékelt ábrán a $\mu_0 = 180,\ \sigma = 10,\ N = 4,\ \Delta\mu_0 = 5$ paraméter értékek mellett megadjuk $140 < \mu < 220$ -re.



1. ábra. *Elfogadás valószínűsége* $\mu_0 = 180$, $\sigma = 10$, $N = 4$, $\Delta\mu_0 = 5$ *esetén*

Az ábráról leolvassuk, hogy a hipotézis elfogadásának a valószínűsége $\mu = \mu_0 = 180$ esetén a legnagyobb. Azt is látjuk, hogy ha μ távolodik μ_0 -tól, akkor a hipotézis elfogadásának a valószínűsége egyre csökken és 0 -hoz tart.

A hipotézis elfogadásának a valószínűsége $\mu = \mu_0$ esetén külön figyelmet érdemel. Jelölésére a p_0 -t vezetjük be.

A p_0 valószínűség könnyen kifejezhető a $\sigma,\ N,\ \Delta\mu_0$ paraméterekkel, hiszen ha $\mu = \mu_0$, akkor

$$\Phi\left(\frac{(\mu_0 + \Delta\mu_0) - \mu}{\frac{\sigma}{\sqrt{N}}} \right) - \Phi\left(\frac{(\mu_0 - \Delta\mu_0) - \mu}{\frac{\sigma}{\sqrt{N}}} \right) = \Phi\left(\frac{\Delta\mu_0}{\frac{\sigma}{\sqrt{N}}} \right) - \Phi\left(\frac{-\Delta\mu_0}{\frac{\sigma}{\sqrt{N}}} \right) =$$

$$= 2\,\Phi\left(\frac{\Delta\mu_0}{\frac{\sigma}{\sqrt{N}}} \right) - 1$$

Tehát

$$p_0 = 2\,\Phi\left(\frac{\Delta\mu_0}{\frac{\sigma}{\sqrt{N}}} \right) - 1$$

Ebből az egyenletből $\Delta\mu_0$ is és N is kifejezhető a többi paraméter segítségével, íme:

$$\Delta\mu_0 = \frac{\sigma}{\sqrt{N}} \cdot \Phi^{-1}\left(\frac{1 + p_0}{2} \right)$$

$$N = \left[\frac{\sigma}{\Delta\mu_0} \cdot \Phi^{-1}\left(\frac{1 + p_0}{2} \right) \right]^2 \qquad \textit{felfelé kerekített értéke}$$

ahol Φ^{-1} a Φ függvény inverzét jelenti.

Ezen a ponton az internet és a számítógép adta lehetőségeket hívjuk segítségül. Az Olvassó nyissa meg a

`http://math.bme.hu/~vetier/2017_osz/A4_vill_2017_osz.html` honlapcímen a lap felső részében a mellékletek között található

`U-próba Elfogadás valószínűsége` nevű linkkel hívható

`U-proba__Elfogadas_vsz-e.xlsx` című Excel fájlt,

és a zöld hátterű cellákban található paraméterek változtatásával vizsgálja meg hogyan változik a hipotézis elfogadása valószínűségének az ábrája. Ebben a nyomtatott jegyzetben az ábrák az alábbi paraméter értékek mellett láthatóak:

1. $\mu_0 = 180$, $\sigma = 10$, $N = 4$, $\Delta\mu_0 = 5$
2. $\mu_0 = 180$, $\sigma = 10$, $N = 4$, $\Delta\mu_0 = 10$
3. $\mu_0 = 180$, $\sigma = 10$, $N = 4$, $\Delta\mu_0 = 2$
4. $\mu_0 = 180$, $\sigma = 10$, $N = 20$, $\Delta\mu_0 = 2$
5. $\mu_0 = 180$, $\sigma = 10$, $N = 100$, $\Delta\mu_0 = 2$

Az első, a második és a harmadik esetben $\mu_0 = 180$, $\sigma = 10$, $N = 4$, és csak a $\Delta\mu_0$ értéke változik:

- $\Delta\mu_0 = 5$
- $\Delta\mu_0 = 10$
- $\Delta\mu_0 = 2$

A harmadik, a negyedik és az ötödik esetben $\mu_0 = 180$, $\sigma = 10$, $\Delta\mu_0 = 2$, és csak az $N = 4$ értéke változik:

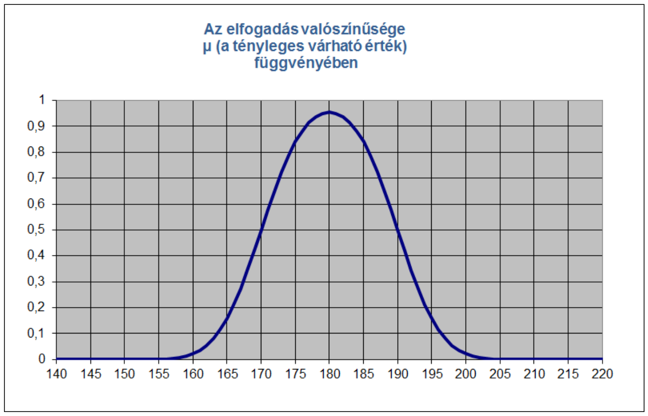
- $N = 4$
- $N = 20$
- $N = 100$

Az Olvasó feladata észrevenni, hogy a N növelésével és $\Delta\mu_0$ csökkentésével elérhető, hogy a grafikon a $\mu_0 = 180$ helyen megközelítse az 1 értéket, egyéb helyeken pedig a 0 -t.

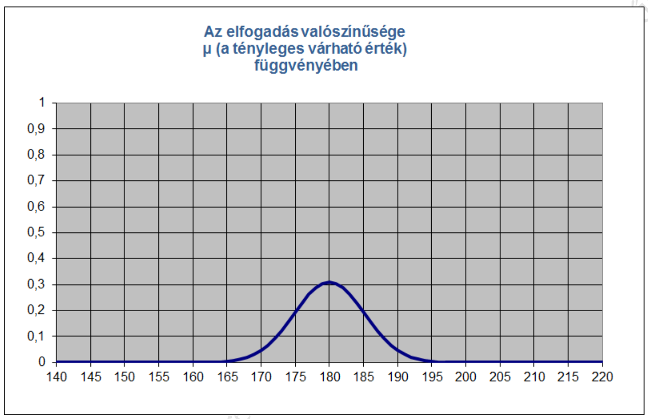
Bár a marslakó naív álmait nem lehet teljesen kielégíteni, az álmokat megközelíteni egészen jól lehet! Ennek persze ára van: a kísérletek számát kell kellően nagyra venni, amit kimondani könnyű, de a gyakorlatban megvalósítani sok esetben nem lehet. Bele kell törődni, hogy a valóság az álmoktól néha messze van. Ilyenkor általában azt szokták csinálni, hogy szubjektív igények alapján először megválasztják a p_0 értéket. Ismételjük: a p_0 megválasztásával azt szabályozzuk be, hogy a hipotézis fennállása, azaz $\mu = \mu_0$ esetén a döntési eljárásunk milyen valószínűséggel fog helyesen, azaz a hipotézis fennállása mellett dönteni. $1 - p_0$ valószínűséggel sajnos hibázni fog, mert a hipotézis tagadása mellett dönt. Valamiért szinte az egész világon a $p_0 = 0.95$ a leghasználatosabb érték. Egyéb μ -k esetén hipotézis elfogadása, ami hibának számit, olyan (a μ tényleges értékétől függő!) valószínűséggel következik be, ahogy azt fentebb kiszámoltuk, és ahogy az

`U-proba__Elfogadas_vsz-e.xlsx` című Excel fájl

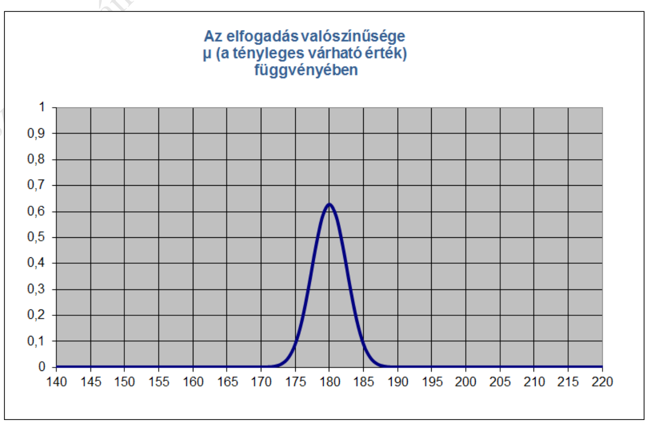
azt kirajzolja.



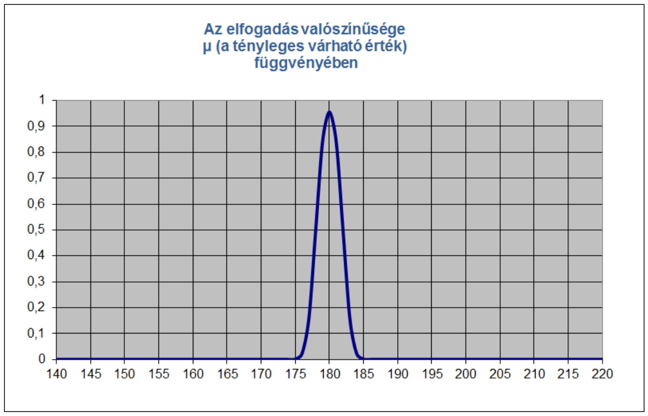
2. ábra. *Elfogadás valószínűsége* $\mu_0 = 180$, $\sigma = 10$, $N = 4$, $\Delta\mu_0 = 10$ *esetén*



3. ábra. *Elfogadás valószínűsége* $\mu_0 = 180$, $\sigma = 10$, $N = 4$, $\Delta\mu_0 = 2$ *esetén*



4. ábra. *Elfogadás valószínűsége* $\mu_0 = 180$, $\sigma = 10$, $N = 20$, $\Delta\mu_0 = 2$ *esetén*



5. ábra. *Elfogadás valószínűsége* $\mu_0 = 180$, $\sigma = 10$, $N = 100$, $\Delta\mu_0 = 2$ *esetén*

1. Feladat: Fűrészgép beállításának ellenőrzése. Egy faüzemben egy automata géppel deszkákat vágnak, melyek hossza 2 méter szeretne lenni, de ez sajnos (legalább) két okból nem teljesül. Egyrészt a fűrészgép lötyög-zötyög, és emiatt a tényleges hossz véletlentől is függ. Feltesszük, hogy ez a valószínűségi változó normális eloszlást követ valamilyen várható értékkel és $\sigma = 4.5$ cm szórással. Másrészt a μ várható érték sem állandó, mert a használat során a fűrész beállítása elmozdul a kívánt $\mu_0 = 200$ értéktől. Reggel ugyan beállítják a várható értéket 200 cm -re, de néhány óra múlva μ értéke a kívánt 200 -tól már jelentősen eltérhet. Ezért időnként ellenőrizni kell, hogy a tényleges μ egyenlő-e a kívánt μ_0 értékkel. Az ellenőrzés céljából valahány – mondjuk 5 – frissen gyártott deszkát leveszenk a futószalagról és lemérnek. Veszik az átlaghosszzukat, és ha az 197 és 203 cm közé esik, akkor a beállítást jónak minősítik, ellenkező esetben leállítják a gépet (hogy - nem kis munkával – újra beállítsák és újraindítsák). Kérdés: hogyan függ a hipotézis elfogadásának s valószínűsége a tényleges μ értéktől?

Megoldás: Ha a beállítás jó, vagyis $\mu = \mu_0\ (= 200)$, akkor a helyes döntés valószínűsége:

$$p_0 = 2\,\Phi\left(\frac{\Delta\mu_0}{\frac{\sigma}{\sqrt{N}}}\right) - 1 = 2\,\Phi\left(\frac{3}{\frac{4.35}{\sqrt{10}}}\right) - 1 = 0.86$$

(A $\Delta\mu_0 = 3$ és $N = 5$ értékeket helyettesítettük be a képletbe.) Mint fentebb kiszámoltuk, tetszőleges μ esetén a hipotézis elfogadásának a valószínűsége

$$\Phi\left(\frac{(\mu_0 + \Delta\mu_0) - \mu}{\frac{\sigma}{\sqrt{N}}}\right) - \Phi\left(\frac{(\mu_0 - \Delta\mu_0) - \mu}{\frac{\sigma}{\sqrt{N}}}\right)$$

Ezért $\mu \neq \mu_0$ esetén ez a képlet adja meg a hibás döntés valószínűségét, amit 1 -ből kivonva kapjuk meg a helyes döntés valószínűségét. Néhány $\mu \neq \mu_0$ értékre táblázatba foglaltuk a helyes döntés valószínűségét:

	μ	Elfogadás valószínűsége
	191	0.00
	192	0.01
	193	0.02
	194	0.07
	195	0.16
	196	0.31
	197	0.50
	198	0.68
	199	0.82
$\mu_0 \;=$	200	**0.86**
	201	0.82
	202	0.68
	203	0.50
	204	0.31
	205	0.16
	206	0.07
	207	0.02
	208	0.01
	209	0.00

Táblázat: *A hipotézis elfogadásának a valószínűsége* μ *függvényében annál a döntési eljárásnál, amikor* $\Delta\mu_0 = 3$ *és* $N = 5$

Tessék meggondolni, hogy az lenne a jó, ha a táblázatban a vastagított **0.86** érték közelebb lenne az 1 -hez, a többi érték pedig a 0 -hoz! Ilyen lesz a gép következő beállítása.

2. Feladat: Fűrészgép beállításának másfajta ellenőrzése. Ha az ellenőrzés céljából nem 5, hanem 20 deszkát vizsgálnak, és annak alapján döntenek, hogy az átlaghosszuk az előzőnél rövidebb $(198; 202)$ intervallumba esik, akkor a $\mu_0 = 200$ esetén a helyes döntés valószínűsége 0.86 -ról 0.95 -re emelkedi. Néhány $\mu \neq \mu_0$ értékre a helyes döntés valószínűségét az alábbi táblázat mutatja:

	μ	Elfogadás valószínűsége
	195	0.00
	196	0.02
	197	0.16
	198	0.50
	199	0.84
$\mu_0 \;=$	200	**0.95**
	201	0.84
	202	0.50
	203	0.16
	204	0.02
	205	0.00

Táblázat: *A hipotézis elfogadásának a valószínűsége* μ *függvényében annál a döntési eljárásnál, amikor* $\Delta\mu_0 = 2$ *és* $N = 20$

Érdemes a két táblázat megfelelő soraiban álló valószínűségeket együtt szemrevételezni, és láttottak jelentésén elgondolkodni.

3. Feladat: Ha jól működik, ne állítgassuk túl gyakran! Elrendeli a főnök: olyan vizsgálatot kell csinálni, hogy

- 10 deszka hosszát mérjék csak meg, mert nem lehet méricskéléssel pocsékolni az időt,
- és amikor jól működik a gép (mert általában azért jól működik!), akkor legalább 99% legyen a valószínűsége annak, hogy a vizsgálat elfogadja a beállítást, és ezért elkerüljük a gép fölösleges állítgatását!

Milyen kritikus értékek mellett teljesülnek ezek a kívánaságok (parancsok)?

Megoldás: A

$$\Delta\mu_0 \;=\; \frac{\sigma}{\sqrt{N}} \cdot \Phi^{-1}\left(\frac{1+p_0}{2}\right)$$

. egyenletből $\Delta\mu_0 \;=\; 3.7$ jön ki, így a kritikus értékek: 203.7 és 196.3 .

	μ	Elfogadás valószínűsége
	192	0.00
	193	0.01
	194	0.05
	195	0.18
	196	0.42
	197	0.69
	198	0.88
	199	0.97
$\mu_0 \;=$	200	**0.99**
	201	0.97
	202	0.88
	203	0.69
	204	0.42
	205	0.18
	206	0.05
	207	0.01
	208	0.00

Táblázat: *A hipotézis elfogadásának a valószínűsége μ függvényében annál a döntési eljárásnál, amikor $\Delta\mu_0 = 3.7$ és $N = 10$*

4. Feladat: Hány deszka hosszából számoljuk ki az átlagot? Elrendeli a főnök, hogy a kritikus határok ne legyenek a 200-tól annyira messze, inkább 198 cm és 202 cm legyenek. Ragaszkodik korábbi elvárásához is, hogy amikor jól működik a gép, akkor legalább 99% legyen a valószínűsége annak, hogy a vizsgálat elfogadja a beállítást. Kérdés most, hogy hány deszka átlaghosszának kiszámolásával lehet ezt garantálni?

Megoldás: A

$$N \;=\; \left[\frac{\sigma}{\Delta\mu_0} \cdot \Phi^{-1}\left(\frac{1+p_0}{2}\right)\right]^2 \qquad \textit{felfelé kerekített értéke}$$

képletből $N = 34$ adódik. Vakarja is a fejét a főnök: 34 deszka méricskélése az bizony sok idő. De beletőrődik, hogy a $\sigma = 4.5$ cm szórás mellett a megadott $\Delta\mu_0 = 2$ cm pontosság és a $p_0 = 0.99$ biztonság elérésének – ha tetszik, ha nem – ez az ára.

	μ	Elfogadás valószínűsége
	196	0.00
	197	0.10
	198	0.50
	199	0.90
$\mu_0 \;=$	200	**0.99**
	201	0.90
	202	0.50
	203	0.10
	204	0.00

Táblázat: *A hipotézis elfogadásának a valószínűsége* μ *függvényében annál a döntési eljárásnál, amikor* $\Delta\mu_0 = 2$ *és* $N = 34$

7.2.2. Döntés az U-érték alapján

Világos, hogy a fentebb vizsgált

$$\left|\, \overline{X}_N - \mu_0 \,\right| \;<\; \Delta\mu_0$$

esemény ekvivalens az

$$\left| \frac{\overline{X}_N - \mu_0}{\frac{\sigma}{\sqrt{N}}} \right| < \frac{\Delta\mu_0}{\frac{\sigma}{\sqrt{N}}}$$

eseménnyel. Ezért ha bevezetjük az *u-érték*nek nevezett

$$U \;=\; \frac{\overline{X}_N - \mu_0}{\frac{\sigma}{\sqrt{N}}}$$

valószínűségi változót és a

$$c \;=\; \frac{\Delta\mu_0}{\frac{\sigma}{\sqrt{N}}}$$

értéket, akkor – egyrészt – a fentebbi

$$\left|\, \overline{X}_N - \mu_0 \,\right| \;<\; \Delta\mu_0$$

eseményt az alábbival helyettesíthetjük:

$$|\, U \,| \;<\; c$$

– másrészt – a fentebbi

$$p_0 \;= 2\,\Phi\left(\frac{\Delta\mu_0}{\frac{\sigma}{\sqrt{N}}} \right) \;-\; 1$$

egyenlet is egyszerűsödik:

$$p_0 \;= 2\,\Phi\,(c) \;-\; 1$$

Ezzel a hozzáállással a döntési eljárás például így tervezhető meg és hajtható végre: a

$$p_0 \;= 2\,\Phi\,(c) \;-\; 1$$

egyenletből adott p_0 -hoz a

$$c \;=\; \Phi^{-1}\left(\frac{1 + p_0}{2} \right)$$

összefüggés alapján meghatározzuk a c "kritikus" értéket, majd az U nagyságát összevetjük c -vel:

- ha $|\, U \,| \; < c$, akkor elfogadjuk a hipotézist,
- ha $|\, U \,| \; > c$, akkor elutasítjuk a hipotézist.

1. Megjegyzés: Mivel az

$$U \;=\; \frac{\overline{X}_N - \mu_0}{\frac{\sigma}{\sqrt{N}}}$$

valószínűségi változót általában U betűvel szokták jelölni, a döntési eljárás neve: *u-próba*. Kivétel az USA, ahol a Z betűt és a *z-próba* elnevezést használják.

2. Megjegyzés: Jó tiszában lenni az alábbi két ténnyel:

- A hipotézis teljesülése, azaz $\mu = \mu_0$ esetén az U érték nem más, mint az $\overline{X}_N$ érték standardizáltja, és az U eloszlása standard normális eloszlás.
- Tetszőleges μ esetén az U eloszlása normális eloszlás

  $$\frac{\mu - \mu_0}{\frac{\sigma}{\sqrt{N}}} \quad \text{várható értékkel és} \quad 1 \quad \text{szórással.}$$

- Ha μ eltér $\mu = \mu_0$ -tól, vagy N értéke elég nagy, akkor U eloszlása jelentősen eltér a standard normális eloszlástól. Igazából ez a tény biztosítja, hogy a próba rendesen működjön.

7.3. T-próba

JÖN MAJD IDE.

De aki a t-próbával kapcsolatan az előadáson elhangzottakon túl kíváncsi még másra is, egyrészt hamar megöregszik, másrészt írja be

`https://www.wikipedia.org/` címen ekérhető angol nyelvű Wikipédiába, hogy

`T-test` és akkor eljut erre a címre:

`https://en.wikipedia.org/wiki/Student%27s_t-test` ahonnan kíváncsiságát kielégítheti.

7.4. Tapasztalati eloszlásfüggvény

Ha egy valószínűségi változóra végzünk mondjuk N kísérletet, akkor minden x valós szám esetén megnézhetjük, hogy a kísérleti eredmények alapján mi a relatív gyakorisága annak az eseménynek, hogy a véletlen X érték kisebb mint a szóbanforgó x. Ha a relatív gyakoriságot x függvéynében ábrázoljuk, akkor egy olyan monoton növekedő "lépcsős" függvényt kapunk, aminek grafikonja vízszintes szakaszokból áll, és a vízszintes szakaszok közötti ugrások nagysága $\frac{1}{N}$. Ezt a függvényt *tapasztalati eloszlásfüggvény*nek nevezzük.

Ha például a kísérleti eredmények:

$$0.73\,, \quad 24.50\,, \quad 10.46\,, \quad 9.81\,, \quad 22.78$$

amik sorba rendezve így festenek:

$$0.73\,, \quad 9.81\,, \quad 10.46\,, \quad 22.78\,, \quad 24.50$$

akkor a tapasztalati eloszlásfüggvény grafikonja úgy fest, ahogy a

`A tapasztalati eloszlásfüggvény 5 kísérletből`

felíratú ábrán láthatjuk. A rajz elkészítését megkönnyíti, ha az ugrásoknál függőleges szakaszokat húzunk, ahogy

`Az ugrásoknál függőleges szakaszok`

felíratú ábrán láthatjuk. Végül

`A tapasztalati eloszlásfüggvény közelíti az elméleti eloszlásfüggvényt`

felíratú ábrán pedig azt a fontos tényt szemléltetjük a $\lambda = 0.1$ paraméterű exponenciális eloszlás példáján, hogy 100 kísérlet esetén a tapasztalati eloszlásfüggvény milyen jól közelíti az elméleti eloszlásfüggvényt.

Az Olvassó megnyithatja a

`http://math.bme.hu/~vetier/2017_osz/A4_vill_2017_osz.html` honlapcímen a lap felső részében a mellékletek között található

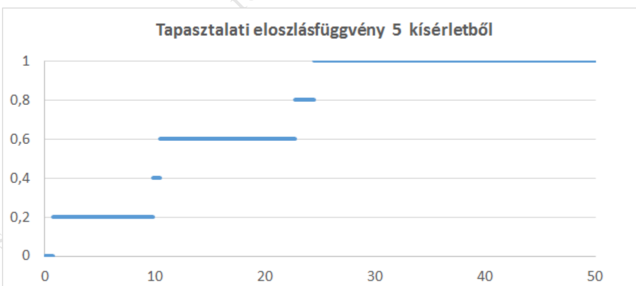
`Tapasztalati_eloszlás_függvény` nevű linkkel hívható

`Tap_el_fv.xlsx` nevű Excel fájlt,

és a számítógép `F9` gombjának ismételt nyomogatásával meggyőződhet, hogy a 100 kísérletre épített tapasztalati eloszlásfüggvény mennyire jól közelíti az elméleti eloszlásfüggvényt.

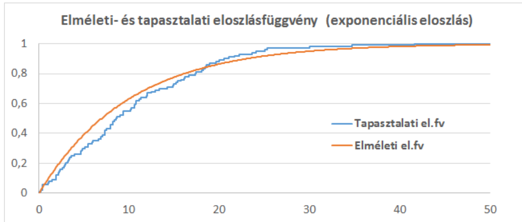
Igaz a következő fontos állítás:

Matematikai statisztika alaptétele: Ha a kísérletek száma tart a végtelenhez, akkor tetszőleges eloszlás esetén a tapasztalati elsozlásfüggvények sorozata 1 valószínűséggel egyenletesen konvergál az elméleti eloszlásfüggvényhez.



6. ábra. *A tapasztalati eloszlásfüggvény 5 kísérletből*

7. ábra. *Az ugrásoknál függőleges szakaszok*



8. ábra. *A tapasztalati eloszlásfüggvény közelíti az elméleti eloszlásfüggvényt*

8. Folytonos, de nem abszolút folytonos eloszlás a síkon *(Extra tananyag)*

Feladat: Három izzó az áramkörben, de csak kettő érdekel. Tegyük fel, két párhuzamosan kötött – mondjuk – piros és kék izzó mellett, velük sorba kötve, van még egy harmadik, ami – mondjuk – fehér. Tegyük fel, hogy a három élettartam egymástól független, és mindegyik élettartam exponenciális eloszlást követ. Azt is feltesszük, hogy a három izzó egyforma típusú, és ezért várható értékeik egyenlőek – mondjuk – 1000 óra várható értékkel.

Az élettartamok függetlensége és azonos eloszlása miatt

- annak a valószínűsége, hogy a fehér lámpa ég ki elsőnek, $1/3$ -dal egyenlő,
- annak a valószínűsége, hogy az egyik színes lámpa ég ki elsőnek, $2/3$ -dal egyenlő.

Jelöljük X -szel azt az időtartamot, amig a piros lámpa világít, Y -nal azt, amig a kék lámpa világít, Z -vel azt, amig a fehér lámpa világít. A piros lámpa két okból fejezheti be fényes működését:

- az egyik az, hogy – puff neki! – kiég, és így nem világít tovább,
- a másik pedig az, hogy a fehér lámpa ég ki, és akkor a soros kötés miatt a piros lámpa se világít tovább.

Ezért nyilván igaz, hogy

$$X = \text{a piros lámpa élettartama és a fehér lámpa élettartama közül a kisebbik}$$

Hasonló okok miatt igaz, hogy

$$Y = \text{a kék lámpa élettartama és a fehér lámpa élettartama közül a kisebbik}$$

Nyilvánvaló, hogy

- ha a piros lámpa ég ki elsőnek, akkor $X < Y$,
- ha a kék lámpa ég ki elsőnek, akkor $X > Y$,
- ha a fehér lámpa ég ki elsőnek, akkor $X = Y$.

Mivel mindegyik lámpa $1/3$ valószínűséggel ég ki elsőnek, az

$$X < Y \qquad\qquad X > Y \qquad\qquad X = Y$$

események mindegyikének $1/3$ a valószínűsége.

Fókuszáljunk az (X, Y) kétdimenziós valószínűségi változóra, vagyis figyeljük meg mennyi ideig világít a piros és mennyi ideig világít a kék lámpa! A fehér lámpával nem törödünk. Nyilván igaz, hogy (X, Y) minden (x, y) értéket 0 valószínűséggel vesz, ezért (X, Y) folytonos valószínűségi változó.

Viszont ennek az (X, Y) valószínűségi változónak – mindjárt látjuk – a korábban folytonosnak nevezett valószínűségi változókkal szemben van valami furcsasága. A korábbi folytonos valószínűségi változók esetében bármely halmaz valószínűsége a kétdimenziós sűrűségfüggvényből kettősintegrállal volt kiszámítható. Egy nulla területű halmazon vett kettős integrál értéke természetesen nulla, ezért a korábbi folytonos valószínűségi változók esetében bármely olyan esemény valószínűsége, mely nulla területű halmazhoz köröődött, nulla volt. Most viszont az az esemény, hogy (X, Y) a nulla területű $x = y$ egyenletű egyenesre esik, nem nulla: mint fentebb leszögeztük, $1/3$ -dal egyenlő!

És akkor most töredelmesen be kell ismernünk, hogy azokat a síkbeli folytonos eloszlásokat, amilyenekkel eddig foglalkoztunk, és csak egyszerűen folytonosnak neveztünk, igazából *a síkon abszolút folytonos* vagy *síkbeli sűrűség-függvénnyel difiniált* eloszlásoknak is szokás hívni. Tehát – ezzel a terminológiával élve azt kell mondanunk, hogy – eddig mi csak abszolút folytonos síkbeli eloszlásokkal foglalkoztunk, de a rövidség kedvéért az "abszolút" jelzőt elhagytuk. Eme pongyolaság utólagos magyarázatául még az alábbi érveket hozzuk fel:

- a "folytonos" jelzőt a "diszkrét" ellenpólusaként használtuk,
- amiket folytonosnak neveztünk, azok tényleg folytonosak voltak, sőt mi több,
- abszolút folytonosak is, de a kezdő diákot nem akartuk az "abszolút" szó – talán akkor még nehezebben emészthető – jelentésével terhelni.

Ezzel a most bevezetett terminológiával élve azt mondhatjuk, hogy az előbb – lámpák segítségével – definiált (X, Y) kétdimenziós valószínűségi változó eloszlása a síkon – bár folytonos –, de nem nem abszolút folytonos,