

2. A valószínűségszámítás axióma rendszere. Relatív gyakoriságok és a valószínűség.

Valószínűségszámítás és statisztika számítógép algebrai támogatással	Készült a PTE PMMIK-n a TÁMOP -4.1.2.A/1-11/1-2011-0098 támogatásával
Klincsik Mihály	

2.1. Bevezetés

BEVEZETÉS

Egy véletlen kísérlet eseményeit mérni a **bekövetkezők gyakoriságaival** tudjuk. Ehhez "nagyszámú" kísérletet kell **elvégezni**. A probléma az, hogy a kísérlet megismétlése során ugyanarra az eseményre mindig **más és más gyakoriság értékek** adódnak. Pontosan azért, mert a jelenség véletlenszerűen változtatja kimeneteleit. Vajon ebben az állandóan változó gyakoriság értékek világában van-e valami **állandóság**? Ez az állandóság a **valószínűség**, amely közelíthető a relatív gyakoriságokkal. A relatív gyakoriság értékét úgy kapjuk, hogy a gyakoriságot elosztjuk az addig elvégzett kísérletek számával. Gyakran előforduló eset az, hogy az események **relatív gyakoriságai ingadoznak** egy jól meghatározott szám körül. Továbbá, ha a megfigyelések számát növeljük, akkor a relatív gyakoriságok **egyre közelebb kerülnek** ehhez a számhoz. Ilyen esetben ezt a számot értjük az **esemény (empirikus) valószínűsége** alatt. Tehát kísérleti úton **közelíteni** tudjuk az esemény valószínűségét. Az **elméleti úton** való számításhoz célszerű **axiómatikus alapokra** helyezni a valószínűségelméletet ugyanúgy, mint azt a geometria esetében már több mint 2000 éve megtették. A valószínűségelmélet axiómái rögzítik azokat az egyszerű tényeket, amelyeket a relatív gyakoriságok **közzvetítenek** számunkra. Nevezetesen a valószínűség mindig egy $[0, 1]$ intervallumba eső szám, a biztos esemény valószínűsége 1 és lehetetlen esemény valószínűsége 0. Az egymást kizáró események összegének valószínűségét a tagok valószínűségeinek összegével számoljuk. Az így kapott $P(A)$ valószínűség függvény ugyanúgy **méri** egy véletlen A esemény valószínűségét, mint ahogyan az $m(T)$ tömegfüggvény méri egy T test fizikai tömegét a Földön. A hasonlítás a tömeg méréshez több, mint analógia, mert itt is a T_1 és T_2 testek egyesítésének tömege is a tömegek összege:

$$m(T_1 \cup T_2) = m(T_1) + m(T_2).$$

Az (Ω, S, P) hármast **valószínűségi mezőnek** nevezzük, amelyben Ω az eseménytér, S az események σ -algebrája és $P : S \rightarrow [0, 1]$ egy valószínűségi mérték az esemény algebrán. Az axiómák alapján szabályokat bizonyítunk a tagadás, az összeg és a különbség műveletekkel kapott események valószínűségének kiszámítására. Megmutatjuk, hogy a valószínűség függvény monoton.

Példákat mutatunk diszkrét és folytonos valószínűségi mezők konstrukciójára. Ezen konstrukciók között szerepel a **Poisson-eloszlás, geometriai eloszlás, egyenletes** és az **exponenciális** eloszlás modellje is. A σ -algebra fogalmának bevezetését indokoljuk a folytonos eseménytér bonyolult halmazainak konstrukciójával. Az eseményekkel végzett műveletek úgy, mint a tagadás, az összeg, a szorzat és a különbség valamint ezek valószínűségei összefüggenek azonosságokon keresztül. Ezeket az összefüggéseket kell alkalmazni, ha adott események valószínűségei alapján ki szeretnénk számolni újabb események valószínűségeit. Ezekre mutatunk példákat!

Célunk az alábbi fogalmak, összefüggések és eljárások megismerése, megértése

A véletlen események gyakoriságának,
relatív gyakoriságának mérése,

Egyenletes eloszlás modellje folytonos
eseménytérben, sűrűségfüggvénye és

megjelenítése	eloszlásfüggvénye
A relatív gyakoriságok és a valószínűség kapcsolata	Poisson-eloszlás valószínűségi mezője
A valószínűségelmélet Kolmogorov-féle axiómái	Geometriai-eloszlás valószínűségi mezője
Valószínűségi mező (Ω, S, P) fogalma. A valószínűség σ -additivitása.	Exponenciális-eloszlás valószínűségi mezője, sűrűségfüggvénye és eloszlásfüggvénye
Elméleti valószínűség, empirikus valószínűség és szubjektív valószínűség megkülönböztetése	Kísérletek ismételt független végrehajtásának valószínűségi mezője: szorzat-tér fogalma
Klasszikus valószínűségi mező	Relatív gyakoriságok Maple szimulációja és grafikonja
Tétel a tagadás esemény valószínűségére $P(\bar{A}) = 1 - P(A)$	Adatok generálása Maple-ben Poisson, egyenletes és exponenciális eloszlással
Tétel az összeg esemény valószínűségére $P(A + B) = P(A) + P(B) - P(A \cdot B)$	Maple <i>Statistics</i> csomag <i>FrequencyTable, Distribution, Uniform, LineChart, ColumnGraph, Histogram, PDF, CDF</i> eljárásai. A <i>piecewise</i> eljárás.
Tétel a különbség esemény valószínűségére $P(A - B) = P(A) - P(A \cdot B)$	A Maple feltételes utasítása: if feltétel then utasítás₁ else utasítás₂ end if
Tétel a részesemény valószínűségére: ha $A \subseteq B$, akkor $P(A) \leq P(B)$	Maple animáció készítése a <i>plots</i> csomagban levő <i>display([képek], insequence = true)</i> eljárással

2.2. Események gyakorisága és relatív gyakorisága

Statisztikai megfigyelések azt mutatják, hogy ha egy véletlen tömegjelenségre (mint a pénzfeldobásra vagy a kockadobásra) nagyon sok kísérletet hajtunk végre, akkor az esemény előfordulásai bizonyos törvényszerűségnek tesznek eleget. Nevezetesen az esemény előfordulásának relatív gyakoriságai ingadoznak valamely 0 és 1 között levő konstans érték körül. Ha a kísérletek számát növelve az eltérések ettől a konstanstól egyre kisebbek lesznek, akkor ezt a számot nevezzük az esemény **tapasztalati** (vagy **empirikus**) **valószínűségének**.

2.2.1. DEFINÍCIÓ. Esemény gyakorisága

Legyen Ω egy véletlen kísérlet eseménytere és A az Ω részhalmaza. Az A eseményre végezzünk **n független kísérletet**. Az **A esemény gyakoriságának** nevezzük azt a $k_n(A)$ számot, ahányszor bekövetkezik az A esemény az n kísérletből. ♦

A $k_n(A) = 0$ eset azt jelent, hogy az A esemény egyszer sem következett be. A másik véglet a $k_n(A) = n$ eset, amikor minden alkalommal az A esemény következett be. Egyébként a $k_n(A)$ gyakoriság értékre mindig teljesül a

$$0 \leq k_n(A) \leq n$$

egyenlőtlenség.

A gyakoriságok számítását tovább **formalizálhatjuk** az alábbiak szerint. Legyen

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_m\}, |\Omega| = m$$

a kísérlet véges sok lehetséges kimenetele és az n kísérlet során bekövetkezett kimenetek vektora

$$X = (x_1, x_2, x_3, \dots, x_n)$$

ahol minden $x_i \in \Omega = \{\omega_1, \omega_2, \dots, \omega_m\}$.

Ha A egy **esemény**, akkor legyen

$$Y = (y_1, y_2, y_3, \dots, y_n),$$

ahol

$$y_i = \begin{cases} 1, & \text{ha } x_i \in A \\ 0, & \text{ha } x_i \notin A \end{cases}$$

az i -ik esemény **indikátor változója**. Az Y vektor értéke ott 1, ahányadikra az A esemény bekövetkezik és ott 0 ahol az A esemény nem következik be. Ezért az Y vektor elemeinek összege éppen az A esemény gyakoriság értéke az n kísérletből

$$k_n(A) = \sum_{i=1}^n y_i = |Y|$$

2.2.2. PÉLDA. Elemi események gyakorisága a pénzfeldobás kísérleténél

Írjunk **szimulációs programot** a pénzérme feldobására és az események gyakoriságainak számítására, valamint megjelenítésére, amelyben legyen a *fej* = 1 és az *írás* = 0.

Az eseménytér $\Omega = \{1, 0\}$ és ezért $m = 2$. Az egyszerűség kedvéért dobjunk fel a pénzérmét $n = 10$ -szer. Ekkor a fenti X vektor 10 véletlenszerűen generált 0 és 1 értéket tartalmaz. Számoljuk meg az $A = \{0\}$ elemi esemény gyakoriság értékeit és szemléltessük!

```
> restart :
> randomize( ) : randomize( ) :
> érme := rand(0..1)
érme := proc( )
    proc( ) option builtin = RandNumberInterface; end proc(6, 2, 1)
end proc
```

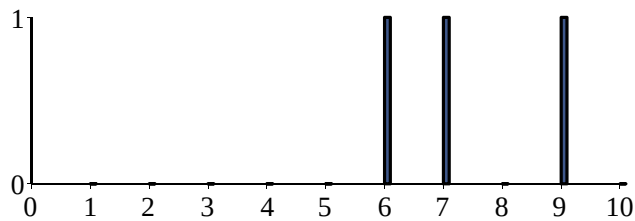
(1.2.1)

```
> dobásszám := 10
> dobások := [seq(érme( ), k = 1..dobásszám)]
dobásszám := 10
dobások := [0, 0, 0, 0, 0, 1, 1, 0, 1, 0]
```

(1.2.2)

A *Statistics* csomag *ColumnGraph* eljárására 1 magasságú oszlopot rajzol azokon a helyeken, ahol a *dobások* lista értéke 1.

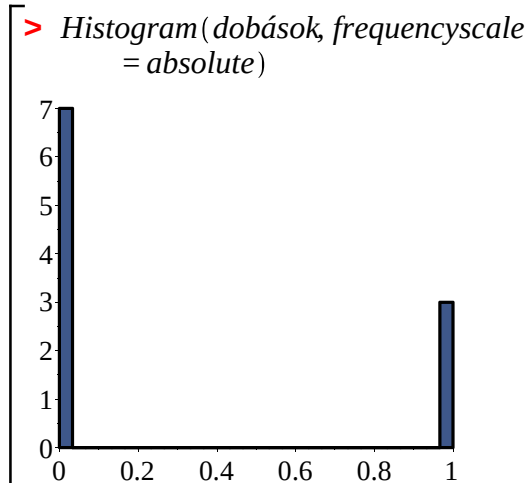
```
> with(Statistics) :
> ColumnGraph(dobások, width = 0.1, distance = 0.9, offset = 1,
    tickmarks = [dobásszám, 2])
```



A *FrequencyTable* eljárásnak a *bins = 2* opciójával azt közöltük, hogy a 0 és 1 értékközt ossza fel 2 egyenlő részre és számolja meg, hogy a $\left[0, \frac{1}{2}\right)$ és $\left[\frac{1}{2}, 1\right]$ intervallumokba hány eleme esik a *dobások* sorozatnak!

```
> evalf (FrequencyTable (dobások, bins = 2), 4);
[ 0..0.5000  7.  70.00  7.  70.00
  0.5000..1.  3.  30.00 10. 100.0 ] (1.2.3)
```

A *Histogram* eljárás olyan magasságú függőleges téglalapot rajzol az $x = 0$ és az $x = 1$ koordinátájú pontokban, ahány darab 0 illetve 1 szerepel a *dobások* listában.



Végül írtunk egy Maple programot, amely megvizsgálja, hogy vajon az i – ik dobás eleme-e a megadott eseménynek. Ha igen, akkor a *gyak* változót megnöveli 1-gyel és ennek értékét hozzáírja a *gyakoriságok* sorozat végéhez. A *LineChart* eljárás felrajzolja vonalas ábrával a *gyakoriságok*at.

```
> esemény := {0} : n := dobásszám :
   gyakoriságok := NULL : gyak := 0 :
   for k from 1 to n do
     if dobások[k] in esemény then gyak := gyak + 1 end if:
     gyakoriságok := gyakoriságok, gyak :
   end do:
   gyakoriságok := [gyakoriságok]
> LineChart(gyakoriságok, gridlines = true, legend = ["a 0 gyakoriságai"]);
   gyakoriságok := [1, 2, 3, 4, 5, 5, 5, 6, 6, 7]
```



2.2.3. DEFINÍCIÓ. Esemény relatív gyakorisága

Legyen Ω egy véletlen kísérlet eseménytere és A az Ω részhalmaza. Az A eseményre végezzünk n független kísérletet. Az A esemény relatív gyakoriságának nevezzük az

$$r_n(A) = \frac{k_n(A)}{n}$$

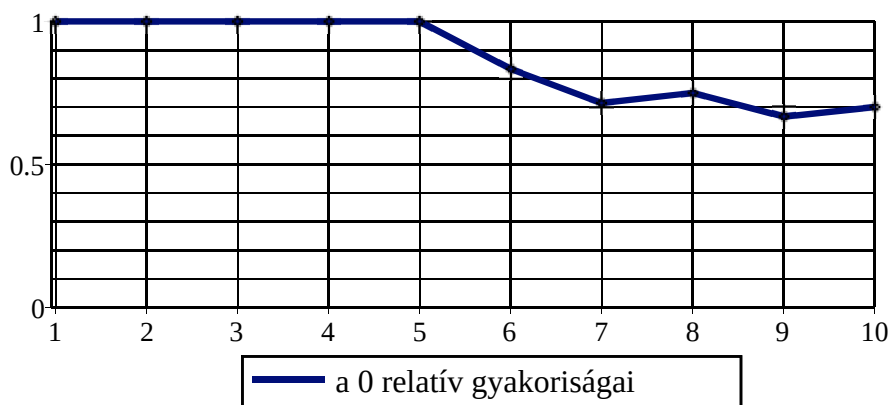
hányadost, ahol $k_n(A)$ az A esemény gyakorisága az n kísérletből. ♦

2.2.4. PÉLDA. Elemi események relatív gyakorisága a pénzfeldobás kísérleténél

Folytassuk a 2.2.2. pontban megkezdett példát és a kiszámolt gyakoriság értékek listájából készítsük el a relatív gyakoriságok listáját. Továbbá rajzoljuk fel relatív gyakoriságok vonalas diagramját!

```
> relativ_gyakorisagok := [seq( $\frac{\text{gyakorisagok}[k]}{k}$ , k = 1..n)];
LineChart(relativ_gyakorisagok, thickness = 3, axis = [gridlines = [10, color = black],
tickmarks = [dobasszam, 3], legend = ["a 0 relatív gyakoriságai"], view = [1
..dobasszam, 0..1])

relativ_gyakorisagok := [1, 1, 1, 1, 1,  $\frac{5}{6}$ ,  $\frac{5}{7}$ ,  $\frac{3}{4}$ ,  $\frac{2}{3}$ ,  $\frac{7}{10}$ ]
```



2.3. Események valószínűsége és a relatív gyakoriságok. Empirikus és szubjektív valószínűségek értelmezése.

Az események valószínűségét a relatív gyakoriságok tulajdonságainak elvonatkoztatásával alkotjuk meg. Ezért fontos megismerni a relatív gyakoriság tulajdonságait.

2.3.1. A relatív gyakoriság értékek alapvető tulajdonságai

Rávilátunk a relatív gyakoriságok néhány alapvető tulajdonságára, amelyek fontosak a valószínűségi fogalmak megalkotása során.

(i) A relatív gyakoriság értékek 0 és 1 között váltakoznak minden A eseményre

$$0 \leq r_n(A) \leq 1 \quad (\text{értékhatárok})$$

(ii) Az $A = \Omega$ eseménytér relatív gyakorisága mindig 1

$$r_n(\Omega) = 1. \quad (\text{egyre normált})$$

(iii) Ha az A és B események egymást kizárók, azaz $A \cdot B = \emptyset$, akkor

$$r_n(A + B) = r_n(A) + r_n(B) \quad (\text{additivitás})$$

az összeg esemény relatív gyakoriságát tagonként számolhatjuk.

Az (i) és (ii) tulajdonságok nem szorulnak magyarázatra. Az (iii) **additivitási** tulajdonság pedig következménye a gyakorisági értékek additivitásának, melyszerint

$$\text{ha } A \cap B = \emptyset, \text{ akkor } k_n(A + B) = k_n(A) + k_n(B).$$

2.3.2. PÉLDA. Szimuláció egymást kizáró események összegének relatív gyakoriságaira

Illusztráljuk a 2.3.1. pont (iii) additivitási tulajdonságát a kocka feldobás $A = \{1, 3\}$ és $B := \{2, 4, 5\}$ kizáró eseményeivel.

Megoldás

Nyilvánvaló, hogy az $A \cap B = \emptyset$ üres halmaz. Generáljunk 20 véletlen kocka dobást, majd gyűjtsük ki a relatív gyakoriságokat rendre az A , B és $C = A \cup B$ eseményekre. Mutassuk meg, hogy az

$$r_k(A + B) = r_k(A) + r_k(B)$$

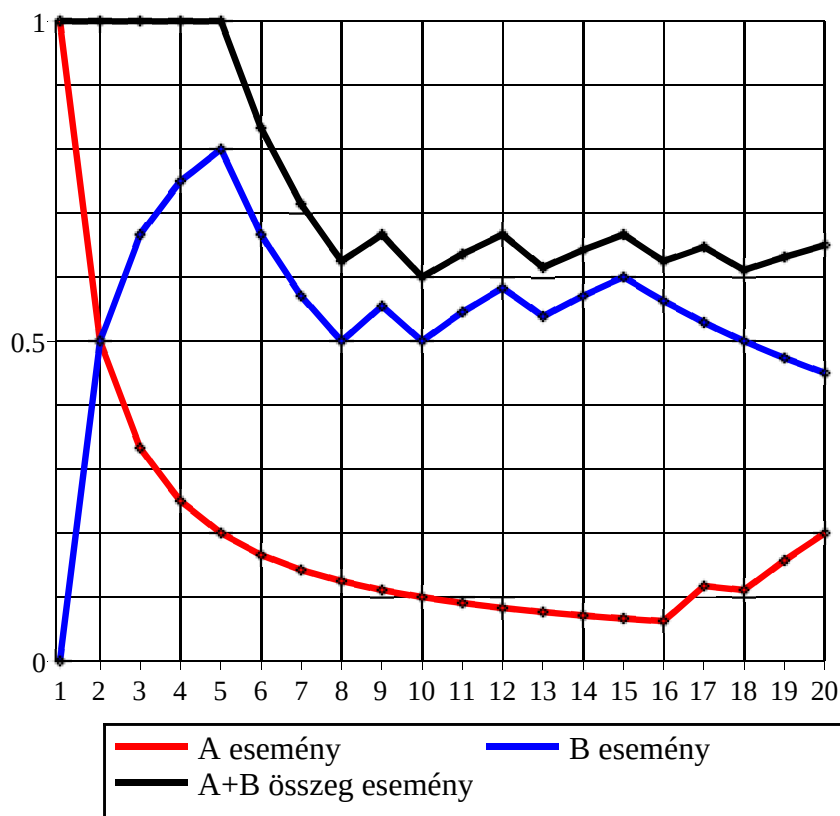
egyenlőség teljesül $k = 1, 2, 3, \dots, 20$ esetek mindegyikére!

```
> restart :
> randomize( ) :
>
```

```

> kocka := rand(1..6) :
> dobásszám := 20
> dobások := [seq(kocka( ), k = 1..dobásszám)]
                                dobásszám := 20
                                dobások := [3, 4, 2, 4, 4, 6, 6, 6, 2, 6, 5, 5, 6, 2, 2, 6, 1, 6, 3, 1]
                                (1.3.1)
> A := {1, 3}; B := {2, 4, 5}; C := A union B;
n := dobásszám:
relativ_gyakoriságok1 := NULL : gyak1 := 0 :
relativ_gyakoriságok2 := NULL : gyak2 := 0 :
relativ_gyakoriságok3 := NULL : gyak3 := 0 :
for k from 1 to n do
  if dobások[k] in A then gyak1 := gyak1 + 1 end if:
  if dobások[k] in B then gyak2 := gyak2 + 1 end if:
  if dobások[k] in C then gyak3 := gyak3 + 1 end if:
  relativ_gyakoriságok1 := relativ_gyakoriságok1,  $\frac{gyak1}{k}$  :
  relativ_gyakoriságok2 := relativ_gyakoriságok2,  $\frac{gyak2}{k}$  :
  relativ_gyakoriságok3 := relativ_gyakoriságok3,  $\frac{gyak3}{k}$  :
end do:
relativ_gyakoriságok1 := [relativ_gyakoriságok1]:
relativ_gyakoriságok2 := [relativ_gyakoriságok2]:
relativ_gyakoriságok3 := [relativ_gyakoriságok3]:
> Statistics[LineChart]([relativ_gyakoriságok1, relativ_gyakoriságok2,
relativ_gyakoriságok3], color = [red, blue, black], thickness = 3, axis = [gridlines
= [10, color = black]], tickmarks = [dobásszám, 3], legend = ["A esemény",
"B esemény", "A+B összeg esemény"], view = [1..dobásszám, 0..1], caption
= Egymást kizáró események és összegük relatív gyakoriságai)
                                A := {1, 3}
                                B := {2, 4, 5}
                                C := {1, 2, 3, 4, 5}

```



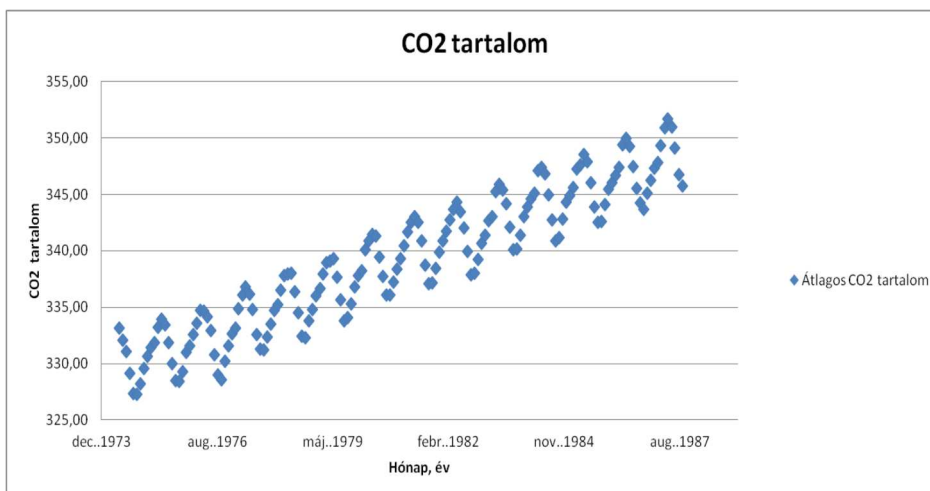
Egymást kizáró események és összegük relatív gyakoriságai

Az A , B és a $C = A + B$ események relatív gyakoriságai láthatók az ábrán. Irassuk most ki az $r_n(A) + r_n(B)$ események relatív gyakoriságainak összegét, majd az $r_n(A + B)$ összeg esemény relatív gyakoriságait is! Látható, hogy a két lista elemről elemre haladva egyezik.

$$\begin{aligned}
 &> \text{relativ_gyakorisagok1} + \text{relativ_gyakorisagok2}; \text{relativ_gyakorisagok3} \\
 &\left[1, 1, 1, 1, 1, \frac{5}{6}, \frac{5}{7}, \frac{5}{8}, \frac{2}{3}, \frac{3}{5}, \frac{7}{11}, \frac{2}{3}, \frac{8}{13}, \frac{9}{14}, \frac{2}{3}, \frac{5}{8}, \frac{11}{17}, \frac{11}{18}, \frac{12}{19}, \frac{13}{20} \right] \\
 &\left[1, 1, 1, 1, 1, \frac{5}{6}, \frac{5}{7}, \frac{5}{8}, \frac{2}{3}, \frac{3}{5}, \frac{7}{11}, \frac{2}{3}, \frac{8}{13}, \frac{9}{14}, \frac{2}{3}, \frac{5}{8}, \frac{11}{17}, \frac{11}{18}, \frac{12}{19}, \frac{13}{20} \right] \quad (1.3.2)
 \end{aligned}$$

2.3.3. A relatív gyakoriságok ingadozása és "határértéke"

Ha a véletlen kísérlet olyan, hogy megismételhető ugyanolyan körülmények között tetszőleges számszor, akkor a relatív gyakoriságok hosszú távú megfigyelésével **tendenciát** tudunk vizsgálni. A következtetés a tendenciára lehet az, hogy periódikusan változik és/vagy növekszik. Lásd például a 2.3.1. ábrán a levegő CO_2 koncentrációjának változását, amelyet mérőműszer automatikusan rögzített.



2.3.1 ábra. A "Mauna Loa" megfigyelőállomás által 1974 és 1987 között rögzített havi átlagos CO_2 koncentráció

Az alábbi 2.3.2. ábrán a New York-i tőzsdén az "Amazon" 2000-2003 közötti árfolyam változása látható a trend vonallal együtt. (http://stockcharts.com/school/doku.php?id=chart_school:chart_analysis:trend_lines). A megfigyelt adatok alapján nem lehet egy értéket prognosztizálni az árfolyamra megbízhatóan.

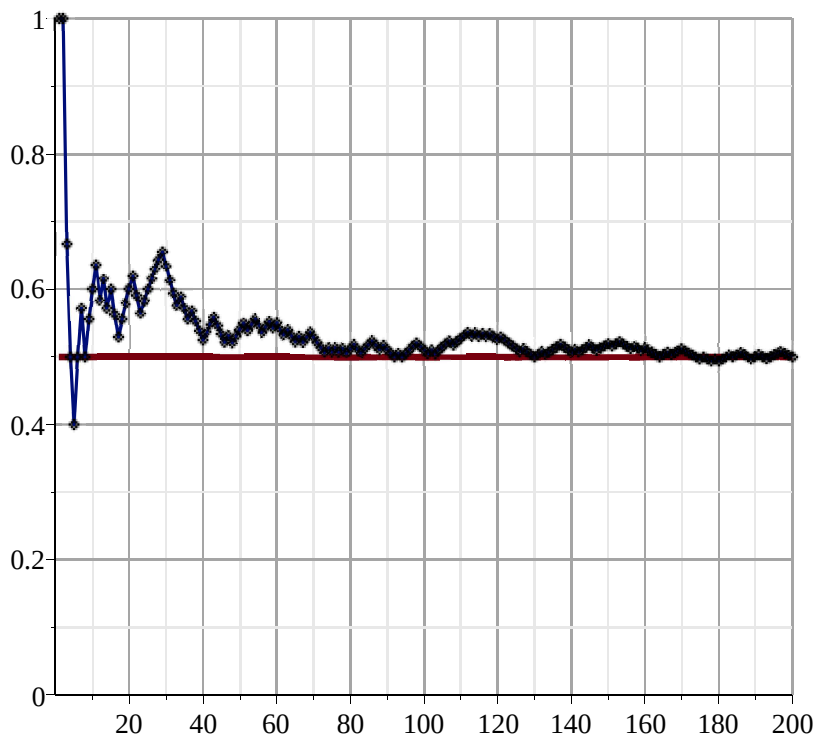


2.3.2. ábra. A tőzsdei árfolyam változása

Leszámítva a fentiekhez hasonló véletlen folyamatokat a relatív gyakoriságok ábrája egy konvergens sorozat határértékének ábrájára emlékeztet. A pénzfeldobás szimulációjára generáljunk $n = 200$ véletlen 0 és 1 értéket egyenletes eloszlásban. Majd rajzoljuk fel a relatív

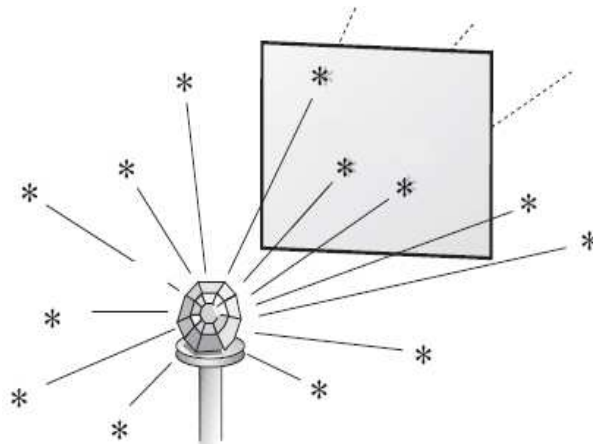
gyakoriságok vonalas ábráját a $p = \frac{1}{2}$ konstans valószínűséggel együtt!

```
> restart :
> randomize( ) : randomize( ) : n := 200 :
> érme := rand(0..1) : a 0 relatív gyakoriságai
> dobások := [seq(érme( ), k=1..n) ] :
esemény := {0} :
relativ_gyakoriságok := NULL : gyak := 0 :
for k from 1 to n do
  if dobások[k] in esemény then gyak := gyak + 1 end if:
  relativ_gyakoriságok := relativ_gyakoriságok,  $\frac{gyak}{k}$  :
end do:
relativ_gyakoriságok := [ relativ_gyakoriságok ] :
> ingadozás := Statistics[LineChart](relativ_gyakoriságok, gridlines = true, legend
= ["a 0 relatív gyakoriságai"], view = [1..n, 0..1]) :
plots[display]([plot( $\frac{1}{2}$ , 1..n, thickness = 3), ingadozás])
```



Ez egy példa a sok olyan véletlen kísérlet közül, amikor a relatív gyakoriságok **ingadozva egyre**

Az időintervallumok száma	57	20 3	383	525	532	408	273	139	45	27	10	6
---------------------------	----	---------	-----	-----	-----	-----	-----	-----	----	----	----	---



2.3.3. ábra. Rutherford és Geiger kísérlete az alfa részecskék megszámlálására

A táblázat alsó sora azon 1/8 perces intervallumok számát mutatja, ahányszor a felette levő részecskeszámot észlelték az adott 1/8 perces időintervallumban. Az, hogy milyen sorrendben jöttek ezek a mérési eredmények a kísérlet szempontjából nem lényegesek.

FELADAT.

Adjuk meg a

$P(a \text{ kisugárzó részecskék száma } 7.5 \text{ mp alatt} = n)$

empirikus valószínűségeket $n = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10$ és 11 -nél több értékekre a relatív gyakoriságok alapján.

Megoldás

Vigyünk be a két adatsort egy-egy listába!

```
> restart :
részecskeszám := [0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11];
gyakoriságok := [57, 203, 383, 525, 532, 408, 273, 139, 45, 27, 10, 6]
részecskeszám := [0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11]
gyakoriságok := [57, 203, 383, 525, 532, 408, 273, 139, 45, 27, 10, 6] (1.3.3)
```

Határozzuk meg a kategóriák számát!

```
> n := nops(gyakoriságok)
n := 12 (1.3.4)
```

Ha összeadjuk az összes időtartam gyakoriságát, akkor megkapjuk a mért 1/8 perces időintervallumok számát. Ezzel osztva a gyakoriságokat megkapjuk a relatív gyakoriság értékeket!

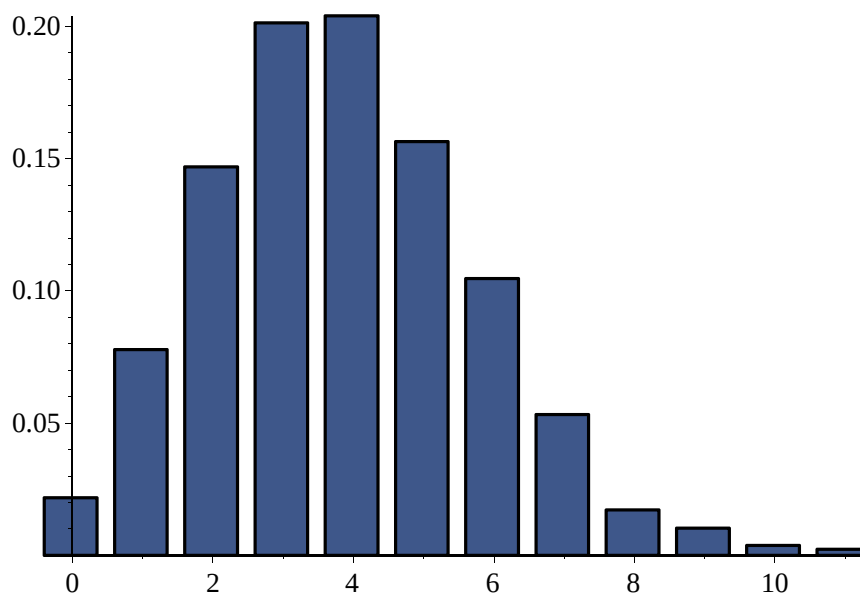
```
> összes_mérés :=  $\sum_{k=1}^n \text{gyakoriságok}_k$ 
összes_mérés := 2608 (1.3.5)
```

$$\begin{aligned}
 &> \text{relatív_gyakoriságok} := \frac{\text{gyakoriságok}}{\text{összes_mérés}} \\
 &\text{relatív_gyakoriságok} := \left[\frac{57}{2608}, \frac{203}{2608}, \frac{383}{2608}, \frac{525}{2608}, \frac{133}{652}, \frac{51}{326}, \frac{273}{2608}, \frac{139}{2608}, \right. \\
 &\quad \left. \frac{45}{2608}, \frac{27}{2608}, \frac{5}{1304}, \frac{3}{1304} \right] \quad (1.3.6)
 \end{aligned}$$

Azt kaptuk, hogy összesen 2068 darab 1/8 percet mértek, amely összesen $\frac{2068}{8}$ perc és ez kb. 4.3 óra.

Rajzoljuk fel relatív gyakoriságok oszlop diagramját!

> Statistics[ColumnGraph](relatív_gyakoriságok, offset = -0.4, caption = Szubjektív valószínűségek)



Szubjektív valószínűségek

Az oszlop diagramok azt mutatják, hogy az egyes $n = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11$ részecske kisugárzásának mennyi a relatív gyakorisága. Mutassuk meg, hogy a relatív gyakoriságok összege 1.

$$\begin{aligned}
 &> \sum_{k=1}^{12} \text{relatív_gyakoriságok}_k = \sum_{k=1}^n \text{relatív_gyakoriságok}_k \\
 &\quad \sum_{k=1}^{12} \text{relatív_gyakoriságok}_k = 1 \quad (1.3.7)
 \end{aligned}$$

Ezért alkalmazhatjuk a feladathoz alábbi (Ω, S, P) **valószínűségi mezőt**. Legyen az eseménytér az

$$\Omega = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11\}$$

a kísérletben az 1/8 perc időtartam alatt kibocsátott alfa-részek **lehetőséges** száma. Az elemi események **empirikus valószínűségeit** a mérésből kapjuk az előző tört értékek közelítésével. (Az elemi események jelölésénél elhagytuk a halmaz megadást szimbolizáló $\{ \}$ zárójeleket!)

$$P(0) = 0.021856 ; P(1) = 0.077837 ; P(2) = 0.14686 ; P(3) = 0.20130 ; P(4) = 0.20399 ; \\ P(5) = 0.15644 ; P(6) = 0.10468 ; P(7) = 0.053298 ; P(8) = 0.017255 ; P(9) = 0.010353 ; \\ P(10) = 0.0038344 ; P(11) = 0.0023006$$

További vizsgálat szükséges ahhoz, hogy az így megjelenő **empirikus eloszlást** milyen **elméleti eloszlás közelít** elfogadhatóan. (lásd 2.8.5. pontban tárgyalt Poisson-valószínűségi mezőt)

2.3.5. Szubjektív valószínűség és a statisztikai számítások

A **statisztikában** olyan adatokkal is kell dolgoznunk, amelyek függenek valamely vizsgálati csoporttól, a kérdéses időpontjától stb. Az ilyen adatok szubjektívek, személytől függőek.

2.3.5.1. PÉLDA

Kérdezzük meg diákjainkat, hogy **naponta átlagosan hány órát** szoktak **tanulni**. Ha a diákok attól tartanak, hogy a tanár visszaél az adatokkal, akkor válaszaikat leírhatják egy papírlapra név nélkül és a papírokat összekeverhetik. Így a tanár nem tudja azonosítani, hogy ki mennyi órát tanul naponta. Tegyük fel, hogy 20 diák adata a következő

5, 6, 3, 3, 2, 4, 7, 5, 2, 3, 5, 6, 5, 4, 4, 3, 5, 2, 5, 3.

Ha a kérdést feltesszük ugyanazoknak a diákoknak fél év múlva, akkor az nem feltétlenül egyezik a régi adatsorral.

```
> restart :
tanulás := [5, 6, 3, 3, 2, 4, 7, 5, 2, 3, 5, 6, 5, 4, 4, 3, 5, 2, 5, 3];
tanulás := [5, 6, 3, 3, 2, 4, 7, 5, 2, 3, 5, 6, 5, 4, 4, 3, 5, 2, 5, 3] (1.3.8)
```

Gyűjtsük ki a gyakoriságokat a *Statistics* csomag *Tally* nevű eljárásával!

```
> with(Statistics) :
gyakoriságok := sort(Tally(tanulás))
gyakoriságok := [2 = 3, 3 = 5, 4 = 3, 5 = 6, 6 = 2, 7 = 1] (1.3.9)
```

A gyakoriságokat olyan listában kaptuk vissza, amelyben egyenlőségek szerepelnek. Az egyenlőség bal oldalán a *tanulás* listában szereplő különböző óraszámok szerepelnek, a jobb oldalon ezek gyakoriságai. Válasszuk el a bal és jobb oldalakat külön listákba! Az *órák* nevű változóba tesszük az előforduló **tanulási óraszámokat** és a *diákokszáma* nevű változóba a **gyakoriságokat**.

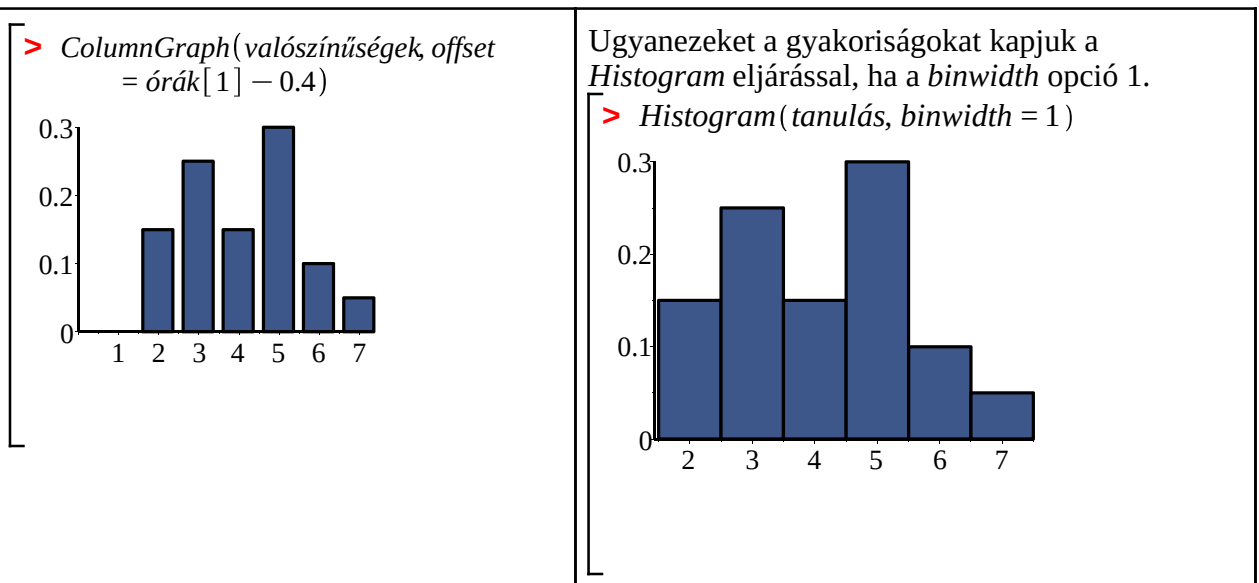
```
> órák := map(lhs, gyakoriságok);
diákokszáma := map(rhs, gyakoriságok);
órák := [2, 3, 4, 5, 6, 7]
diákokszáma := [3, 5, 3, 6, 2, 1] (1.3.10)
```

A relatív gyakoriságokat úgy kapjuk, ha az egyes kategóriákba tartozó diákszámot elosztjuk a teljes létszámmal, azaz 20-al. Így olyan valószínűségeket kapunk, amelyek összege 1.

```
>
```

$$\begin{aligned}
 & \text{> létszám} := \sum_{k=1}^{\text{nops}(\text{diákokszáma})} \text{diákokszáma}[k]; \\
 & \text{> valószínűségek} := \text{evalf}\left(\frac{\text{diákokszáma}}{\text{létszám}}\right); \\
 & \sum_{k=1}^{\text{nops}(\text{diákokszáma})} \text{'valószínűségek'}_k = \sum_{k=1}^{\text{nops}(\text{diákokszáma})} \text{valószínűségek}[k] \\
 & \text{létszám} := 20 \\
 & \text{valószínűségek} := [0.1500000000, 0.2500000000, 0.1500000000, 0.3000000000, \\
 & \quad 0.1000000000, 0.0500000000] \\
 & \sum_{k=1}^6 \text{valószínűségek}_k = 1.0000000000 \quad (1.3.11)
 \end{aligned}$$

A *ColumnGraph* eljárással olyan oszlop diagramot rajzolunk, amelyben az oszlopok magasságai a valószínűségek. Mivel nem 1-gyel kezdődik az óraszám, ezért eltoljuk a grafikont az x-tengelyen annyival, hogy az első oszlop 2 óránál kezdődjön.



A 2.3.4. pontban mondottaknak megfelelően megadhatunk egy (Ω, S, P) **valószínűségi mezőt** a feladathoz. Ebben az eseménytér

$$\Omega = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24\}$$

az egy nap alatt tanulással eltölthető órák száma. Az $S = 2^\Omega$ az eseménytér összes részhalmazainak halmaza. A példában az elemi események valószínűsége $P(2) = 0.15, P(3) = 0.25, P(4) = 0.15, P(5) = 0.3, P(6) = 0.1, P(7) = 0.05$ és $P(k) = 0$ egyébként.

Miben különbözik a fenti feladat és a 2.3.4. pontban vizsgált részecske bomlástól? Mivel a részecske bomlás folyamatát fizikai törvények írják le, ezért az adatokhoz feltett kérdésekre a magyarázatot ezekkel kell megválaszolni. A tanulási idő valószínűségeihez feltett kérdésekre a választ a diákok más elfoglaltságai, érdeklődési köre, előképzettsége, stb. adhatják meg. Ezek a területek az egyéntől függenek. Ezért a kapott valószínűség a **szubjektív** kategóriába tartozik és nem tévesztendő össze az **empirikus valószínűséggel**!

Megjegyzés. A *Statistics* csomag *FrequencyTable* eljárásával is kigyűjthettük volna a gyakoriság,

relatív gyakoriság, sőt az összegzett értékeket is. Arra kell azonban ügyelni, hogy az egyes intervallumokba esés gyakoriságainál $[a, b)$ balról zárt és jobbról nyitott intervallummal számol, kivéve az utolsó intervallumot, ahol az intervallum $[a, b]$ jobbról is zárt! Ezért a megfelelő gyakoriságokat úgy kaptuk volna meg, ha a *range* opcióban az adatok értékkészletéhez hozzáadunk 1-et és a *bins* opciónál is ugyanezt tesszük. Próbálja ki az olvasó az utasítást a +1 nélkül is!

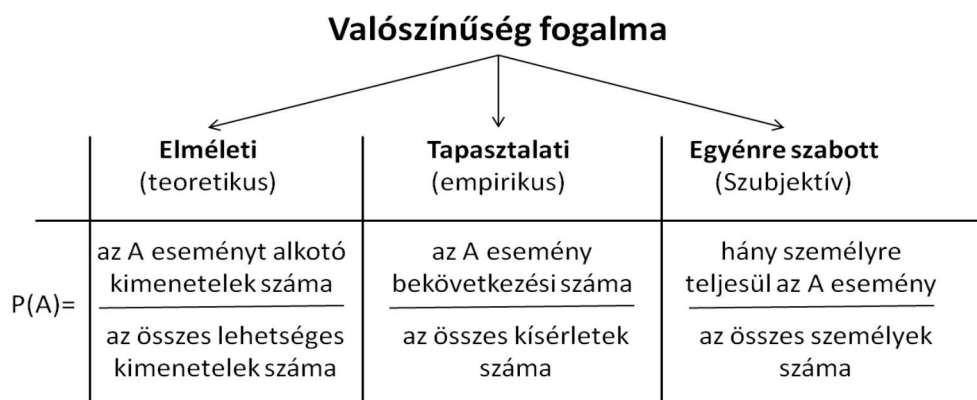
```
> FrequencyTable(tanulás, range = min(tanulás) ..max(tanulás) + 1, bins
= round(Range(tanulás)) + 1)
```

```

2...3. 3. 15.00000000 3. 15.00000000
3...4. 5. 25.00000000 8. 40.00000000
4...5. 3. 15.00000000 11. 55.00000000
5...6. 6. 30.00000000 17. 85.00000000
6...7. 2. 10.00000000 19. 95.00000000
7...8. 1. 5.00000000 20. 100.00000000
```

(1.3.12)

Miután áttekintettük példákon keresztül a valószínűség fogalmának lehetséges értelmezéseit, ezután megpróbáljuk egyértelműen megadni a közöttük levő különbséget és hasonlóságot. Ezt az alábbi grafikon szemlélteti.



A $P(A)$ valószínűségek mindhárom esetben hányadosként adtuk meg.

Az **elméleti valószínűséget** egy matematikai (vagy **sztochasztikus**) **modell** alapján számoljuk és ez klasszikus esetben az A eseményt alkotó kedvező kimenetek **számának** aránya az összes lehetséges kimenetek számához viszonyítva. Végtelen sok kimenetelre ennek nincs értelme. Folytonos esetben a kimenetek száma alatt a megfelelő halmaz (számsíki intervalluma vagy sík tartománya vagy térbeli tartomány) **mértékét** (hosszát vagy területét vagy térfogatát) kell érteni!

A **tapasztalati valószínűségnél** egy **valós folyamathoz** megtervezett **kísérletet** ismételnünk többször egymástól függetlenül. Figyeljük az A esemény bekövetkezési számát az n kísérletből és ezek aránya a valószínűség az összes kísérletszámhoz viszonyítva.

A **szubjektív valószínűségnél** az emberekhez kapcsolható A esemény bekövetkezését vizsgáljuk egy populációban. Az A eseményt teljesítő személyek számát elosztva a populáció n számával megkapjuk a valószínűséget.

Látható, hogy éles határt nem tudunk húzni a három fogalom között.

2.4. Kolmogorov-axiómák. Valószínűségi mező. Klasszikus valószínűségi mező.

A valószínűségelmélet ma is használatos axiómáit **A.N. Kolmogorov** (1903–1987) orosz matematikus fektette le egy 1933-ban megjelent könyvében. Az axiómák 3 részre oszthatók. Első rész az Ω eseménytérre vonatkozik. A második rész az események olyan S rendszerét írja elő, amelyen értelmezhető a valószínűség. Itt a σ -algebra fogalmát kell használni a folytonos eseménytér kezelhetősége miatt. (lásd 2.7. részt). A harmadik rész a $P : S \rightarrow [0, 1]$ valószínűség függvényt értelmezi az eseményeken úgy, hogy az a relatív gyakoriságok jól ismert tulajdonságainak teljesülését írja elő feltételként.

A.1. AXIÓMA. Feltételek az eseménytérre

Az Ω nem üres halmaz neve **eseménytér**, amely a véletlen **kísérlet** összes lehetséges **kimenetelét** tartalmazza. Az eseménytér elemei a kísérlet lehetséges kimenetelei.

A.2. AXIÓMA. Feltételek az eseményalgebrára

Legyen S az Ω eseménytér részhalmazából álló halmazrendszer, melynek elemeit **eseményeknek** nevezzük. Az S teljesítse a **σ -algebra** alábbi 3 feltételét

(A.2.1) $\emptyset \in S$

(A.2.2) Ha $A \in S$, akkor $A^c = \Omega \setminus A \in S$.

(A.2.3) Ha az $A_1, A_2, \dots, A_n, \dots$ (végtelen sok) esemény mindegyike eleme az S – halmazrendszernek, akkor legyen az uniójuk is S -ben, azaz $\left(\bigcup_{n=1}^{\infty} A_n \right) \in S$ teljesüljön.

A.3. AXIÓMA. Feltételek a valószínűségre

Az S σ -algebra minden $A \in S$ eseményéhez hozzárendelünk egy $P(A)$ **valószínűséget**, amely eleget tesz a következő feltételeknek

(A.3.1) $0 \leq P(A) \leq 1$

(A.3.2) $P(\Omega) = 1$

(A.3.3) Ha az $A_1, A_2, A_3, \dots$ események egymást páronként kizárók $A_k \cdot A_i = \emptyset$ ($k \neq i$), akkor az események összegének valószínűségét a tagok valószínűségeinek összegével számolhatjuk, azaz

$$P\left(\sum_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n) . \square$$

Az (A.3.3) feltétel teljesülése esetén azt mondjuk, hogy a $P : S \rightarrow [0, 1]$ valószínűségi függvény **σ -additív**.

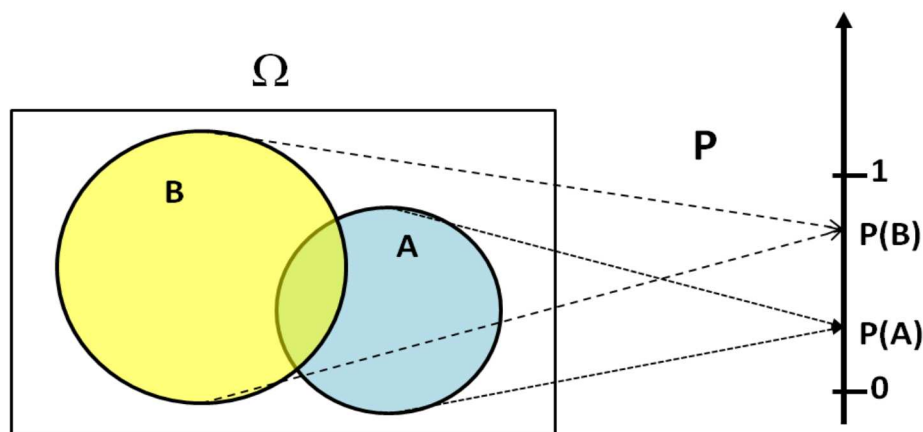
A valószínűségre kiemelt σ (szigma) jelző arra utal, hogy az additivitás nemcsak véges sok tagra érvényes, hanem megszámlálhatóan végtelen sok tagra is!

2.4.1. DEFINÍCIÓ. VALÓSZÍNŰSÉGI MEZŐ

Az (Ω, S, P) hármast **valószínűségi mezőnek** nevezzük, amelyben

- Ω az eseménytér,
- S az események σ -algebrája
- $P : S \rightarrow [0, 1]$ a σ -additív valószínűségi mérték.

A 2.4.1. ábra azt mutatja, hogyan képezi le a P valószínűség függvény az Ω eseménytér A és B részhalmazait a számegyenes $[0, 1]$ intervallumának egy-egy számértékére. Nagyobb mértékű halmazhoz nagyobb valószínűségi érték tartozik.



2.4.1. ábra. A P valószínűség függvény Ω részhalmazait leképezi a $[0, 1]$ intervallum értékeire.

2.4.2. A klasszikus valószínűségi mező

Legyen a véges Ω eseménytér $n = |\Omega|$ elemű és az elemei a következők

$$\Omega = \{\omega_1, \omega_2, \omega_3, \dots, \omega_n\}.$$

Jelölje

$$A_1 = \{\omega_1\}, A_2 = \{\omega_2\}, A_3 = \{\omega_3\}, \dots, A_n = \{\omega_n\}$$

az Ω különböző elemeihez tartozó **elemi eseményeket**! Mint láttuk, az $A_1, A_2, A_3, \dots, A_n$ események teljes eseményrendszert alkotnak.

DEFINÍCIÓ. KLASSZIKUS VALÓSZÍNŰSÉGI MEZŐ

Az (Ω, S, P) valószínűségi mezőt **klasszikus valószínűségi mezőnek** nevezzük, ha a véges Ω eseménytér **minden elemi eseményének valószínűsége ugyanakkora**, tehát

$$P(A_1) = P(A_2) = P(A_3) = \dots = P(A_n) = p. \blacklozenge$$

Mivel az elemi események egymást páronként kizárók és összegük az Ω eseménytér, ezért az (A.3.2) és (A.3.3) axiómák alapján

$$P(\Omega) = P(A_1 + A_2 + A_3 + \dots + A_n) = P(A_1) + P(A_2) + P(A_3) + \dots + P(A_n) = n \cdot p = 1.$$

Ahonnán

$$p = \frac{1}{n}$$

Ha az $A = \left\{ \omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_k} \right\}$ halmaz az Ω eseménytér valamely k – elemű részhalmaza

$$(1 \leq k \leq n), \text{ akkor } A = \left\{ \omega_{i_1} \right\} + \left\{ \omega_{i_2} \right\} + \dots + \left\{ \omega_{i_k} \right\} = A_{i_1} + A_{i_2} + \dots + A_{i_k}.$$

A $P : S \rightarrow [0, 1]$ valószínűségi mérték additivitása miatt

$$P(A) = P(A_{i_1} + A_{i_2} + \dots + A_{i_k}) = P(A_{i_1}) + P(A_{i_2}) + \dots + P(A_{i_k}) = \frac{1}{n} + \frac{1}{n} + \dots + \frac{1}{n} = \frac{k}{n}.$$

Tehát egy klasszikus valószínűségi mező valamely k kimenetelt tartalmazó A eseménynek valószínűségét

$$P(A) = \frac{\text{kedvező esetek száma az } A \text{ eseményre}}{\text{összes esetek száma}} = \frac{|A|}{|\Omega|}$$

arányal kell számolni.

Ezért, ha **számítógépes szimulációval** véletlen számokat generálunk 1 és n között és egy szám

jelöli a neki megfelelő sorszámú kimenetelt, akkor az A esemény méretének megfelelő gyakorisággal esik a szám az A halmazba. Ha ugyanis az A halmaz kevés elemet tartalmaz, akkor a generált szám ritkán esik bele a halmazba. Ha az A esemény sok kimenetelt tartalmaz, akkor a generált szám is sűrűn esik az A halmazba.

Sok véletlen alapú **játék** egyik alapfeltétele az, hogy az **elemi események esélyei egyenlők** legyenek. Például

- kocka játéknál felteszik, hogy a kocka szabályos,
- kártya játéknál a lapok kiosztásának esélye egyforma,
- ruletnél a golyó ugyanolyan eséllyel állhat meg bármelyik rekeszben,
- a LOTTÓ-nál a 90 szám közül ugyanolyan esélye van bármelyik 5 szám kihúzásának.

Ezek mind a klasszikus valószínűségi mező elméleti modelljeit feltételezik a gyakorlatban.

2.5. A tagadás, az összeg és a különbség műveletek valószínűségi szabálya

A valószínűségelmélet alappillérei az axiómák. Az axiómák felhasználásával bizonyítunk számítási szabályokat a tagadás, az összeg és a különbség műveletekre. Majd összehasonlítjuk két egymást tartalmazó esemény valószínűségét.

2.5.1. TÉTEL. Tagadás esemény valószínűsége

Tetszőleges A esemény tagadásának valószínűségét a

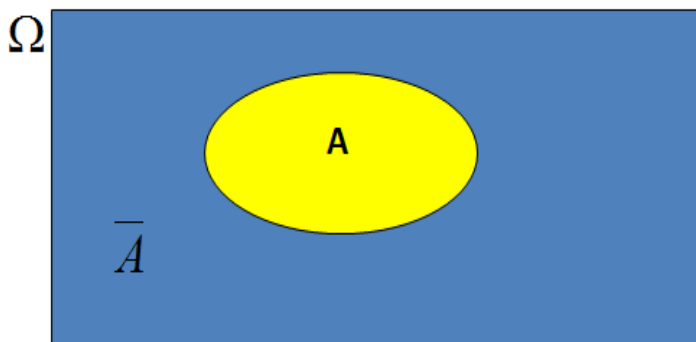
$$P(\bar{A}) = 1 - P(A)$$

képlettel számolhatjuk.

BIZONYÍTÁS

Az A esemény és az $A^c = \bar{A}$ tagadás esemény uniója egyrészt a teljes eseménytér, másrészt egymást kizárók (lásd az 1.6.2. táblázatot)

$$A \cup A^c = \Omega, A \cap A^c = \emptyset.$$



A valószínűség függvény additív tulajdonsága és az eseménytér valószínűsége miatt

$$P(A \cup A^c) = P(A) + P(A^c) = P(\Omega) = 1.$$

Az egyenletet átrendezve kapjuk a tétel állítását $P(A^c) = 1 - P(A)$. Tehát a tagadás esemény valószínűségét megkapjuk, ha az eredeti esemény valószínűségét kivonjuk egyből. ♦

2.5.2. Következmény

A lehetetlen esemény valószínűsége nulla, azaz $P(\emptyset) = 0$.

Bizonyítás.

Mivel a biztos esemény tagadása a lehetetlen esemény, ezért

$$P(\emptyset) = P(\bar{\Omega}) = 1 - P(\Omega) = 1 - 1 = 0. \quad \blacklozenge$$

Tehát a lehetetlen esemény nulla valószínűségű. Ez fordítva nem feltétlenül igaz.

2.5.3. TÉTEL. Összeg esemény valószínűsége

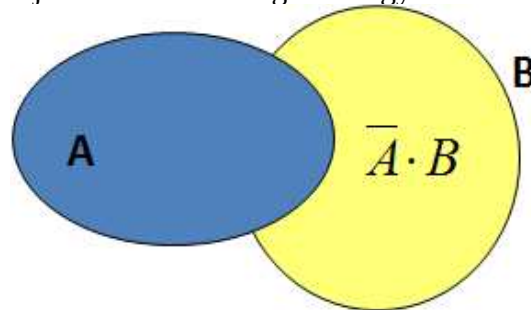
Tetszőleges A és B események összegének valószínűsége a

$$P(A + B) = P(A) + P(B) - P(A \cdot B)$$

képlettel számolható.

BIZONYÍTÁS.

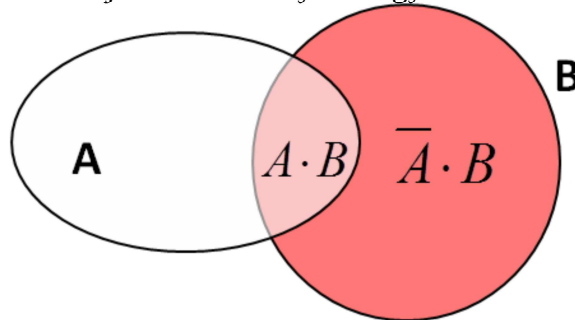
Első lépésként felbontjuk az $A + B$ összeget két egymást kizáró esemény összegére



A Venn-diagram alapján teljesül az $A + B = A + A^c \cdot B$ azonosság, melyben a jobb oldali két tag metszete üres. Ezért a valószínűség additív tulajdonsága miatt

$$(2.5.1) \quad P(A + B) = P(A) + P(A^c \cdot B).$$

Második lépésként felbontjuk a B eseményt két egymást kizáró esemény összegére.



A Venn-diagram és halmazalgebrai azonosságok alapján teljesül a

$$B = B \cdot \Omega = B \cdot (A + A^c) = B \cdot A + B \cdot A^c$$

azonosság, melyben a jobb oldali két tag metszete üres. Ezért a valószínűség additív tulajdonsága miatt

$$(2.5.2) \quad P(B) = P(A \cdot B) + P(A^c \cdot B).$$

A (2.5.2) egyenletből rendezéssel kapjuk a $P(A^c \cdot B) = P(B) - P(A \cdot B)$ előállítást, amelyet behelyettesítve a (2.5.1) azonosságba adódik a

$$P(A + B) = P(A) + P(A^c \cdot B) = P(A) + (P(B) - P(A \cdot B))$$

bizonyítandó egyenlőség. ♦

2.5.4. Események különbségeinek valószínűsége

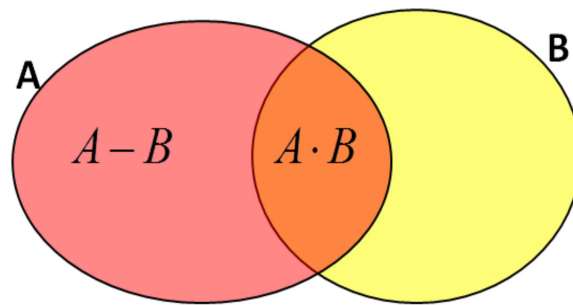
TÉTEL. Bármely A és B eseményekre $P(A - B) = P(A) - P(A \cdot B)$.

BIZONYÍTÁS

Az alábbi Venn-diagramról leolvasható az

$$A = (A - B) + A \cdot B$$

halmaz azonosság igaz tetszőleges A és B eseményre.



Mivel az $(A - B)$ és $A \cdot B$ események egymást kizárók, ezért A3.3. additivitási axióma alapján a valószínűséget a jobb oldalon tagonként vehetjük

$$P(A) = P(A - B) + P(A \cdot B).$$

Az egyenlet mindkét oldalából kivonva a $P(A \cdot B)$ valószínűséget kapjuk a keresett azonosságot.



2.5.5. Részesemény valószínűsége kisebb vagy egyenlő, mint az őt tartalmazó esemény valószínűsége

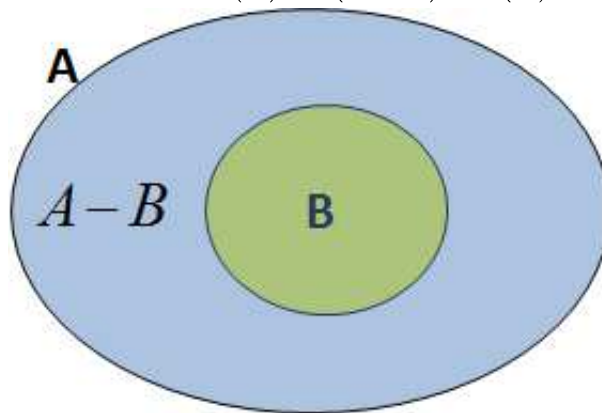
Az alábbi tétel a valószínűség függvény monotonitását mutatja.

TÉTEL. Ha $B \subseteq A$, akkor $P(B) \leq P(A)$.

BIZONYÍTÁS

Mivel $B \subseteq A$, ezért $A = (A - B) + B$, ahol a tagok egymást kizárók $(A - B) \cdot B = \emptyset$. Ezért az (A.3.3) axióma alapján

$$P(A) = P(A - B) + P(B).$$



Az (A.3.1) axióma alapján $P(A - B) \geq 0$, ahonnan kapjuk a

$$P(A) = P(A - B) + P(B) \geq P(B)$$

bizonyítandó egyenlőtlenséget. ◆

2.5.6. Összeg valószínűségének általánosítása három tagú összegre

Ha két tagú összeg helyett **három tagú összeg** valószínűségét szeretnénk kiszámolni, akkor az $A + B + C = (A + B) + C$ zárójelezéssel két tagú összeget kapunk és a két tagú összegre vonatkozó 2.5.3. tétel alapján

$$P(A + B + C) = P(A + B) + P(C) - P((A + B) \cdot C).$$

A jobb oldalon szereplő a $P(A + B)$ valószínűség helyébe a 2.5.3. tétel alapján $P(A) + P(B) - P(A \cdot B)$ képlet kerülhet. Az $(A + B) \cdot C = A \cdot C + B \cdot C$ disztributivitást és a 2.5.3. tételt használva újból

$$P(A \cdot C + B \cdot C) = P(A \cdot C) + P(B \cdot C) - P(A \cdot B \cdot C).$$

A képleteket egymásba helyettesítve és rendezve a tagokat

$$P(A + B + C) = P(A) + P(B) + P(C) - P(A \cdot B) - P(A \cdot C) - P(B \cdot C) + P(A \cdot B \cdot C).$$

Tehát pozitív előjellel kell venni az A , B és C események valószínűségeit, negatív előjellel kell venni az összes két tényezős $A \cdot B$, $A \cdot C$ és $B \cdot C$ szorzat események valószínűségeit és végül pozitív előjellel kell venni a három tényezős $A \cdot B \cdot C$ szorzat esemény valószínűségét. ♦

Több tagú összeg valószínűségére hasonló formula adható, amelyet szita-formulának vagy Poincare-formulának nevezünk.

2.5.7. TÉTEL. Poincare-formula vagy szita-formula vagy kizárás-beszámítás formulája

Az $A_1, A_2, \dots, A_n$ események összegének valószínűségét az alábbi formula adja meg

$$P(A_1 + A_2 + \dots + A_n) = \sum_{k=1}^n (-1)^{k-1} \cdot S_k$$

ahol

$$S_k = \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} P(A_{i_1} \cdot A_{i_2} \cdot \dots \cdot A_{i_k}).$$

Azt mondjuk, hogy váltakozó előjellel össze kell adni az összes k -tényezős szorzat esemény valószínűségét, ahol mindegyik k -tényezős szorzat valószínűségének előjele azonos. Ügyelni kell arra, hogy a halmazok különböző szorzatai közül mindegyik szorzat csak **egyszer forduljon elő**. Ezt úgy értük el, hogy el az összegzésben résztvevő indexek szigorúan növekszenek. Tehát, ha például az $A_1 \cdot A_2$ szorzat szerepel a két tényezős szorzatok között, akkor a fordított sorrendben írt $A_2 \cdot A_1$ szorzat nem szerepelhet.

Sok tagú $A_1 + A_2 + \dots + A_n$ összeg esetén a 2.5.7. tétel jobb oldalán szereplő összeg nagyon sok tagból áll. Ezért mindegyik tag valószínűségének kiszámítása hosszadalmas vagy akadályokba ütközik. Helyette sok esetben egyszerűbb becsléseket használni.

2.5.8. TÉTEL. Boole- és Bonferoni- egyenlőtlenségek

Tetszőleges $A_1, A_2, \dots, A_n$ ($n \geq 2$) események összegének és szorzatának valószínűségére teljesülnek a

$$(i) \quad P(A_1 + A_2 + \dots + A_n) \leq P(A_1) + P(A_2) + \dots + P(A_n) \quad (\text{Boole-egyenlőtlenség})$$

$$(ii) \quad P(A_1 \cdot A_2 \cdot \dots \cdot A_n) \geq P(A_1) + P(A_2) + \dots + P(A_n) - n + 1$$

(Bonferoni-egyenlőtlenség)
egyenlőtlenségek.

BIZONYÍTÁS

(i) Boole-egyenlőtlenség

Teljes indukcióval bizonyítunk. Ha $n = 2$, akkor a 2.5.3. tétel alapján

$$P(A_1 + A_2) = P(A_1) + P(A_2) - P(A_1 \cdot A_2) \leq P(A_1) + P(A_2)$$

teljesül, mert $P(A_1 \cdot A_2) \geq 0$.

Tegyük fel, hogy a Boole-egyenlőtlenség igaz $n \leq k$ mellett. Mutassuk meg, hogy teljesül $n = k + 1$ esetén is. Az összeg asszociatív tulajdonsága alapján

$$P((A_1 + A_2 + \dots + A_k) + A_{k+1}) \leq P(A_1 + A_2 + \dots + A_k) + P(A_{k+1}),$$

ahol felhasználtuk, hogy $n = 2$ mellett az egyenlőtlenség igaz. Vegyük most figyelembe, hogy $n = k$ esetén az egyenlőtlenség a feltételezés miatt fennáll, azaz

$$P(A_1 + A_2 + \dots + A_k) \leq P(A_1) + P(A_2) + \dots + P(A_k).$$

Ezt behelyettesítve az előző egyenlőtlenségbe kapjuk a bizonyítandó

$$P(A_1 + A_2 + \dots + A_k + A_{k+1}) \leq P(A_1) + P(A_2) + \dots + P(A_k) + P(A_{k+1})$$

egyenlőtlenséget.

Az (ii) **Bonferoni-egyenlőtlenséget** a De Morgan-azonosság, a jelen tétel (i) Boole-egyenlőtlensége és a tagadás esemény valószínűségének számítása alapján kapjuk

$$\begin{aligned} P(A_1 \cdot A_2 \cdot \dots \cdot A_n) &= 1 - P(A_1^c + A_2^c + \dots + A_n^c) \geq 1 - (P(A_1^c) + P(A_2^c) + \dots + P(A_n^c)) = \\ &= 1 - (1 - P(A_1) + 1 - P(A_2) + \dots + 1 - P(A_n)) = P(A_1) + P(A_2) + \dots + P(A_n) - n + 1. \quad \blacklozenge \end{aligned}$$

2.5.9. PÉLDA. Bonferoni-egyenlőtlenség élessége

Tegyük fel, hogy van $n=10$ eseményünk, amelyekre

$$P(A_1) = P(A_2) = P(A_3) = P(A_4) = P(A_5) = P(A_6) = P(A_7) = P(A_8) = P(A_9) = P(A_{10}) = p.$$

Becsüljük meg a

$$P(A_1 A_2 A_3 A_4 A_5 A_6 A_7 A_8 A_9 A_{10})$$

szorzat esemény valószínűségét,

(i) ha $p = 0.99$

(ii) ha $p = 0.01$

MEGOLDÁS

(i) Mivel az $A_1 \cdot A_2 \cdot A_3 \cdot A_4 \cdot A_5 \cdot A_6 \cdot A_7 \cdot A_8 \cdot A_9 \cdot A_{10} \subseteq A_1$ és a részesemény valószínűsége kisebb vagy egyenlő, mint a nagyobb esemény valószínűsége (lásd 2.5.5. tétel), ezért

$$> P(A_1 \cdot A_2 \cdot A_3 \cdot A_4 \cdot A_5 \cdot A_6 \cdot A_7 \cdot A_8 \cdot A_9 \cdot A_{10}) \leq 0.99$$

$$P(A_1 A_2 A_3 A_4 A_5 A_6 A_7 A_8 A_9 A_{10}) \leq 0.99 \quad (1.5.1)$$

Nézzük meg, hogy a 2.5.8. (ii) Bonferoni-egyenlőtlenséget használva milyen alsó korlátot kapunk a fenti valószínűségekre

$$> P(A_1 \cdot A_2 \cdot A_3 \cdot A_4 \cdot A_5 \cdot A_6 \cdot A_7 \cdot A_8 \cdot A_9 \cdot A_{10}) \geq 10 \cdot 0.99 - 10 + 1$$

$$0.90 \leq P(A_1 A_2 A_3 A_4 A_5 A_6 A_7 A_8 A_9 A_{10}) \quad (1.5.2)$$

Tehát a szorzat esemény kétoldalú becslése

$$0.9 \leq P(A_1 A_2 A_3 A_4 A_5 A_6 A_7 A_8 A_9 A_{10}) \leq 0.99$$

Az eltérés a két oldal között $0.99 - 0.9 = 0.09$, amely jó becslésnek mondható.

(ii)

Ismételjük meg a becslést, ha a valószínűségek mindegyike $p = 0.01$. Ekkor a 2.5.8. (ii)

Bonferoni-egyenlőtlenséget használva

$$> P(A_1 \cdot A_2 \cdot A_3 \cdot A_4 \cdot A_5 \cdot A_6 \cdot A_7 \cdot A_8 \cdot A_9 \cdot A_{10}) \geq 10 \cdot 0.01 - 10 + 1$$

$$-8.90 \leq P(A_1 A_2 A_3 A_4 A_5 A_6 A_7 A_8 A_9 A_{10}) \quad (1.5.3)$$

Mivel a valószínűség mindig nem-negatív, ezért a negatív alsó korlát nem ad használható információt. Ami azt jelenti, hogy a (ii) Bonferoni-egyenlőtlenséget "kis" valószínűségek mellett nem célszerű alkalmazni. Használjuk tehát a Bonferoni-egyenlőtlenséget nagy valószínűségű események szorzatának becslése során.

2.6. Példák diszkrét valószínűségi mezők alkalmazására és ezek Maple

szimulációja

Diszkrét mező esetén egy esemény valószínűségének kiszámításához az alábbi **problémákra** kell választ keresnünk.

1. lépés. Létre kell hozni az eseménytér halmazát!

2. lépés. Az elemi események valószínűségeit meg kell határozni!

3. lépés. A feladatban szereplő eseményt elő kell állítani elemi események összegeként!

4. lépés. A feladatban szereplő esemény valószínűségét additivitási szabály alapján tudjuk meghatározni!

2.6.1 PÉLDA. Szerencsejáték 3 kockával és annak szimulációja

A játék és szabályai

A játékos mond egy egész számot 1 és 6 között, majd feldob 3 szabályos kockát. A játékos akkor **nyer**, ha az általa választott szám megjelenik valamelyik kockán. Ezzel összhangban a játékos akkor **veszít**, ha az általa mondott szám nem jön ki egyik kockán sem.

Mekkora valószínűséggel **nyer** a játékos?

Az alábbi **4 különböző megoldás felvonultatása során** fokozatosan függetlenítjük magunkat a Maple-program használatától. Kezdetben csak a Maple-re támaszkodunk. Később csak számolunk vele, majd a végén elméleti úton megkeressük a valószínűséget.

1. MEGOLDÁS. Szimuláció Maple programmal

Mielőtt választ találnánk a fenti problémára, azelőtt készítsünk **szimulációs** vagy **Monte Carlo-programot**, amely a **nyerés relatív gyakoriságait** felrajzolja!

Használjuk a játékos által gondolt szám megadásához ugyanazt a `kocka:=rand(1..6)` eljárást, amelyet korábban már működtettünk.

```
> restart : randomize( ) : kocka := rand(1..6) : játékos := kocka( )  
                                     játékos := 5 (1.6.1)
```

A 3 kocka feldobását egy 3 elemű *dobás* nevű listával szimbolizáljuk, amelyben háromszor meghívjuk a `kocka()` eljárást.

```
> dobás := [kocka( ), kocka( ), kocka( )]  
                                     dobás := [3, 5, 4] (1.6.2)
```

Nézzük meg, hogy a *dobás* lista 3 elemét halmazzá alakítva tartalmazza-e a játékos által kigondolt számot?

```
> if játékos in convert(dobás, set) then "A játékos győzött" else "A játékos veszített" end if  
                                     "A játékos győzött" (1.6.3)
```

Miután a programrészeket külön-külön kipróbáltuk és azok működnek, csak azután tegyük egy Maple futás csoportba! Képezzünk 200 dobás sorozatot és rajzoljuk fel a relatív gyakoriságok vonalas grafikonját!

```
> dobásszám := 200 :  
   relativ_gyakoriságok := NULL : gyak := 0 :  
   for k from 1 to dobásszám do  
     dobás := [kocka( ), kocka( ), kocka( ) ] :
```



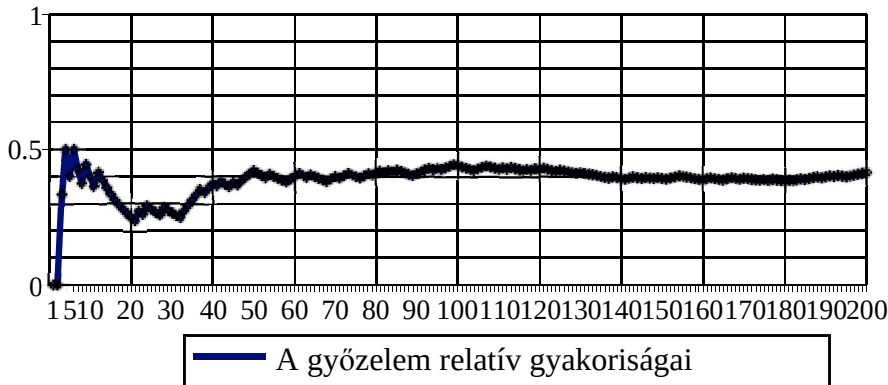
```
if játékos in convert(dobás, set) then gyak := gyak + 1 end if:
```

```
relativ_gyakoriságok := relativ_gyakoriságok,  $\frac{gyak}{k}$  :
```

```
end do:
```

```
relativ_gyakoriságok := [relativ_gyakoriságok] :
```

```
> Statistics[LineChart](relativ_gyakoriságok, thickness = 3, axis = [gridlines = [10, color = black]], tickmarks = [dobásszám, 3], legend = ["A győzelem relatív gyakoriságai"], view = [1 ..dobásszám, 0 ..1])
```



A relatív gyakoriságok nagy n esetén egyre közelebb kerülnek 0.4-hez vagy egy 0.4-hez közeli számhoz. A valószínűség **pontos** vagy **elméleti** értékét nem tudjuk ilyen módon meghatározni!

2. MEGOLDÁS. A lépések végrehajtása Maple programmal

Próbáljuk bizonyítani, hogy a játékos nyerésének esélye valóban "kevéssel" tér el 0.4-től!

1. lépés. Az eseménytér az összes 3 elemű $[d_1, d_2, d_3]$ listák halmaza, melyben a lista elemek az 1-től 6-ig terjedő kockadobások lehetnek $d_i \in \{1, 2, 3, 4, 5, 6\}$ és $i = 1, 2, 3$. Az összes ilyen listát 3 egymásba ágyazott sorozat képző *seq* eljárással tudjuk megadni.

```
> eseménytér := {seq(seq(seq([i, j, k], k = 1 ..6), j = 1 ..6), i = 1 ..6)} : N
:= nops(eseménytér)
```

$N := 216$

(1.6.4)

Látható, hogy az eseménytér elemeinek száma $216 = 6^3$.

2. lépés. Mivel a három kocka szabályos, ezért az eseménytér mindegyik kimenetele egyformán valószínű. Tehát minden elemi esemény valószínűsége egyformán $\frac{1}{216}$

$$P(\{[d_1, d_2, d_3]\}) = \frac{1}{216}, \text{ ahol } d_1, d_2, d_3 \in \{1, 2, 3, 4, 5, 6\}.$$

3. lépés. Adjuk meg az eseménytérben a játékos nyerésének eseményét! Mivel nem ismerjük a játékos által kigondolt számot, ezért tetszőlegesen felvehetjük azt 1 és 6 között. Vegyük most az 1 értéket. Próbálja ki az olvasó, ha ezt a számot átírja 2, 3, 4, 5 vagy 6 értékre, akkor is ugyanazt a végeredményt kapja!

```
> játékos := 1 : kedvező_esetek := NULL :
```

```
for k from 1 to N do
```

```
dobás := eseménytér[k] :
```

```
if játékos in convert(dobás, set) then kedvező_esetek := kedvező_esetek, dobás end if:
```

end do:

> $kedvező_esetek_száma := nops([kedvező_esetek])$
 $kedvező_esetek_száma := 91$ (1.6.5)

> $P(\text{nyerés}) = \frac{kedvező_esetek_száma}{N}; evalf(\%)$
 $P(\text{nyerés}) = \frac{91}{216}$
 $P(\text{nyerés}) = 0.4212962963$ (1.6.6)

3. MEGOLDÁS. Elméleti megoldás szita-formulával

Vajon hogyan kaphatjuk meg **elméleti úton** a nyerés kedvező eseteinek 91 számát, amelyet **Maple-program** segítségével az előbb **leszámoltunk**?

Ennek a kérdésnek a megválaszolásához használjuk a szita formulát. Tekintsük az

$$A_k = \{a \mid a \text{ a } k \text{-ik dobás értéke megegyezik a játékos által kigondolt számmal}\}$$

eseményeket, ahol $k = 1, 2, 3$ lehet. A játékszabály szerint a nyerés eseménye az $A_1 + A_2 + A_3$ összeg esemény, mert ekkor valamelyik kockán megkapjuk a kigondolt számot. A szita formula szerint ennek valószínűsége

$$P(A_1 + A_2 + A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1 \cdot A_2) - P(A_1 \cdot A_3) - P(A_2 \cdot A_3) + P(A_1 \cdot A_2 \cdot A_3)$$

képlettel számolható.

Annak valószínűsége, hogy valamelyik dobás megegyezik a gondolt számmal $\frac{1}{6}$, mert összesen 6 lehetőség van egyforma eséllyel és ebből csak egy a gondolt szám. Tehát

$$P(A_1) = P(A_2) = P(A_3) = \frac{1}{6}.$$

Annak valószínűsége, hogy a 3 dobás közül valamelyik kettő megegyezik a gondolt számmal $\frac{1}{36}$, mert összesen 36 lehetőség van egyforma eséllyel és ebből csak egy esetben egyezik mindkét szám a gondolt számmal. Tehát

$$P(A_1 \cdot A_2) = P(A_1 \cdot A_3) = P(A_2 \cdot A_3) = \frac{1}{36}.$$

Annak valószínűsége, hogy mindhárom dobás egyezik a gondolt számmal $\frac{1}{216}$, mert összesen 216 lehetőség van egyforma eséllyel és ebből csak egy olyan van, amelyben mindhárom szám megegyezik a gondolt számmal. Tehát

$$P(A_1 \cdot A_2 \cdot A_3) = \frac{1}{216}.$$

A szita formula alapján

$$\begin{aligned} P(A_1 + A_2 + A_3) &= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} - \frac{1}{36} - \frac{1}{36} - \frac{1}{36} + \frac{1}{216} = \frac{3}{6} - \frac{3}{36} + \frac{1}{216} \\ &= \frac{36 \cdot 3 - 6 \cdot 3 + 1}{216} = \frac{91}{216} \end{aligned}$$

Ezzel elméleti úton is igazoltuk, hogy a játékos nyerésének esélye kb. 0.42129629, amelyet a szimulációs program "jól" közelített.

4. MEGOLDÁS. Tagadás esemény valószínűségének számításával

Abban az esetben, ha túl bonyolult az eredeti esemény valószínűségének számítása, akkor

próbálkozzunk a tagadás esemény valószínűségének számításával

$$P(a \text{ játékos nyer}) = 1 - P(a \text{ játékos veszít}).$$

Korábban felírtuk a nyerés eseményét

$$A \text{ játékos nyer} = A_1 + A_2 + A_3,$$

melynek valószínűségét kiszámolni csak több lépésb után tudtuk.

Ennek tagadása a De Morgan-azonosság alapján

$$A \text{ játékos veszít} = (A_1 + A_2 + A_3)^c = A_1^c \cdot A_2^c \cdot A_3^c.$$

Szavakkal azt mondhatjuk, hogy a játékos akkor veszít, ha egyik dobás sem lesz egyenlő a kigondolt számmal. Ez mindhárom dobásnál ötféleképpen következhet be, amikor kimarad a gondolt szám. A kombinatorikánál (lásd 3.1. fejezet) alkalmazott szorzás szabályt használva ezen esetek száma az 5 elem harmad osztályú ismétléses variációja: $V_5^{3,i} = 5 \cdot 5 \cdot 5 = 5^3 = 125$.

Ha az eseménytér összes 216 esetéből elvesszük a veszítés 125 esetét, akkor $216 - 125 = 91$ kedvező esetet kapunk, amit Maple programmal is és szita-formulával is már korábban megkaptunk

$$P(a \text{ játékos nyer}) = 1 - \frac{125}{216} = \frac{91}{216} \approx 0.42129629.$$

Látható, hogy ebben a feladatban könnyebben számoltuk ki a tagadás esemény valószínűségét, mint az eredeti esemény valószínűségét. Érdekes tehát mindig átgondolni még a számítás részleteinek megkezdése előtt, hogy a tagadás esemény valószínűségének számítása milyen nehézségű, az eredeti esemény valószínűségéhez viszonyítva.

2.6.2 PÉLDA. Monthy Hall játék eseménytere és a nyerés valószínűsége

A játék és a szabályok. (A játék internetes változata megtalálható a <http://www.stayorswitch.com> címen)

Show műsor keretében egy játékosnak felajánlják, hogy nyerhet egy új autót, ha szerencséje van és jól dönt. A játékszabályok a következők:

1. szabály.	Három zárt ajtó van a színpadon. Kettő ajtó mögött egy-egy kecskét helyeztek el és a harmadik ajtó mögött található az új autó. Az elhelyezések véletlenszerűek, de a műsorvezető tudja, hogy melyik mögött milyen nyeremény található.
2. szabály.	A játékos rámutat a három ajtó közül az egyikre.
3. szabály.	A játékos által ki nem választott 2 ajtó közül a játékvezető kinyitja azt az ajtót, amelyik mögött kecske van. Ha mindkettő ilyen, akkor szabadon dönthet a játékvezető.
4. szabály.	A megmaradt 2 ajtó egyike biztosan az autót rejti. A játékos eldönti, hogy marad-e az eredeti döntésénél vagy megváltoztatja azt.
5. szabály.	A játékvezető végül megmutatja a nyereményt a játékos által másodszorra kiválasztott ajtó mögött, majd ellenőrzésképpen kinyitja a harmadik ajtót is.

Megoldás

Az ajtókat jelölje a , b és c . Legyen az autó az a jelű ajtó mögött minden esetben. Válasszuk az eseménytérnek az

$$\Omega = \{a, b, c\} \times \{a, b, c\}$$

Descartes-szorzatot, amelyet részletesen kiírunk az alábbi táblázatban

Monthy Hall	A játékvezető választása
--------------------	--------------------------

játék eseménytere		a	b	c
A játékos választása	a	(a, a)	(a, b)	(a, c)
	b	(b, a)	(b, b)	(b, c)
	c	(c, a)	(c, b)	(c, c)

2.6.1. táblázat. A Monty Hall játék kimeneteleiből álló eseménytér

Tehát a rendezett pár első értéke **a játékos választása**, míg a második érték **játékvezető döntése**.

(i) Adjuk meg a Monty Hall elemi eseményeinek valószínűségeit táblázattal!

A **valószínűségek meghatározásánál** figyelemmel kell lenni a játékszabályok alapján az alábbiakra

(i_1) a játékos nem tudja, hogy az a jelű ajtó mögött van az autó

(i_2) a játékvezető soha nem választja az a jelű ajtót, mert tudja, hogy mögötte az autó van

(i_3) a játékvezető nem választhatja azt az ajtót, mint amit a játékos már kiválasztott

(i_4) ha a játékos az a ajtót választotta, akkor szabadon dönthet a játékvezető a b és c ajtó között.

Figyelemmel a fentiekre az alábbi valószínűségi táblázatot kapjuk

Monthy Hall játék elemi eseményeinek valószínűsége		A játékvezető választása		
		a	b	c
A játékos választása	a	$P(a, a) = 0$	$P(a, b) = \frac{1}{6}$	$P(a, c) = \frac{1}{6}$
	b	$P(b, a) = 0$	$P(b, b) = 0$	$P(b, c) = \frac{1}{3}$
	c	$P(c, a) = 0$	$P(c, b) = \frac{1}{3}$	$P(c, c) = 0$

2.6.2. táblázat. A Monty Hall játék elemi eseményeihez rendelt valószínűségek

A táblázatban szereplő valószínűségeket az $(i_1) - (i_4)$ követelmények alapján kaptuk.

Az (i_2) feltétel alapján az első oszlopban csupa 0 valószínűség áll: $P(x, a) = 0$, ahol $x \in \{a, b, c\}$.

Az (i_3) feltétel alapján az főátlóban csupa 0 valószínűség áll: $P(x, x) = 0$, ahol $x \in \{a, b, c\}$.

Az (i_4) feltétel alapján az első sor két utolsó valószínűsége egyenlő: $P(a, b) = P(a, c)$.

Az (i_1) feltétel alapján

$$P(\{a \text{ játékos az "a" ajtót választja}\}) = P(\{a \text{ játékos a "b" ajtót választja}\}) = P(\{a \text{ játékos a "c" ajtót választja}\}) = \frac{1}{3}.$$

A valószínűség additivitási axiómája alapján

$$P(\{a \text{ játékos az "a" ajtót választja}\}) = P(a, a) + P(a, b) + P(a, c) = 2 \cdot P(a, b) = \frac{1}{3},$$

ezért $\frac{1}{6}$ az első sor két utolsó valószínűsége. A második és harmadik sorban van két-két nulla

valószínűség és az összeg $\frac{1}{3}$ kell, hogy legyen. Ezért mindkét hiányzó valószínűség $\frac{1}{3}$.

(ii) Adjuk meg a Monty Hall játékban a játékos nyerési esélyét, ha a játékos nem változtatja meg eredeti döntését

Feltételezve, hogy a játékos az **eredeti döntését nem változtatja meg**, akkor vajon meg tudjuk-e adni a játékos nyerési eseményét és ezzel a valószínűségét is az $\Omega = \{a, b, c\} \times \{a, b, c\}$ eseménytérben?

Az aggodalom jogos, mert az eseménytérben csak a játékos első döntése és a játékvezető döntése látható és nincs beépítve a játékos második döntése. Ez a példa elegánsan mutatja azt, hogyan kell az eseménytér kimeneteleit minél kevesebb elemből megszerkeszteni úgy, hogy mégis a felvetett problémákra egyértelműen válaszolni tudjunk.

A játékos akkor **nem változtatja meg** eredeti döntését,

- ha az első választása az a ajtó volt, akkor a játékvezető döntése után is az a ajtót választja
- ha az első választása a b ajtó volt, akkor a játékvezető döntése után is a b ajtót választja
- ha az első választása a c ajtó volt, akkor a játékvezető döntése után is a c ajtót választja

A feltételezések szerint az autó az a ajtó mögött van. Ezért ha a játékos nem változtat döntésén, akkor úgy nyerhet, ha induláskor az a ajtót választotta. Ugyanis, ha elsőre az a ajtót választja a játékos és marad ennél a döntésénél, akkor megnyeri az autót az a ajtó kinyitása után. Minden más esetben veszít a játékos, ha marad az eredeti választásánál. Ezért a nyerés eseménye az első sorban levő két nem nulla valószínűségű kimenetel.

A játékos nyerési eseménye, ha a játékos marad az eredeti döntésénél		A játékvezető választása		
		a	b	c
A játékos választása	a	(a, a)	(a, b)	(a, c)
	b	(b, a)	(b, b)	(b, c)
	c	(c, a)	(c, b)	(c, c)

2.6.3. táblázat. A játékos **nyerési eseménye sárgával jelölve**, ha a játékos **marad az eredeti döntésénél**

A 2.6.2. táblázat valószínűségei alapján

$$P(\text{nyerés, ha a játékos nem változtat}) = P(a, b) + P(a, c) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

Tehát $\frac{1}{3}$ a játékos nyerési esélye, ha nem változtatja meg eredeti döntését.

(iii) Adjuk meg a Monty Hall játékban a játékos nyerési esélyét, ha a játékos megváltoztatja eredeti döntését

Ha a játékos **megváltoztatja eredeti döntését**, akkor úgy nyerhet, ha induláskor vagy a b ajtót

vagy a c ajtót választja. Ugyanis, ha **elsőre az a ajtót választja** a játékos és **megváltoztatja** döntését, akkor nem nyerheti meg az autót, amely az a ajtó mögött van.

Ha a játékos elsőre a **b ajtót választja**, akkor a játékvezető a c ajtót kell, hogy kinyissa, mert az mögött nincs az autó. Mivel a játékos megváltoztatja döntését, ezért az a ajtót kell, hogy válassza, mert a c ajtót már kinyitotta a játékvezető. Ekkor tehát nyer a játékos.

Hasonló gondolat menettel, ha a játékos elsőre a **c ajtót választja**, akkor a játékvezető a b ajtót kell, hogy kinyissa, mert az mögött nincs az autó. Mivel a játékos megváltoztatja döntését, ezért az a ajtót kell, hogy válassza, mert a b ajtót már kinyitotta a játékvezető. Ekkor tehát szintén nyer a játékos.

Ezért a nyerés eseménye az eseménytér második és harmadik sorában szereplő két nem nulla valószínűségű kimenetel.

A játékos nyerési eseménye, ha a játékos megváltoztatja eredeti döntését		A játékvezető választása		
		a	b	c
A játékos választása	a	(a, a)	(a, b)	(a, c)
	b	(b, a)	(b, b)	(b, c)
	c	(c, a)	(c, b)	(c, c)

2.6.4. táblázat. A játékos **nyerési eseménye sárgával jelölve**, ha a játékos **megváltoztatja az eredeti döntését**

A 2.6.2. táblázat valószínűségei alapján

$$P(\text{nyerés, ha a játékos megváltoztatja döntését}) = P(b, c) + P(c, b) = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$$

Tehát $\frac{2}{3}$ a játékos nyerési esélye, ha megváltoztatja eredeti döntését.

Összesítve az eddigieket azt kaptuk, hogy **kétszer akkora a játékos nyerési esélye**, ha az **eredeti döntését megváltoztatja**. Tehát a Monty Hall típusú játéknál mindig érdemes megváltoztatni eredeti döntésünket, mert így kétszer akkora eséllyel nyerhetünk a játékban.

(iv) Monthy Hall játék szimulációja

Most, hogy a Monty Hall játékban a játékos nyerési esélyeit a különböző stratégiák mellett meghatároztuk, írjunk **szimulációs programot a változtatás és a maradás** stratégiáira.

Használjuk az $ajto := rand(1..3)$ véletlenszám generátort az 1, 2 és 3 számokra, mely a 3 ajtó közül véletlenszerűen kiválasztja valamelyiket, amelyik mögé az autót rejtik.

Készítünk két darab $N = 200$ elemű listát. Az egyik lista az *auto* nevű, mely tartalmazza annak az ajtónak a sorszámát, amely mögé az autót rejtették. A másik a *játékos* nevű lista, mely tartalmazza a játékos első tippjét az ajtó sorszámára.

```
> restart ;
> randomize( ) ;
> ajto := rand(1..3) ;
> N := 200 ;
> auto := [seq(ajto( ), k = 1..N) ] ;
> játékos := [seq(ajto( ), k = 1..N) ] ;
auto := [1, 3, 2, 2, 1, 1, 3, 3, 3, 1, 3, 3, 1, 3, 2, 3, 1, 3, 3, 2, 1, 3, 3, 1, 1, 2, 2, 2, 1, 1, 3,
```

```

3, 1, 2, 3, 2, 1, 3, 1, 2, 3, 2, 3, 3, 2, 1, 3, 1, 1, 2, 1, 3, 1, 1, 2, 3, 3, 3, 1, 1, 1, 1, 2, 1,
1, 2, 2, 1, 1, 1, 2, 3, 1, 1, 1, 1, 3, 3, 3, 1, 3, 1, 2, 3, 1, 3, 2, 2, 1, 2, 1, 3, 3, 2, 3, 2, 1,
1, 3, 2, 3, 2, 2, 2, 1, 1, 3, 1, 1, 3, 2, 3, 2, 2, 2, 3, 3, 2, 2, 1, 1, 2, 2, 2, 3, 3, 3, 3, 1, 3,
3, 1, 1, 1, 3, 3, 2, 1, 1, 1, 1, 2, 3, 1, 1, 3, 2, 2, 2, 2, 3, 1, 1, 1, 1, 1, 3, 2, 1, 2, 2, 3, 3,
2, 1, 3, 2, 1, 3, 3, 1, 2, 1, 1, 2, 2, 1, 2, 2, 1, 3, 3, 1, 2, 2, 1, 3, 1, 1, 1, 1, 3, 2, 1, 2, 3,
2, 1, 3, 1]

```

```

játékos := [1, 2, 2, 1, 2, 3, 2, 2, 1, 3, 1, 2, 3, 1, 3, 2, 3, 1, 3, 2, 1, 1, 2, 3, 2, 2, 3, 2, 2, 2, (1.6.7)
2, 1, 2, 2, 2, 3, 3, 2, 3, 3, 2, 3, 3, 1, 2, 2, 3, 3, 3, 2, 3, 2, 1, 3, 1, 2, 1, 2, 3, 2, 2, 2, 2,
2, 3, 1, 2, 3, 1, 3, 2, 2, 3, 1, 2, 3, 1, 2, 2, 3, 2, 3, 1, 2, 2, 3, 3, 2, 2, 3, 2, 2, 1, 1, 1, 3,
1, 1, 2, 1, 1, 3, 3, 3, 2, 2, 3, 3, 1, 1, 3, 2, 2, 1, 2, 3, 1, 2, 1, 1, 3, 3, 1, 1, 3, 3, 3, 2, 2,
2, 3, 2, 2, 1, 3, 3, 3, 2, 3, 1, 2, 3, 3, 1, 2, 1, 3, 3, 3, 1, 3, 2, 2, 1, 1, 3, 1, 2, 1, 1, 3, 1,
1, 1, 1, 3, 3, 1, 1, 3, 3, 2, 1, 2, 2, 2, 3, 3, 3, 1, 3, 1, 1, 2, 1, 2, 2, 2, 2, 3, 2, 2, 3, 2, 3,
1, 3, 3, 2, 2]

```

1. Stratégia. Ha a játékos **marad** az eredeti döntésénél

A játékos ilyen stratégia mellett akkor **nyer**, ha azt az ajtót választja elsőre, amelyik mögött az autó van. Hiszen a játékos nem változtat az eredeti döntésén és így a második döntés után megnyeri az autót.

```

> nyerések1 := NULL :
for k from 1 to N do
  if auto[k] = játékos[k] then
    nyerések1 := nyerések1, 1 else nyerések1 := nyerések1, 0
  end if:
end do:
nyerések1 := [nyerések1]
nyerések1 := [1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 0, 0, 0, 0, 1, 0, 1, 0, (1.6.8)
0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0,
1, 0, 0, 0, 1, 0, 1, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 1, 1, 0, 1, 0, 1, 0, 0, 0, 0, 1, 1, 1, 0,
0, 0, 1, 0, 0, 1, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0,
0, 0, 0, 1, 1, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0]

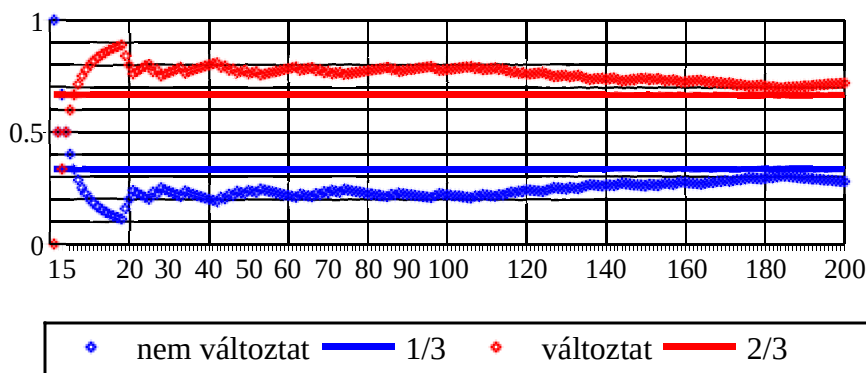
```

Látható, hogy az elméleti meggondolások megjelentek a program írása során is, de a program szervezése jobban közelít a játékhoz, mint az elméleti meggondolás. Itt ugyanis véletlenszerűen tudjuk változtatni az **autó helyzetét**, míg az elmélet során fixen feltételeztük, hogy az (a) ajtó mögött van. A program ellenőrzi, hogy a két lista megfelelő elemei rendre egyenlők-e! Ha ez teljesül, akkor a *nyerések1* lista végéhez hozzátesz egy 1 számot, egyébként egy 0 számot tesz hozzá. Jelezve a nyeres és veszítés tényét!

2. Stratégia. A játékos **megváltoztatja** eredeti döntését

A játékos ilyen stratégia mellett akkor **veszít**, ha azt az ajtót választja elsőre, amelyik mögött az autó van. Hiszen a játékos megváltoztatja az eredeti döntését és így a második döntés után egy kecskét tartalmazó ajtóra mutat. Ha elsőre a két kecskét tartalmazó ajtó közül valamelyikre mutat, akkor a második döntése után az autót tartalmazó ajtót választja, mert egyrészt megváltoztatja döntését és másrészt a játékvezető időközben kinyitotta azt az ajtót, amelyik mögött a másik kecske van.

= ["nem változtat", "1/3", "változtat", "2/3"])



A **kék** színű pontok a **maradás** stratégiája mellett a relatív gyakoriságok, amelyek a kék színű vízszintes $\frac{1}{3}$ valószínűséghez tartanak.

A **piros** színű pontok a **változtatás** stratégiája mellett a relatív gyakoriságok, amelyek a piros színű vízszintes $\frac{2}{3}$ valószínűséghez tartanak.

Tehát a Maple programmal szimulációkat végezve olyan a relatív gyakoriság értékeket kaptunk, amelyek alátámasztják az elméleti megfontolásokat.

Látható, hogy a **szimulációs programrészek** és az **elméleti megfontolások** kiegészítik egymást. Egyik megerősíti a másiknál kapott eredményeket. Kezdek ellenőrizhetik elméleti számításait a Maple programmal, ha a programot megfelelően írják meg.

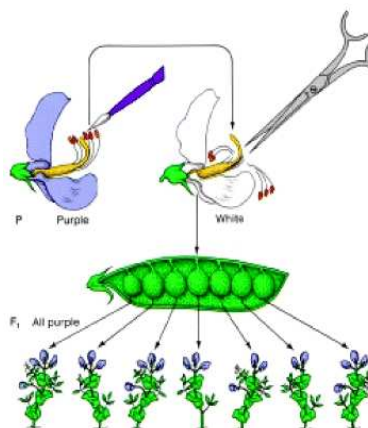
2.6.3 PÉLDA. Mendel kísérlete az öröklítés vizsgálatára

Gregor Mendel (1822-1884) **borsókkal** végzett előre jól megtervezett keresztezési **kísérleteket**. Mendel idejében az öröklésre a keveredés elmélet volt a jellemző, melyszerint az utód öröklí mindkét szülő tulajdonságait keverten, biztosítva ezzel az élővilág sokféleségét. A fajok fennmaradására pedig Darwin szelekciós elmélete volt a meghatározó. Ezzel ellentétben Mendel megfigyelte a kolostorban, hogy a zöldborsók virága vagy teljesen fehér vagy teljesen lila színű, soha sem foltos vagy rózsaszínű. Kereste azt az információ **örökítő anyagot** (amit ma **génnek** nevezünk), amely felelős a virág **szín tulajdonságának** örökítéséért a szülőkről az utódokra. Ezért a borsók egy részénél eltávolította a virágokból a porzó (vagy férfi) nemi szervet és így **anyai** (bibés, magtermő) szülőket hozott létre. Hasonlóan a borsók másik részénél eltávolította a virágokból a bibét (vagy női) nemi szervet és így **apai** (porzós, megtermékenyítő) szülőket hozott létre. (lásd 2.6.5. ábra) Az így kapott szülőkkel irányított keresztezési kísérleteket tudott végezni, megakadályozva ezzel a **véletlenszerű beporzásokat**. Addig végzett irányított keresztezést, amíg létrehozta a tiszta **lila** virágú növényeket (amelyek csak lila-virágú utódokat hoztak létre) és a tiszta **fehér** virágú növényeket (amelyek csak fehér-virágú utódokat hoztak létre). Az ilyen egyedeket mai szóhasználatnál „tiszta származéksorú” (vagy homozigóta) egyedeknek nevezünk.



2.6.5. Ábra. Zárvertermő növények virágjának felépítése

A kétféle homozigóta egyedek vegyes (fehér virágúak és a lila virágúak) keresztezésével létrehozta az első vagyis F1 (filius, gyermeki) hibrid generációt. Ezután az F1 egyedek egymás közötti keresztezésével létrehozta a második (F2) hibrid generációt.



2.6.6. Ábra. Mendel kísérlete a tiszta virágzatok létrehozására, majd a keresztező beporzásra

Mendel meglepődve tapasztalta, hogy az F1 utódok mindegyike lila virágú lett. Nagyszámú F2 utódot vizsgált és statisztikai szempontból azt kapta, hogy 3:1 arányban lila:fehér virágúak lettek. Vizsgálta, hogy az utódokban hogyan jelenik meg a szülők (P=parents) örökítő anyaga. Kétféle örökítő anyagot, mai szóhasználattal gént feltételezett, nevezzük ezeket L=lila és f=fehér színű virág örökítő anyagnak, génnek. A megfigyelései alapján feltételezte, hogy minden utód egy gént örököl apai ágon és egy gént anyai ágon. Így az F1 generáció génállománya a következőképpen néz ki az ún. **Punnett négyzet** alapján.

--	--

Az F1 generáció génállománya		Apai gének (♂) 	
		L	L
Anyai gének (♀) 	f	(L,f) 	(L,f) 
	f	(L,f) 	(L,f) 

2.6.7. Táblázat. Homozigóta egyedek keresztezésével létrejött F1 egyedek génállománya

A 2.6.7 táblázat azt mutatja, hogy az anya génje (f,f) tiszta fehér virágzatú és az apa génje (L,L) tiszta lila virágzatú. Mindegyik esetben az örökítés (L,f) párt hoz létre az F1 generációban. Azonban az L lila gén **domináns** az f fehér **recesszív** génnel szemben. Ezért mindegyik egyed lila virágzatú lett. Az L betűt azért írtuk nagybetűvel, mert a gén domináns, míg az f gén recesszív, ezért őt kis betűvel írtuk.

Folytatva ezt a leírást az F2 generációra a következő Punnett négyzetet kapjuk.

Az F2 generáció génállománya		Apai gének (♂) 	
		L	f
Anyai gének (♀) 	L	(L,L) 	(L,f) 
	f	(f,L) 	(f,f) 

2.6.8. Táblázat. Heterozigóta egyedek keresztezésével létrejött F2 egyedek génállománya

A 2.6.8 táblázat szerint az apa is és az anya is F1 egyed, ezért mindegyik rendelkezik egy L és egy f génnel. Ezért az F2 generáció egyedei az alábbi kombinációban öröklük a gén párokat

$$\Omega = \{ (L, L), (L, f), (f, L), (f, f) \}$$

amely egyben a **kísérlet eseménytere**. Mivel az L gén domináns, ezért az

$$A = \{ \text{a borsó virágja lila} \} = \{ (L, L), (L, f), (f, L) \}$$

esemény valószínűsége $P(A) = \frac{|A|}{|\Omega|} = \frac{3}{4}$. A számítás során feltételeztük, hogy mindegyik

"kimenetel" egyformán valószínű. A lila virágú esemény tagadása a

$$B = \{ \text{a borsó virágja fehér} \} = \{ (f, f) \}$$

esemény és ennek valószínűsége

$$P(B) = \frac{|B|}{|\Omega|} = \frac{1}{4}. \text{ Ez azt jelenti, hogy a lila virágú és fehér}$$

virágú egyedek aránya az F2 generációban $P(A) : P(B) = \frac{3}{4} : \frac{1}{4} = 3 : 1$, amely "jól" egyezik a

Mendel által tapasztalt $\frac{705}{224} \approx 3.147$ aránnyal.

Mendel feltételezte, hogy a borsók virágjának színe az örökölt két géntől függ és nem attól, hogy melyik gént melyik szülőtől örökölte. Ezért az (L, f) és az (f, L) egyedek egyformák öröklési szempontból. Tehát Mendel joggal gondolhatta volna azt, hogy az Ω négy elemű eseménytér helyett az $\Omega_1 = \{LL, Lf, ff\}$ három elemű eseménytérrel kellene dolgoznia. Alkalmaztuk itt a genetika jelölés rendszerét, mely szerint az örölt gének betűit egymásután írjuk és nem teszünk közéjük vesszőt, valamint a sorrend tetszőleges. Viszont ilyen eseménytér mellett, ha egyenlő esélyeket adunk a 3 kimenetelnek, akkor a lila és fehér virágok aránya $\frac{2}{3}$ lenne, ellentétben a kísérlethez jobban igazodó $\frac{3}{4}$ helyett. Mendel ezen gondolata azért vezetett hibás eredményre, mert az elemi események kimenetelei az Ω_1 eseménytérben nem egyformán valószínűek, hanem

$$P(\{LL\}) = P(\{ff\}) = \frac{1}{4} \text{ és } P(\{Lf\}) = \frac{1}{2}.$$

Ahonnán az L gén dominanciája miatt

$$P(\{a \text{ borsó virágja lila}\}) = P(\{LL, Lf\}) = P(\{LL\}) + P(\{Lf\}) = \frac{1}{4} + \frac{1}{2} = \frac{3}{4}$$

$$P(\{a \text{ borsó virágja fehér}\}) = P(\{ff\}) = \frac{1}{4}$$

ugyanazt kapjuk, mint az Ω eseménytérben. (Vesd össze az 1.5.9. példánál mondott ajánlással!)

2.7. Példák folytonos valószínűségi mezők alkalmazására és ezek Maple szimulációja

Folytonos eseménytér mellett módosulnak azok a feladatok, amelyeket a 2.6. pontban megfogalmaztunk az események valószínűségének kiszámítása érdekében.

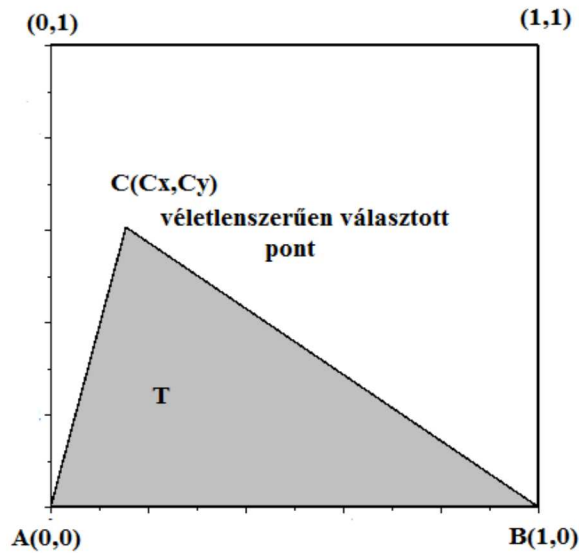
1. lépés.	A folytonos Ω eseménytér kiválasztása és $m(\Omega)$ mértékének számítása
2. lépés.	A feladatban szereplő E esemény meghatározása az eseménytéren belül
3. lépés.	Az E esemény $m(E)$ mértékének számítása
4. lépés.	A $P(E) = \frac{m(E)}{m(\Omega)}$ hányadossal számoljuk az E esemény valószínűségét.

2.7.1. PÉLDA. Véletlen háromszögek területe

Véletlenszerűen választunk egy $C(Cx, Cy)$ pontot a $(0, 0)$, $(1, 0)$, $(1, 1)$ és $(0, 1)$ csúcspontú egységnégyzet belsejében, vagyis $0 \leq Cx \leq 1$, $0 \leq Cy \leq 1$. Tekintsük az $A(0, 0)$, $B(1, 0)$ és $C(Cx, Cy)$ csúcspontok által meghatározott ABC_{Δ} háromszög T területét.

(i) Adja meg a kísérlet eseményterét!

(ii) Mekkora a valószínűsége az $E = \left\{ T \leq \frac{1}{4} \right\}$ eseménynek?



2.7.1. ábra. A C véletlenszerűen választott pont és az ABC_{Δ} háromszög T területe

MEGOLDÁS

Írjunk **szimulációs programot** a C pont véletlenszerű választására és ábrázoljuk az ABC háromszögeket az egységnyezetben. Ehhez generálni kell a C pont Cx és Cy koordinátáit a $[0, 1]$ intervallumban. Folytonos Cx és Cy generálása érdekében tudni kell az eloszlását. Az eloszlások mindegyike egyenletes a $[0,1]$ intervallumban, amelyet a *Statistics* csomag *Distribution(Uniform(0, 1))* eljárásával képezhetünk. Az egyenletesség azt jelenti, hogy a $[0,1]$ intervallum **bármely** $[a, b]$ részintervallumába az intervallum $(b - a)$ hosszával arányosan esnek pontok. Ez a választás megfelel a feladat szövegében szereplő "véletlenszerűen választunk" kifejezésnek. Képezzünk mindkét koordinátára $N = 200$ véletlen számot a *Sample* eljárással és tegyük azokat az X és Y vektorokba!

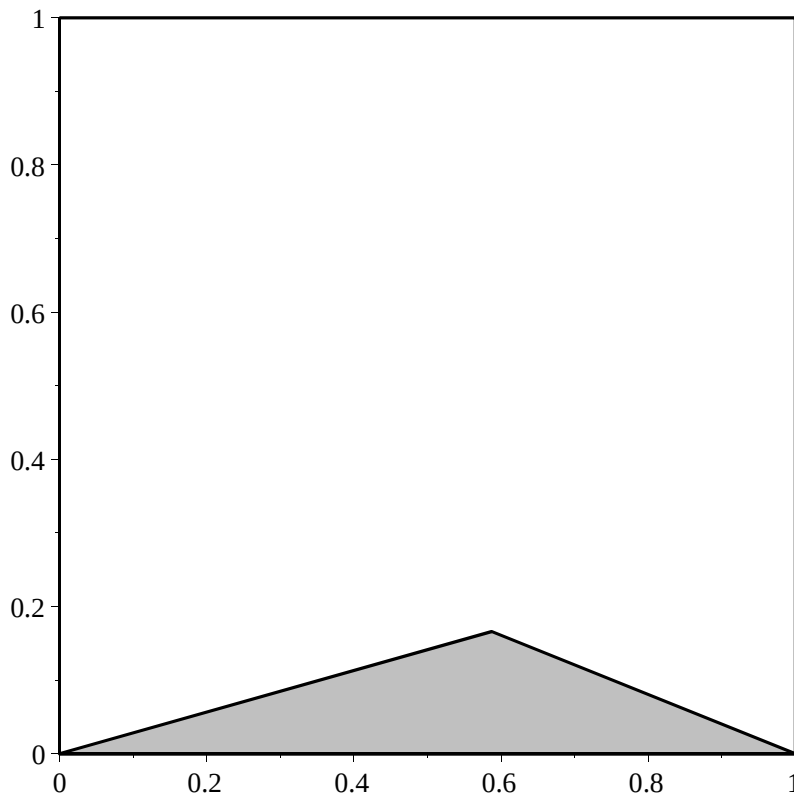
```
> restart
> with(Statistics) : with(plots) :
> Cx := Distribution(Uniform(0, 1)) :
> Cy := Distribution(Uniform(0, 1)) :
> N := 200 : randomize() :
  X, Y := Sample(Cx, N), Sample(Cy, N)
```

$$X, Y := \begin{bmatrix} 1 \dots 200 \text{ Vector}_{\text{row}} \\ \text{Data Type: float}_8 \\ \text{Storage: rectangular} \\ \text{Order: Fortran_order} \end{bmatrix}, \begin{bmatrix} 1 \dots 200 \text{ Vector}_{\text{row}} \\ \text{Data Type: float}_8 \\ \text{Storage: rectangular} \\ \text{Order: Fortran_order} \end{bmatrix} \quad (1.7.1)$$

A Maple nem mutatja az X és Y vektorokban levő értékeket, mert túl nagy helyet foglalnak. Csak helykitöltővel jelzi, hogy mennyi érték van a vektorban. Megnézhetjük viszont az értékeket egy külön ablakban, ha duplán klikkelünk a helykitöltő objektumon.
A háromszögeket a *plots* csomag *polygonplot* eljárásával rajzoljuk fel sorozatban a négyzettel

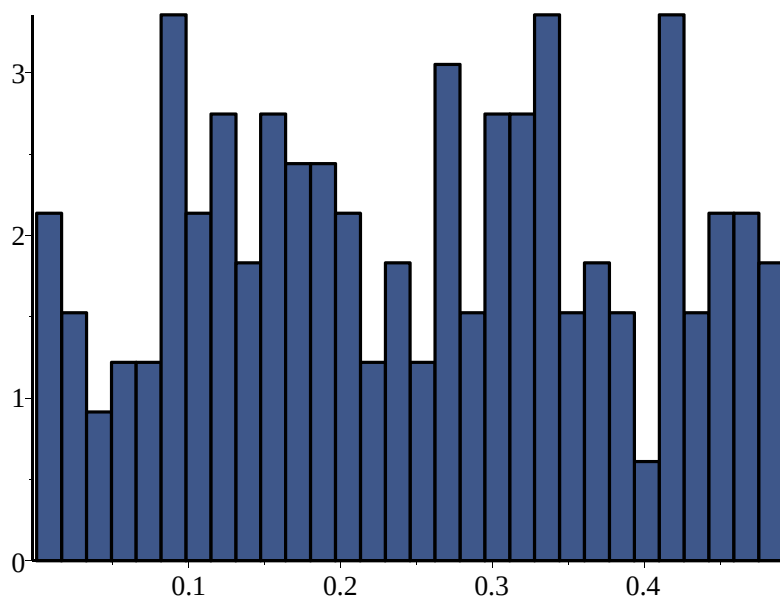
együtt, majd a *display* eljárás megjeleníti a képeket animációval az *insequence = true* opciónak köszönhetően.

```
> negyzet := plot([ [0, 0], [1, 0], [1, 1], [0, 1]], color = black) :  
> kepek := seq(display([negyzet, polygonplot([ [0, 0], [1, 0], [X[k], Y[k]]], color = grey)]), k  
= 1..20) :  
> display(kepek, insequence = true, view = [0..1, 0..1], scaling = constrained)
```



A háromszög T területének képlete $\frac{\text{alap} \cdot \text{magasság}}{2}$. Az *alap* = 1 minden esetben és a *magasság* = Cy a véletlen pont y -koordinátája. Ezért a háromszög területe $T = \frac{Cy}{2}$. Számítsuk ki mind a 200 háromszög területét és rajzoljuk fel a területek relatív gyakoriság hisztogramját!

```
> teruletek := [ seq( ( Y[k] / 2 ), k = 1..N ) ] :  
Histogram(teruletek)
```



Látható, hogy a T területek csak a generált Cy koordinátától függenek és nincs rá hatással Cx értéke. Ezért a véletlen területek leírásához elegendő valamilyen egy dimenziós halmaz és nincs szükség két dimenzióra. Az eseménytér tartalmazza a véletlen kísérlet összes lehetséges kimenetelét. A véletlen kísérlet kimenetele a Cy véletlen szám, amelynek lehetséges értékei a $[0, 1]$ intervallumba esnek. Tehát az **eseménytér** lehet a $[0, 1]$ intervallum, amelynek mértéke a hossza és ez $m(\Omega) = 1$.

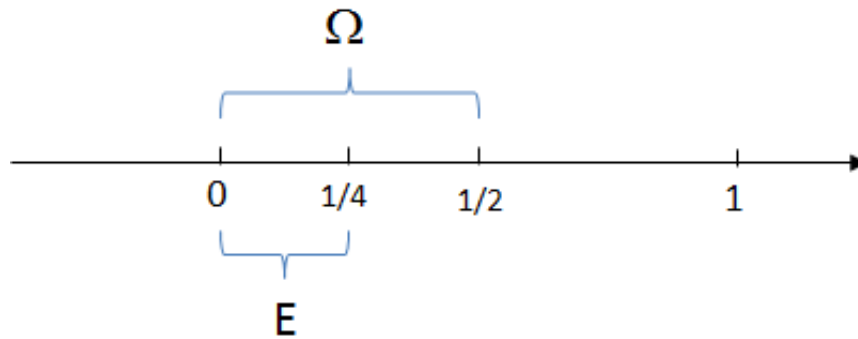
Az **(ii)** rész megoldásához meg kell adni az $\Omega = [0, 1]$ intervallumban azt az E halmazt, amelyre a T terület kisebb $\frac{1}{4}$ -nél

$$E = \left\{ Cy \in [0, 1] \mid T = \frac{Cy}{2} \leq \frac{1}{4} \right\} = \left[0, \frac{1}{2} \right]$$

Nilvánvaló, hogy a keresett E esemény a $\left[0, \frac{1}{2} \right]$ intervallum, melynek hossza $m(E) = \frac{1}{2}$.

Ezért az E esemény valószínűsége a két mérték aránya

$$P(E) = \frac{m(E)}{m(\Omega)} = \frac{\frac{1}{2}}{1} = \frac{1}{2}.$$

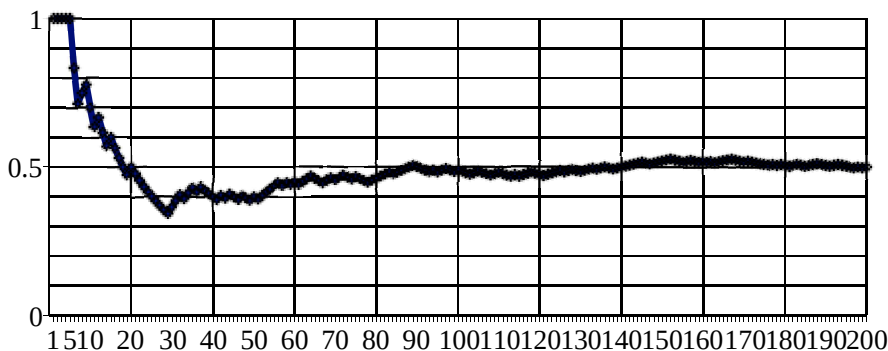


2.7.2. ábra. A 2.7.1. példa eseménytere és benne a keresett E esemény

Ellenőrizzük a korábban kiszámolt véletlen területekre, hogy a relatív gyakoriságok valóban $\frac{1}{2}$ körül ingadoznak-e!

```
> relativ_gyakoriságok := NULL : gyak := 0 :
for k from 1 to N do
  if területek[k] ≤ 1/4 then gyak := gyak + 1 end if:

  relativ_gyakoriságok := relativ_gyakoriságok, gyak/k :
end do:
relativ_gyakoriságok := [relativ_gyakoriságok] :
> Statistics[LineChart](relativ_gyakoriságok, thickness = 3, axis = [gridlines = [10, color
= black]], tickmarks = [N, 3], legend = ["A háromszög területek relatív gyakoriságai"],
view = [1..N, 0..1])
```



Látható, hogy 200 véletlen terület relatív gyakoriságai ingadoznak az $\frac{1}{2}$ érték körül.

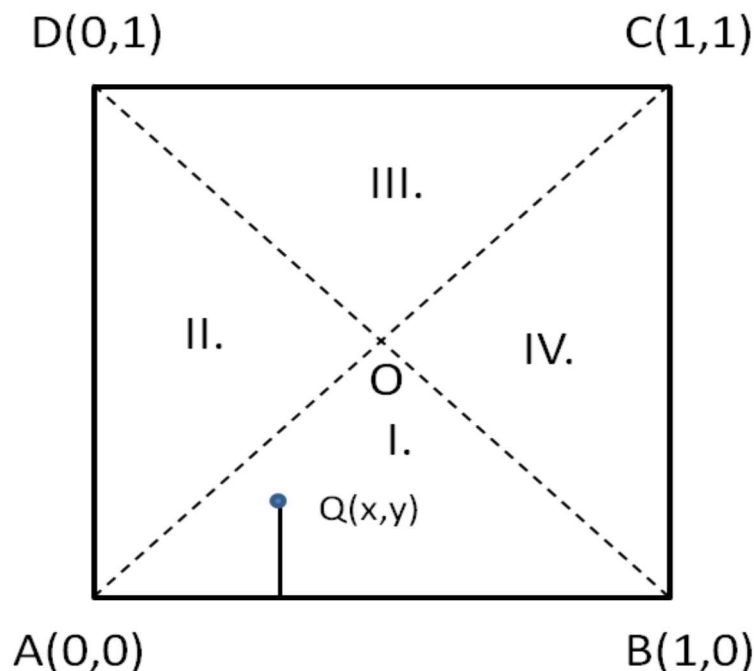
Az előző feladatot kissé módosítva olyan eseménytérhez juthatunk, amelyhez szükséges 2 dimenziós eseményteret használni.

2.7.2. PÉLDA. Véletlen pont távolsága az egységnégyzet hozzá legközelebb eső oldalától

Az $A(0, 0)$, $B(1, 0)$, $C(1, 1)$ és $D(0, 1)$ csúcspontú egységnégyzet belsejében véletlenszerűen választunk egy $Q(x, y)$ pontot. Jelölje d a Q pont távolságát az $ABCD$ négyzet hozzá legközelebb eső oldalától! (lásd 2.7.3 ábrát)

(i) Adja meg az Ω eseményteret!

(ii) Mekkora a valószínűsége az $E = \left\{ d \leq \frac{1}{4} \right\}$ eseménynek?



2.7.3. ábra. Az egységnégyzet belsejében levő $Q(x,y)$ pont d távolsága a négyzet legközelebbi oldalától

Megoldás

A 2.7.3. ábrán berajzoltunk az egységnégyzetbe egy $Q(x, y)$ pontot, amely olyan helyzetű, hogy az AB oldalhoz van legközelebb. Ez a távolság $d = y$. A négyzet AD , DC , CB többi oldalaitól mért távolságok rendre x , $(1 - y)$ és $(1 - x)$. Ezek közül a legrövidebb a d távolság, melynek képletet az x és y véletlen számokkal

$$(2.7.1) \quad d = \min(x, y, 1 - x, 1 - y).$$

Az előző feladathoz képest a különbség szembeötlő! Nevezetesen a kísérlet kimenetele a d érték, **függ mindkét generált véletlen számtól**.

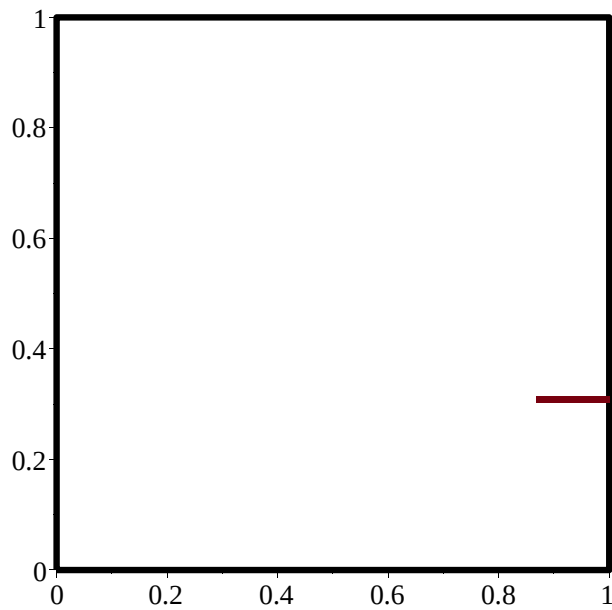
Generáljunk most is $N = 200$ véletlen számot egyenletes eloszlással a x és y koordinátákra.

```
> restart :  
> with(Statistics) : with(plots) :  
> x := Distribution(Uniform(0, 1)) :  
> y := Distribution(Uniform(0, 1)) :  
> N := 200 : randomize( ) :  
X, Y := Sample(x, N), Sample(y, N)
```

$$X, Y := \left[\begin{array}{l} 1 \dots 200 \text{ Vector}_{\text{row}} \\ \text{Data Type: float}_8 \\ \text{Storage: rectangular} \\ \text{Order: Fortran_order} \end{array} \right], \left[\begin{array}{l} 1 \dots 200 \text{ Vector}_{\text{row}} \\ \text{Data Type: float}_8 \\ \text{Storage: rectangular} \\ \text{Order: Fortran_order} \end{array} \right] \quad (1.7.2)$$

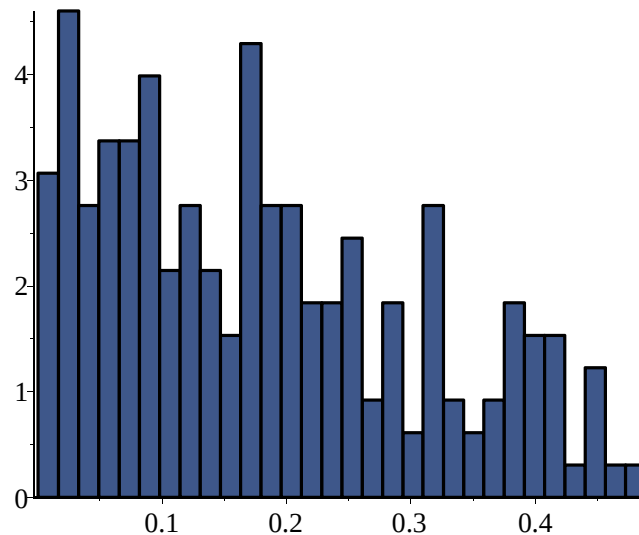
Az előző példához hasonlóan, most a négyzet belsejébe szeretnénk berajzolni a minimumot adó távolságot a 4 távolság közül. Ehhez meg kell vizsgálni, hogy a négyzet átlóival négy részre osztott háromszögek melyikébe esik a Q pont. (lásd a 2.7.3. ábrán az I.,II.,III. és IV. római számokkal jelölt szektorokat!) Ennek megfelelően a program 4 feltételes **if ... then ... end if** utasítást tartalmaz.

```
> negyzet := plot([ [0, 0], [1, 0], [1, 1], [0, 1], [0, 0]], color = black, thickness = 3) :
kepek := NULL :
for k from 1 to N do
    d := min(X[k], Y[k], 1 - X[k], 1 - Y[k])
> if (Y[k] ≤ X[k]) and (Y[k] ≤ 1 - X[k]) then vonal := plot([ [X[k], Y[k]], [X[k], 0]],
    thickness = 3) end if:
> if (Y[k] ≥ X[k]) and (Y[k] ≤ 1 - X[k]) then vonal := plot([ [X[k], Y[k]], [0, Y[k]]],
    thickness = 3) end if:
> if (Y[k] ≥ X[k]) and (Y[k] ≥ 1 - X[k]) then vonal := plot([ [X[k], Y[k]], [X[k], 1]],
    thickness = 3) end if:
> if (Y[k] ≤ X[k]) and (Y[k] ≥ 1 - X[k]) then vonal := plot([ [X[k], Y[k]], [1, Y[k]]],
    thickness = 3) end if:
    kepek := kepek, display([ negyzet, vonal]) :
end do:
display(kepek, insequence = true, view = [0..1, 0..1], scaling = constrained)
```



A távolságot a $d = \min(x, y, 1 - x, 1 - y)$ képlettel számoljuk mind a 200 esetben. Rajzoljuk fel a d távolságok gyakoriság hisztogramját!

> *távolságok* := [seq(min($X[k]$, $Y[k]$, $1 - X[k]$, $1 - Y[k]$), $k = 1..N$)] :
Histogram(*távolságok*, *caption* = A távolságok gyakorisága)



A távolságok gyakorisága

A $d = \min(x, y, 1 - x, 1 - y)$ képlet és a hisztogram alapján látható, hogy $0 \leq d \leq \frac{1}{2}$ teljesül a távolságokra. A hisztogram most nem mutat egyenletes eloszlást. Azt sejtjük a diagram alapján, hogy d értéke gyakrabban vesz 0-hoz közeli értékeket, mint $\frac{1}{2}$ - hez közelebbieket.

Az eseménytér kitalálásához elegendő a d távolság (2.7.1) képletében szereplő x és y véletlenszámok szerepét megvizsgálni. Mivel nincs semmilyen korlátozás x és y értékeire a $[0, 1]$ intervallumban, ezért az eseménytér ehhez a feladathoz az

$$\Omega = [0, 1] \times [0, 1] = \{(x, y) \mid 0 \leq x \leq 1, 0 \leq y \leq 1\}$$

egységnégyzet, melynek mértéke a területe és ez $m(\Omega) = 1$.

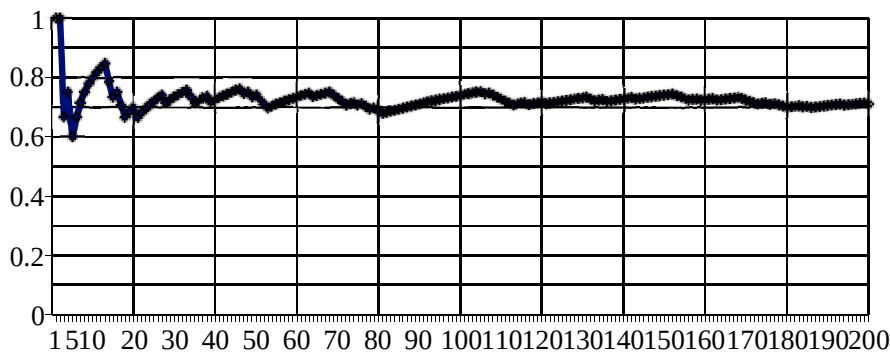
A példa (ii) részének megválaszolásához előbb számoljuk ki az

$$E = \left\{ (x, y) \mid \min(x, y, 1 - x, 1 - y) < \frac{1}{4} \right\}$$

esemény relatív gyakoriságait a generált 200 pontra, majd rajzoljuk fel vonalas diagrammal.

```
> relativ_gyakorisagok := NULL : gyak := 0 :
for k from 1 to N do
  if távolsagok[k] ≤ 1/4 then gyak := gyak + 1 end if:

  relativ_gyakorisagok := relativ_gyakorisagok, gyak/k :
end do:
relativ_gyakorisagok := [relativ_gyakorisagok] :
> Statistics[LineChart](relativ_gyakorisagok, thickness = 3, axis = [gridlines = [10, color
= black]], tickmarks = [N, 6], legend
= ["A pontok távolságának relatív gyakoriságai a legközebbi oldaltól mérve"], view
= [1..N, 0..1])
```



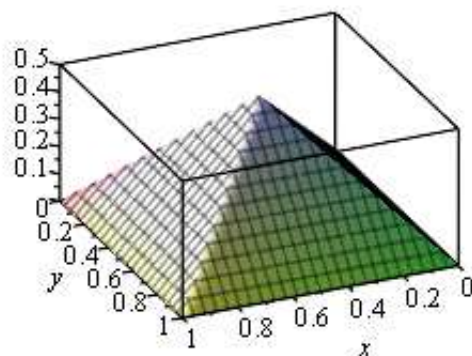
A relatív gyakoriságok ingadoznak és egyre közelebb kerülnek a $p = 0.75$ körüli értékhez.

Nézzük, hogyan lehet felrajzolni az E halmazt az egységnégyzetben!

A 2.7.3. ábrán levő I. tartományban $d = y$, a II. tartományban $d = x$, a III. tartományban $d = 1 - y$ és végül a IV. tartományban $d = 1 - x$. Szakaszonként értelmezett függvény segítségével ezt a $d(x, y)$ kétváltozós függvényt meg tudjuk adni a Maple *piecewise* eljárásával. Rajzoljuk fel $d(x, y)$ 3D-s grafikonját!

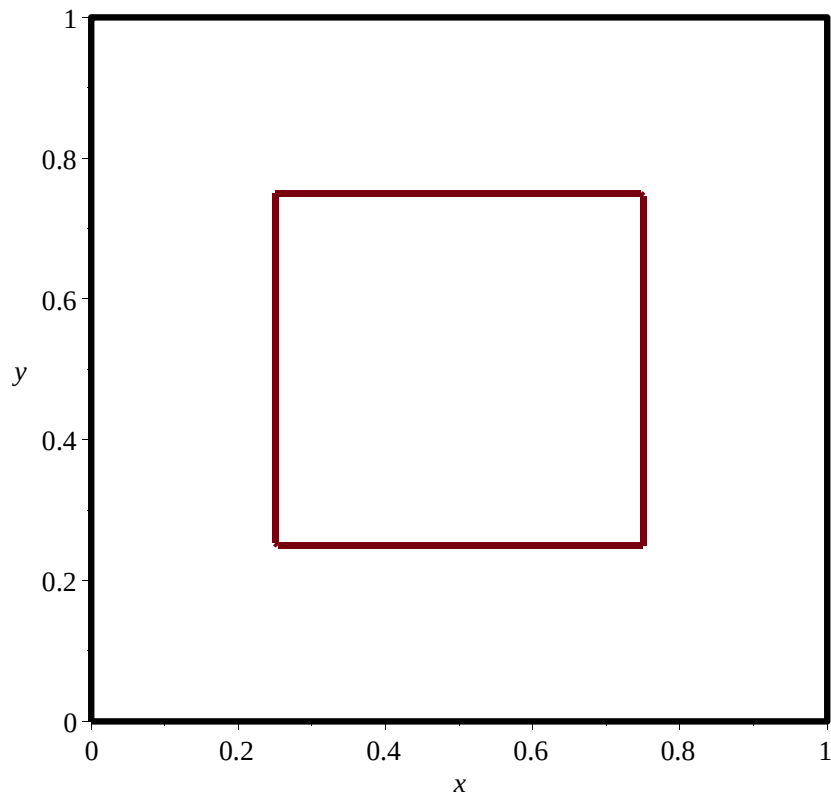
```
> d := (x, y) → piecewise( (y ≤ x) and (y ≤ 1 - x), y, (y ≥ x) and (y ≤ 1 - x), x, (y
    ≥ x) and (y ≥ 1 - x), 1 - y, (y ≤ x) and (y ≥ 1 - x), 1 - x, 0) : `d(x,y)` = d(x,
    y)
> plot3d(d(x, y), x = 0 .. 1, y = 0 .. 1, scaling = constrained, axes = boxed, orientation = [66,
    65])
```

$$d(x,y) = \begin{cases} y & y \leq x \text{ and } y \leq 1-x \\ x & x \leq y \text{ and } y \leq 1-x \\ 1-y & x \leq y \text{ and } 1-x \leq y \\ 1-x & y \leq x \text{ and } 1-x \leq y \\ 0 & \text{otherwise} \end{cases}$$



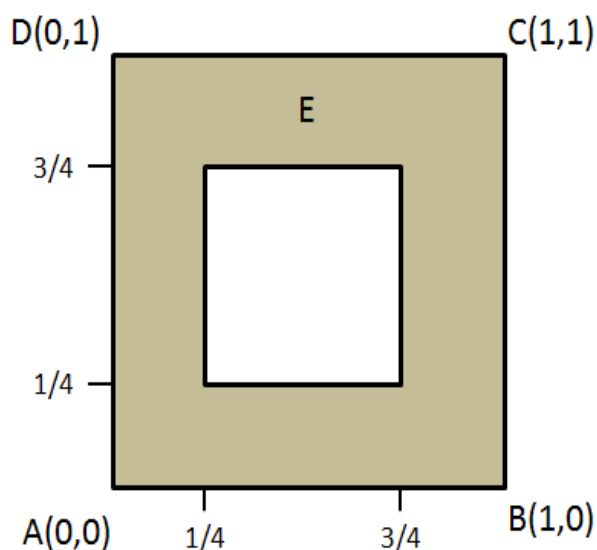
A "piramis" azt mutatja, hogy a piramis (vagy gúla) alaplapjának belsejében hogyan változik a távolság az oldalaktól és ez megegyezik a z magassággal. A négyzet határvonalán ez a távolság 0 a legkisebb értékű és legnagyobb a négyzet középpontjában, ahol $\frac{1}{2}$. Ha a $d(x, y) = \frac{1}{4}$ szintvonalat felrajzoljuk, akkor a megoldáshoz közelebb kerülünk!

```
> display( [ contourplot( d(x, y), x = 0 .. 1, y = 0 .. 1, contours = [ 1/4 ], numpoints = 5000,
    thickness = 3 ), negyzet ], caption = Az d = 1/4 magasságú szintvonal rajza )
```



Az $d = 1/4$ magasságú szintvonal rajza

Az olvasó valószínű már kitalálta közben a választ. Hiszen azok a belső pontok, amelyek $\frac{1}{4}$ -nél közelebb vannak valamelyik oldalhoz, azok az egységnégyzet oldalaival párhuzamosan húzott és tőle $\frac{1}{4}$ távolságra levő belső négyzet és a külsőnégyzet oldalai között elhelyezkedő "gyűrűszerű" tartomány. Ezt mutatja a 2.7.4. ábra.



2.7.4. ábra. A négyzet oldalaihoz $1/4$ -nél közelebb eső pontok E halmazát szürkével rajzoltuk. Most már csak az E tartomány területét vagyis az $m(E)$ mértéket kell kiszámolni. Ezt megkapjuk, ha kivonjuk az egységnyezet 1 területéből a belső négyzet területét. A belső négyzet oldalhossza

$$a = 1 - \frac{1}{4} - \frac{1}{4} = \frac{1}{2}.$$

Tehát a keresett terület $m(E) = 1 - \frac{1}{2} \cdot \frac{1}{2} = 1 - \frac{1}{4} = \frac{3}{4}$.

Ennélfogva az E esemény valószínűsége

$$P(E) = \frac{m(E)}{m(\Omega)} = \frac{\frac{3}{4}}{1} = \frac{3}{4}.$$

Ez a valószínűség egyezik a szimulációnál kapott relatív gyakoriságok "határértékével" is!

2.8. Valószínűségi mezők konstrukciója. Geometriai, Poisson, egyenletes és exponenciális eloszlások

A valószínűségi mezők konstrukciójához meg kell adni az **eseményteret**, az események σ -**algebráját** és az események **valószínűségeit**. Az eseménytérnél figyelembe kell venni a véletlen kísérlet lehetséges kimeneteleit. Az események σ -algebráját a feladatban szereplő események segítségével kell meghatározni úgy, hogy a valószínűségi számítások egyszerűek legyenek. Meg kell vizsgálni, hogy a kapott valószínűségi mező klasszikus mező-e!

2.8.1. PÉLDA. Pénzérme feldobás valószínűségi mezője

A legegyszerűbb véletlen kísérlet, a pénzfeldobás kísérlete, melynek kimeneteleit megadja az $\Omega = \{0, 1\}$ halmaz. A fej dobást jelölje 1 és az írást a 0 . Az események σ -algebrája az

$$S = \{\Omega \text{ összes részhalmaza}\} = 2^\Omega = \{\emptyset, \{0\}, \{1\}, \{0, 1\}\}$$

hatványhalmaz. Ha a fej dobás valószínűsége $p = P(\{1\})$ ($0 < p < 1$), akkor az írás dobás valószínűsége $P(\{0\}) = 1 - p = q$ kell, hogy legyen a komplementer esemény 2.5.1. tétele

alapján. Az axiómák egyértelműen meghatározzák a többi esemény valószínűségét, mert

$$P(\emptyset)=0, P(\{0, 1\}) = 1.$$

Ez a valószínűségi mező pontosan akkor **klasszikus**, ha a {fej} és {írás} elemi események egyformán $p = q = \frac{1}{2}$ valószínűek. Ez jelzi azt, hogy a pénzérme szabályos. Tehát szabályos pénzérme esetén klasszikus valószínűségi mezőt kapunk. Nem szabályos érme valószínűségi mezője nem klasszikus.

2.8.2. Szabályos kocka egyszeri feldobásának klasszikus valószínűségi mezője

A kockadobás eseménytere az $\Omega = \{1, 2, 3, 4, 5, 6\}$. Ha feltesszük, hogy a kocka szabályos, akkor ez azzal egyenértékű, hogy mindegyik oldalára ugyanolyan valószínűséggel esik

$$P(\{1\}) = P(\{2\}) = P(\{3\}) = P(\{4\}) = P(\{5\}) = P(\{6\}) = \frac{1}{6}.$$

Tehát a valószínűségi mező **klasszikus**. Az események σ -algebráját, mint az 1.8.8. példánál láttuk a keresett esemény valószínűségéhez célszerű igazítani. Ha minden elemi eseményt mérni szeretnénk, akkor σ -algebrának a legnagyobb elemszámú $S = 2^\Omega$ hatványhalmazt kell választani, melynek $2^6 = 64$ eleme van. Ha a páros dobás

$$\text{páros dobás} = \{2, 4, 6\}$$

eseményének valószínűségét keressük, akkor a

$$P(\text{páros dobás}) = \frac{\text{kedvező esetek száma}}{\text{összes esetek száma}} = \frac{\#\{2, 4, 6\}}{\#\{1, 2, 3, 4, 5, 6\}} = \frac{3}{6} = \frac{1}{2}$$

formulával számolhatunk, ahol a '#' (andrás kereszt vagy hashmark) jel arra utal, hogy a mögöttes szereplő halmaz elemeinek számát kell venni.

2.8.3. Ismételt pénzfeldobás véges sokszor. Szorzat-mező.

Dobjuk fel a pénzérmét egymástól függetlenül n -szer (az $n = 2, 3, 4, \dots$ lehet). Egyszeri feldobásnál a 2.8.1. pontban megadott (Ω, S, P) valószínűségi mezőt kapjuk:

$$\Omega = \{0, 1\}, S = 2^\Omega, P(1) = p, P(0) = 1 - p = q.$$

Az n feldobás eseménytere az $\Omega = \{0, 1\}$ halmaz Descartes-szorzata n -szer önmagával

$$\Omega^n = \Omega \times \Omega \times \dots \times \Omega = \left\{ \omega = (\omega_1, \omega_2, \dots, \omega_n) \mid \omega_i = 0 \text{ vagy } 1 \right\}$$

és az események S^n σ -algebrája az Ω^n összes részhalmazainak halmaza.

A P^n valószínűségi mértéket definiáljuk az elemi eseményekre a

$$P^n(\{\omega\}) = p^{\sum_{i=1}^n \omega_i} \cdot (1-p)^{n - \sum_{i=1}^n \omega_i}$$

képlettel, ahol $\omega = (\omega_1, \omega_2, \dots, \omega_n)$ egy n -elemű 0 és 1 számokból álló vektor és $n_\omega = \sum_{i=1}^n \omega_i$ az ω

vektorban szereplő 1 számok összege. Az így kapott (Ω^n, S^n, P^n) valószínűségi mező az n -szeres független pénzfeldobás kísérletének valószínűségi mezője. Bármely $A \in S^n$ esemény valószínűségét a

$$P^n(A) = \sum_{\omega \in A} p^{\sum_{i=1}^n \omega_i} \cdot (1-p)^{n - \sum_{i=1}^n \omega_i}$$

az (A.3.3) additivitási tulajdonságnak megfelelően értelmezzük.

Nézzük, hogyan lehet $n = 2$ esetén a konstrukciót megadni. Ekkor az eseménytér

$$\Omega^2 = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$$

az összes rendezett 0 és 1 értékeket tartalmazó párok halmaza. Az S^2 halmaz megadását az olvasóra bízunk, megjegyezve, hogy ez Ω^2 összes részhalmazainak halmaza, melynek $2^4 = 16$ eleme van. Az elemi események valószínűségeit az alábbi képletek adják meg

$$P(\{(0, 0)\}) = q^2, P(\{(0, 1)\}) = q \cdot p, P(\{(1, 0)\}) = p \cdot q, P(\{(1, 1)\}) = p^2.$$

Az összes elemi esemény valószínűségének összege 1, mert

$$q^2 + q \cdot p + p \cdot q + p^2 = (p + q)^2 = (p + 1 - p)^2 = 1.$$

Ezzel igazoltuk, hogy $P^2(\Omega^2) = 1$ az eseménytér valószínűsége 1. Tehát (Ω^2, S^2, P^2) egy

valószínűségi mező. Látható itt is, hogy a kapott mező, akkor klasszikus, ha $p = q = \frac{1}{2}$ vagyis szabályos érme esetén.

2.8.4. Pénzfeldobás végtelen sokszor (1.5.5. példa folytatása). Geometriai eloszlás.

Tekintsük a pénz ismételt feldobásának kísérletét, amelyet addig folytatunk függetlenül egymástól, amíg fej kimenetelt nem kapunk. Az eseménytér megszámlálhatóan végtelen lesz

$$\Omega = \{F, IF, IIF, IIIF, \dots\}.$$

Milyen valószínűségeket rendeljünk az egyes elemi eseményekhez, ha egyszeri alkalommal a fej dobás valószínűsége $p = P(F)$ és az írás dobás valószínűsége $P(I) = q = 1 - p$. A dobások **függetlensége** miatt (lásd az 5. fejezetben) a szorzat esemény valószínűsége egyenlő a tényezők valószínűségeinek szorzatával $P(IF) = P(I) \cdot P(F) = q \cdot p$. Hasonlóan kapjuk a

$$P(IIF) = P(I) \cdot P(I) \cdot P(F) = q^2 \cdot p$$

valószínűséget. Általában $(n - 1)$ írás dobás és az n -ik dobásra fej esemény valószínűségére

$$P(II \dots IF) = q^{n-1} \cdot p. \quad (n = 1, 2, 3, \dots)$$

Mutassuk meg, hogy az így értelmezett elemi eseményekkel az Ω eseménytér valószínűsége 1 lesz

$$P(\Omega) = P(F + IF + IIF + \dots) = P(F) + P(IF) + P(IIF) + \dots = p + p \cdot q + p \cdot q^2 + \dots = \sum_{n=1}^{\infty} p \cdot q^{n-1}.$$

A kapott végtelen összeg egy **mértani sor**, melynek általános alakjára és összegképletére emlékeztetünk

$$a + a \cdot q + a \cdot q^2 + \dots = \frac{a}{1 - q}, \text{ ha } -1 < q < 1.$$

Mivel $a = p$ a sor első tagja és kvóciense $0 < q < 1$, ezért összege

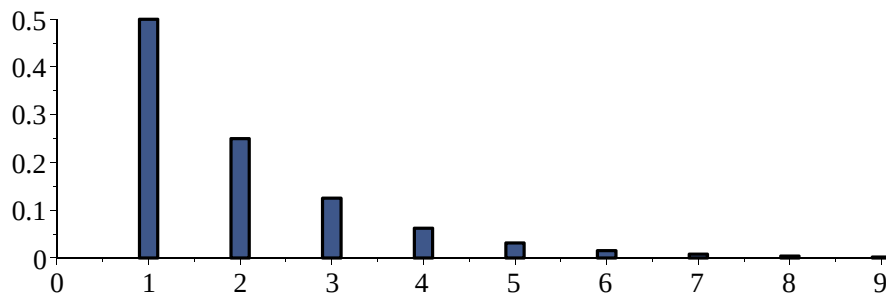
$$\sum_{n=1}^{\infty} p \cdot q^{n-1} = \frac{p}{1 - q} = \frac{p}{1 - (1 - p)} = \frac{p}{p} = 1.$$

Tehát a végtelen eseménytér elemi eseményeihez ilyen módon rendelt valószínűségi értékek olyanok, hogy a teljes eseménytér valószínűsége 1. Ezért a (Ω, S, P) hármas a fenti Ω eseménytérrel, valamint az $S = 2^\Omega$ hatványhalmazzal és az elemi eseményeken definiált P valószínűségi mértékkel egy valószínűségi mezőt alkot. Későbbiekben ezt valószínűségi mezőt a **geometriai eloszlás** modelljének nevezzük.

Az alábbi grafikon **szabályos érme** esetén mutatja, hogyan csökken az esélye annak, hogy a pénzérme n -edik feldobására kapjunk **először fej értéket**!

> seq $\left(\frac{1}{2^n}, n = 1..9\right)$; Statistics[ColumnGraph]([%], offset = 0.9, width = 0.2, distance = 0.8, caption = Fej dobás első előfordulásának valószínűsége szabályos érme esetén)

$$\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{32}, \frac{1}{64}, \frac{1}{128}, \frac{1}{256}, \frac{1}{512}$$



Fej dobás első előfordulásának valószínűsége szabályos érme esetén

2.8.5. Poisson-valószínűségi mező. Nem klasszikus valószínűségi mező.

Poisson - eloszlású modelleket sikeresen lehet alkalmazni egy **várakozási sorba** érkezők számának vizsgálatára, adott hosszúságú független időintervallumokat vizsgálva. Ilyen várakozó sor alakulhat ki például a számítógépes feladatok kiszolgálásánál is, mint például az internetes kérések vagy a nyomtatási sorok stb. Az (Ω, S, P) valószínűségi mező ebben az esetben diszkrét és végtelen, amelyre $\Omega = \{0, 1, 2, 3, 4, \dots\}$ a várakozó sorba érkezők száma az adott időtartam alatt. Az események σ -algebrája Ω összes részhalmazának halmaza. Az elemi események valószínűségét a

$$P(\omega) = \frac{\lambda^\omega}{\omega!} \cdot e^{-\lambda}, \omega = 0, 1, 2, 3, \dots$$

képlettel adjuk meg, ahol $\omega! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot \omega$ (ejtsd omega faktoriális) a természetes számok szorzata 1-től ω -ig. A modell paramétere a $\lambda > 0$, amelyet úgy választunk meg, hogy a valószínűségek jól illeszkedjenek a megfigyelt adatokhoz.

Annak bizonyítása, hogy az Ω eseménytér valószínűsége 1 az e^λ függvény Taylor-sorfejtésén múlik

$$\begin{aligned} P(\Omega) &= P(0) + P(1) + \dots + P(n) + \dots = e^{-\lambda} + \lambda \cdot e^{-\lambda} + \frac{\lambda^2}{2!} \cdot e^{-\lambda} + \dots + \frac{\lambda^n}{n!} \cdot e^{-\lambda} + \dots = \\ &= e^{-\lambda} \cdot \left(1 + \lambda + \frac{\lambda^2}{2!} + \dots + \frac{\lambda^n}{n!} + \dots \right) = e^{-\lambda} \cdot e^\lambda = 1. \end{aligned}$$

A kapott valószínűségi mező nemcsak azért nem klasszikus típusú, mert az eseménytér nem véges, hanem azért is, mert az elemi események nem egyformán valószínűek.

A Maple Statistics csomagja tartalmazza a Poisson-eloszlást *Poisson* - néven. Hozzunk létre egy X Poisson-eloszlású változót.

```
> restart : with(Statistics) :
X := Distribution(Poisson(lambda))
X := module( )
option Distribution, Discrete;
export Conditions, ParentName, Parameters, CDF, CharacteristicFunction, CGF,
Kurtosis, Mean, MGF, ProbabilityFunction, Skewness, Support, Variance,
```

(1.8.1)

VariationCoefficient, CDFNumeric, QuantileNumeric, RandomSampleMethod, RandomSampleSetup, RandomVariate, MaximumLikelihoodEstimate;

end module

Az X nevű Maple változó egy *modul*. A modulhoz tartoznak eljárások, amelyek az eloszlás tulajdonságait adják meg. Így a *Parameters* eljárás megmondja az eloszlás λ paraméterének nevét. A *Conditions* eljárás megadja a $\lambda > 0$ feltételt, amit a paraméterről tudni kell. A

ProbabilityFunction(k) eljárás megadja az $\{X = k\}$ esemény $P(\{X = k\})$ valószínűségére a $\frac{\lambda^k}{k!} \cdot e^{-\lambda}$ képletet.

> X :-Parameters, X :-Conditions, X :-ProbabilityFunction(k)

$$[\lambda], [0 < \lambda], \begin{cases} 0 & k < 0 \\ \frac{\lambda^k e^{-\lambda}}{k!} & \text{otherwise} \end{cases} \quad (1.8.2)$$

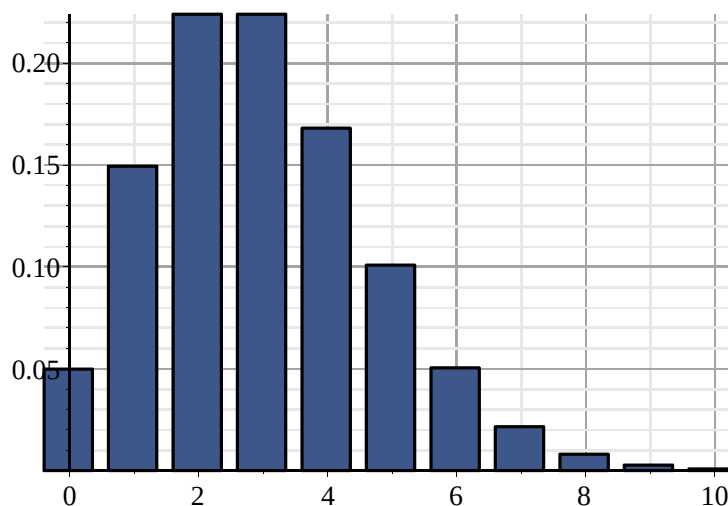
Számoljuk ki és rajzoljuk fel az *Poisson*-eloszlású X változó $P(X = k)$ valószínűségeit $k = 0, 1, \dots, 10$ értékekre, ha a paraméter $\lambda = 3$.

> $X := \text{Distribution}(\text{Poisson}(3))$:

> valószínűségek := [seq(evalf(X :-ProbabilityFunction(k)), $k = 0 \dots 10$)];

ColumnGraph(valószínűségek, offset = -0.4, gridlines = true)

valószínűségek := [0.04978706837, 0.1493612051, 0.2240418077, 0.2240418077, 0.1680313557, 0.1008188134, 0.05040940672, 0.02160403145, 0.008101511796, 0.002700503932, 0.0008101511796]



Ha összevetjük ezt az ábrát a 2.3.4. pontban Rutherford-alfa részek kisugárzására kapott empirikus eloszlásával, akkor a hasonlóság szembeötlő. Ezért felvetődik a kérdés, hogy vajon a

radióaktív sugárzásnál a részecskék száma Poisson-eloszlású modellt követ-e valamilyen λ -paraméterrel? A válaszra várni kell a statisztikai tesztek vizsgálatáig.

2.8.6. Folytonos egyenletes eloszlás valószínűségi mezője. Sűrűségfüggvény és eloszlásfüggvény.

Legyen $\Omega = [a, b]$ egy folytonos eseménytér! Vizsgálni szeretnénk, hogy milyen esetben lesz az intervallum "klasszikus" valószínűségi mező eseménytere. Ha minden $x \in [a, b]$ kimenetel **valószínűsége ugyanakkora**, akkor ezek valószínűsége csak 0 lehet, azaz $P(\{x\}) = 0$. Ha ugyanis $P(\{x\}) = p > 0$ pozitív valószínűségű lenne minden $x \in [a, b]$ esetén, akkor az (A.3.3) σ -additivitási feltétel miatt

$$P\left(\bigcup_{n=1}^{\infty} \{x_n\}\right) = \sum_{n=1}^{\infty} P(\{x_n\}) = p + p + \dots = \infty$$

Tehát sérülne az (A.3.1) axióma.

Az (Ω, S, P) folytonos egyenletes eloszlású valószínűségi mezőt úgy célszerű definiálni, hogy tetszőleges $[c, d] \subseteq [a, b]$ intervallum valószínűsége

$$P([c, d]) = \frac{m([c, d])}{m([a, b])} = \frac{(d - c)}{(b - a)} = \int_c^d \frac{1}{b - a} dx = \int_c^d f(x) dx.$$

Itt az $f(x) = \frac{1}{b - a}$, ha $x \in [a, b]$ függvényt az $[a, b]$ intervallumon egyenletes eloszlás

sűrűség függvényének nevezzük. Látható, hogy a teljes $\Omega = [a, b]$ intervallum valószínűsége 1

$$P([a, b]) = \frac{(b - a)}{(b - a)} = 1.$$

Így az előbb említett $P(\{x\}) = \int_x^x f(t) dt = 0$ feltétel is teljesül minden $x \in [a, b]$ esetén.

Az $[a, x]$ intervallum valószínűségére kapunk egy $F(x)$ függvényt, amelyet az egyenletes eloszlás **eloszlásfüggvényének** nevezünk és képlete

$$F(x) = P([a, x]) = \int_a^x \frac{1}{b - a} dt = \left[\frac{1}{b - a} \cdot t \right]_{t=a}^{t=x} = \frac{x - a}{b - a}, \text{ ha } a \leq x \leq b.$$

A differenciál-és integrálszámítás alaptétele jelenik meg az alábbi deriválási formulában

$$F'(x) = \frac{d}{dx} F(x) = \frac{1}{b - a} = f(x).$$

Tehát az eloszlásfüggvény deriváltja a sűrűségfüggvény.

A **folytonos egyenletes eloszlás** a *Statistics* csomagban *Distribution(Uniform(a, b))* néven beépített modul. Hívjuk meg speciálisan $a = 0$ és $b = 1$ paraméterekkel.

```
> restart : with(Statistics) :
U := Distribution(Uniform(0, 1))
U := module( )
option Distribution, Continuous;
export Conditions, ParentName, Parameters, CDF, InverseSurvivalFunction, Mean,
Median, MGF, Mode, PDF, Quantile, Support, Variance, RandomSample,
RandomSampleSetup, RandomVariate, MaximumLikelihoodEstimate;
```

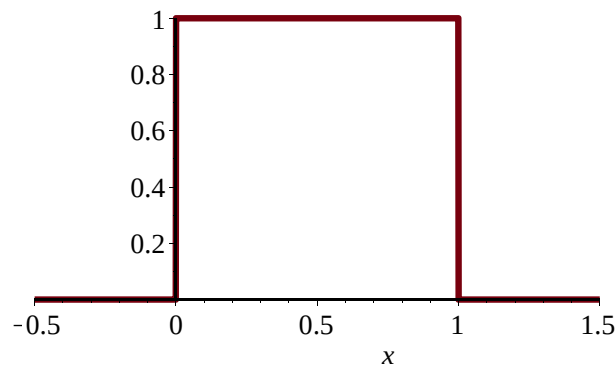
(1.8.3)

end module

Kaptunk az U változóban egy *modul* – t, amely tartalmazza a felsorolt nevű **eljárásokat**. Így többek között a $PDF(U, x)$ eljárás megadja az U egyenletes eloszlás **sűrűségfüggvényének** (angolul *Probability Density Function*) értékét az x helyen.

```
> f := x → PDF(U, x) :  
'f(x)' = f(x); plot(f(x), x = -0.5..1.5, thickness = 3, scaling = constrained, caption  
= A [0,1] intervallumon egyenletes eloszlás sűrűség függvénye)
```

$$f(x) = \begin{cases} 0 & x < 0 \\ 1 & 0 \leq x < 1 \\ 0 & \text{otherwise} \end{cases}$$

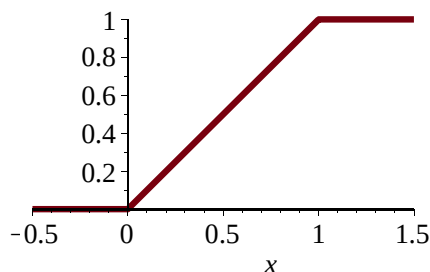


A [0,1] intervallumon egyenletes eloszlás sűrűség függvénye

A $CDF(U, x)$ eljárás megadja az U egyenletes eloszlás **eloszlásfüggvényének** (angolul *Cumulative Distribution Function*) értékét az x helyen.

```
> F := x → CDF(U, x) : 'F'(x) = F(x);  
plot(F(x), x = -0.5..1.5, thickness = 3, scaling = constrained, caption  
= A [0,1] intervallumon egyenletes eloszlás eloszlás függvénye)
```

$$F(x) = \begin{cases} 0 & x < 0 \\ x & 0 \leq x < 1 \\ 1 & \text{otherwise} \end{cases}$$



$A [0,1]$ intervallumon egyenletes eloszlás eloszlás függvénye

A $[0,1]$ intervallumon U egyenletes eloszlásra n **véletlen értéket generálni** a $\text{Sample}(U, n)$ hívással lehet.

```
> Sample(U, 10)
[0.814723686393179, 0.905791937075619, 0.126986816293506, 0.913375856139019,
0.632359246225410, 0.0975404049994095, 0.278498218867048, 0.546881519204984,
0.957506835434298, 0.964888535199277]
```

(1.8.4)

Végül az S σ -algebra az egyenletes eloszláshoz a $\mathcal{B}([a, b])$ -vel jelölt $[a, b]$ intervallum Borel-halmazainak rendszere, amely az összes $[c, d] \subseteq [a, b]$ részintervallumok által generált legszűkebb σ -algebra. (Lásd a 2.9. pontot!) Tetszőleges $B \in \mathcal{B}([a, b])$ Borel-halmaz valószínűségét a

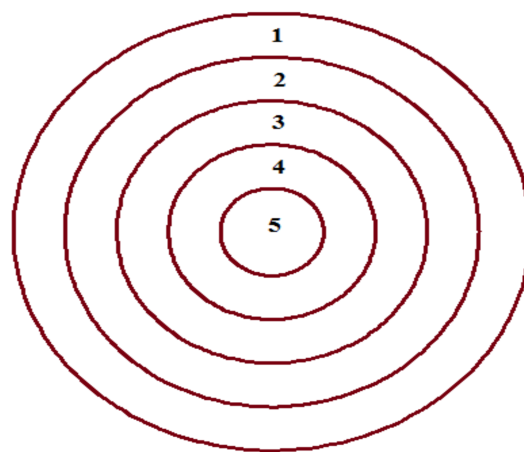
$$P(B) = \int_B f(x) \, dx = \int_B \frac{1}{b-a} \, dx = \frac{m(B)}{b-a}$$

(Lebesgue) integrállal számolhatjuk ki.

2.8.7. PÉLDA. Céltáblára lövés kísérletének nem klasszikus valószínűségi mezője

A darts játékban egy kör alakú táblára hegyes nyilakat dobunk és a részekre osztott táblán eltalált résznek megfelelő pontszám jár a játékosnak. Egy tartomány belsejét annál nehezebb eltalálni, minél kisebb a területe. Ezért a találat valószínűsége arányos a tartomány területével. Az egyszerűbb tárgyalás érdekében tegyük fel, hogy a céltábla egy r sugarú kör lemez, amelyet $n = 5$ koncentrikus körrel osztunk fel úgy, hogy a sugarak közötti **lépték** mindegyike $h = \frac{r}{5}$. A legkülső körgyűrű tartomány eltalálása esetén a játékos 1 pontot kap és befelé haladva 2, 3, 4 és 5 pont jár. (lásd a 2.8.1. ábrát)

FELADAT. Adjuk meg az i pontszám elérésének valószínűségét $i = 1, 2, 3, 4, 5$ esetén!



2.8.1. ábra. A céltábla és a pontszámok

Megoldás

Feltétel szerint az $A_i = \{a \text{ találat } i \text{ pontot ér}\}$ esemény valószínűségére

$$P(A_i) = \frac{\text{az } i - \text{ik tartomány területe}}{\text{a céltábla területe}} \quad (i = 1, 2, 3, 4, 5)$$

területek arányát kell alkalmazni. Látszólag a valószínűségek számítása a folytonos valószínűségi mező számítási szabálya alapján történik, azonban az események csak diszkrét értékeket vehetnek fel.

Így

$$P(A_1) = \frac{r^2 \cdot \pi - \left(\frac{4}{5} \cdot r\right)^2 \cdot \pi}{r^2 \cdot \pi} = 1 - \left(\frac{4}{5}\right)^2 = \frac{9}{25}.$$

Általánosan

$$P(A_i) = \frac{\left(\frac{6-i}{5} \cdot r\right)^2 \cdot \pi - \left(\frac{5-i}{5} \cdot r\right)^2 \cdot \pi}{r^2 \cdot \pi} = \frac{(6-i)^2 - (5-i)^2}{5^2}$$

minden $i = 1, 2, 3, 4, 5$ esetén. Tehát a valószínűségek függetlenek a tábla r sugarától és π értékétől is.

Nézzük a találatok valószínűségeinek Maple számításait!

> restart :

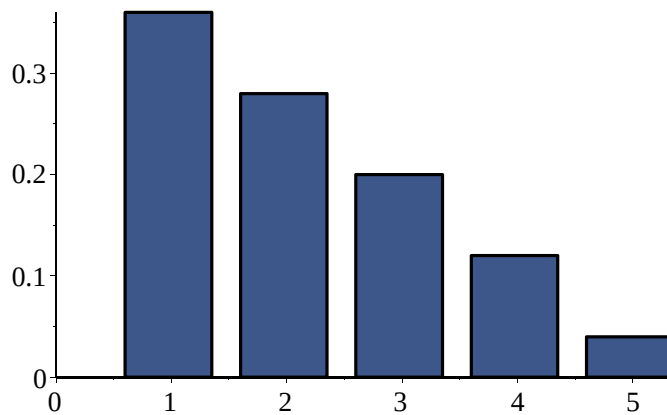
```
találatok := [ seq( ( (6-i)^2 - (5-i)^2 ) / 5^2, i = 1..5 ) ];
```

```
evalf(%, 4);
```

```
Statistics[ColumnGraph](találatok, offset = 0.6, caption  
= A pontszámok elérésének valószínűségei)
```

```
találatok := [ 9/25, 7/25, 1/5, 3/25, 1/25 ]
```

```
[0.3600000000, 0.2800000000, 0.2000000000, 0.1200000000, 0.0400000000]
```



A pontszámok elérésének valószínűségei

Látható, hogy a valószínűségek sorozata szigorúan monoton csökkenő, amely egyezik a tapasztalattal. Tehát legnehezebb eltalálni a céltábla közepét, melynek valószínűsége $0.04 = \frac{1}{25}$.

Ezért, ha egyenletes eloszlással dobáljuk a nyílhegyet a táblára, akkor átlagosan minden 25 dobásból 1 talál a közepébe.

Mutassuk meg, hogy az 5 valószínűség összege 1.

$$\begin{aligned} &> \sum_{i=1}^5 'találatok'_i = \sum_{i=1}^5 találatok_i \\ &\qquad\qquad\qquad \sum_{i=1}^5 találatok_i = 1 \end{aligned} \qquad (1.8.5)$$

Ez azt jelenti, hogy kaptunk egy diszkrét (Ω, S, P) valószínűségi mezőt, ahol az eseménytér

$$\Omega = \{1, 2, 3, 4, 5\}$$

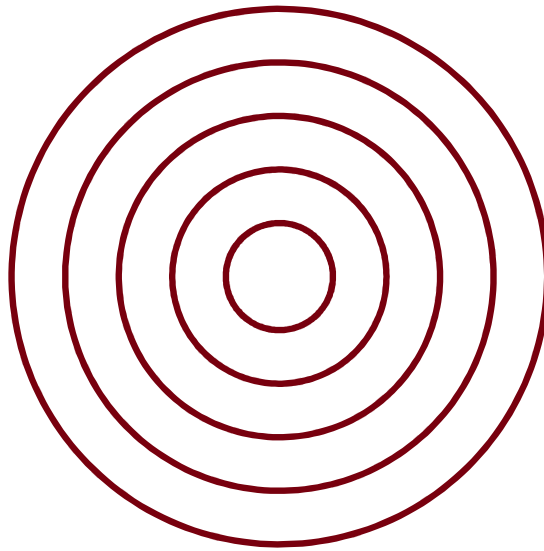
a dobható pontszámok. Az S események σ -algebrája az Ω összes részhalmazainak halmaza

$S = 2^\Omega$. Az elemi események a fent megadott $A_i = \{i\} = \{a \text{ találat } i \text{ pontot ér}\}$ események. Az elemi események valószínűségeit számoltuk ki az előzőekben és megmutattuk, hogy $P(\Omega) = 1$.

A kapott valószínűségi mező egy újabb példa **nem klasszikus valószínűségi mezőre**. Ugyanis itt az egyes elemi események kimenetelei **nem egyformán valószínűek**.

Rajzoljuk fel a céltábla koncentrikus köreit Maple-ben!

```
> restart
> with(plots) :
> n := 5 :
> korok := NULL :
  for k from 1 to n do
    korok := korok, plot( [ k*cos(t), k*sin(t), t = 0..2*pi ] ) :
  end do:
display( [ korok ], tickmarks = [ 2*n, 2*n ], thickness = 3, axes = none, scaling = constrained)
```

2.8.8. Valószínűségi mező exponenciális-eloszláshoz. Exponenciális eloszlás előállítása egyenletessel.

Exponenciális - eloszlású modell alkalmazható egy **várakozási sorban** az egymás után érkezők között eltelt időtartam hosszának vizsgálatára. Ebben az esetben az (Ω, S, P) valószínűségi mező **folytonos** és az eseménytér $\Omega = [0, \infty)$ az egymás után a várakozási sorba érkezők között eltelt lehetséges **időtartam**. Az eseménytérben elhelyezkedő tetszőleges $[a, b]$ intervallum valószínűségét az

$$P([a, b]) = \int_a^b \lambda \cdot e^{-\lambda \cdot x} dx \quad (0 \leq a < b < \infty)$$

integrállal értelmezzük. Itt az $f(x) = \lambda \cdot e^{-\lambda \cdot x}$ függvényt az exponenciális eloszlás **sűrűség függvényének** nevezzük és $\lambda > 0$ a paramétere. Ekkor a $[0, \infty)$ teljes eseménytér valószínűsége valóban 1 lesz

$$P(\Omega) = P([0, \infty)) = \int_0^{\infty} \lambda \cdot e^{-\lambda \cdot x} dx = \lim_{t \rightarrow \infty} [-e^{-\lambda \cdot x}]_{x=0}^{x=t} = \lim_{t \rightarrow \infty} (-e^{-\lambda \cdot t} + 1) = 1,$$

mert $\lambda > 0$. Az **integrál additivitása** miatt a valószínűség is additív. Ugyanis, ha $[a, b]$ és $[c, d]$ intervallumok részei a $[0, \infty)$ félegyenesnek és metszetük üres $[a, b] \cap [c, d] = \emptyset$, akkor

$$P([a, b] \cup [c, d]) = \int_{[a, b] \cup [c, d]} \lambda \cdot e^{-\lambda \cdot x} dx = \int_a^b \lambda \cdot e^{-\lambda \cdot x} dx + \int_c^d \lambda \cdot e^{-\lambda \cdot x} dx = P([a, b]) + P([c, d]).$$

Az $S = \mathcal{B}([0, \infty))$ σ -algebra ebben az esetben az összes $[a, b] \subset [0, \infty)$ részintervallum által generált Borel-féle σ -algebra. Tetszőleges $B \in \mathcal{B}([0, \infty))$ Borel-halmaz valószínűségét a

$$P(B) = \int_B f(x) dx = \int_B \lambda \cdot e^{-\lambda \cdot x} dx$$

(Lebesgue) integrállal számolhatjuk ki. Ekkor az integrál σ -additív tulajdonsága miatt teljesül az

(A.3.3) axióma.

A $[0, x]$ intervallum valószínűségére kapunk egy $F(x)$ függvényt, amelyet az exponenciális eloszlás **eloszlásfüggvényének** nevezünk

$$F(x) = P([0, x]) = \int_0^x \lambda \cdot e^{-\lambda \cdot t} dt = [-e^{-\lambda \cdot t}]_{t=0}^{t=x} = 1 - e^{-\lambda \cdot x}, \text{ ha } x \geq 0.$$

A differenciál-és integrálszámítás alaptétele jelenik meg az alábbi deriválási formulában

$$F'(x) = \frac{d}{dx} F(x) = \lambda \cdot e^{-\lambda \cdot x} = f(x).$$

Tehát az eloszlásfüggvény deriváltja a sűrűségfüggvény.

Vigyázat! Az exponenciális eloszlás Maple-be beépített képlete **eltér** a fentiektől. Ugyanis a λ paraméter helyett az $\frac{1}{\lambda}$ paramétert kell az $Exponential(\lambda)$ eljárásba beírni.

```
> restart : with(Statistics) :
```

```
> E := Distribution(Exponential( $\lambda$ ))
```

```
E := module( )
```

(1.8.6)

```
option Distribution, Continuous;
```

```
export Conditions, ParentName, Parameters, CharacteristicFunction, CDF, CGF, Mean,  
Median, MGF, Mode, PDF, Quantile, Support, Variance, CDFNumeric, QuantileNumeric,  
RandomSample, RandomSampleSetup, RandomVariate, MaximumLikelihoodEstimate;
```

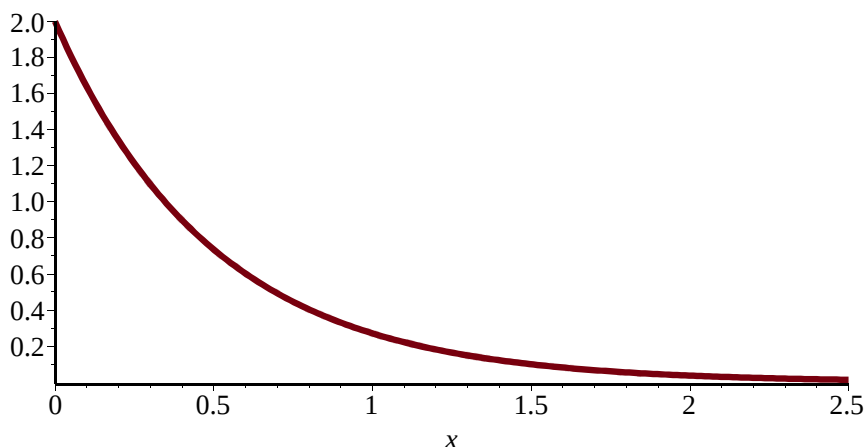
```
end module
```

```
>  $f_X := unapply(PDF(E, x), x) : f_X'(x) = f_X(x);$ 
```

```
plot(eval( $f_X(x)$ ),  $\lambda = 0.5$ ),  $x = 0 .. 2.5$ , thickness = 3, caption
```

```
= Az exponenciális eloszlás sűrűség függvénye  $\lambda=1/2$  mellett)
```

$$f_X(x) = \begin{cases} 0 & x < 0 \\ \frac{e^{-\frac{x}{\lambda}}}{\lambda} & \text{otherwise} \end{cases}$$



Az exponenciális eloszlás sűrűség függvénye $\lambda=1/2$ mellett

2.8.9. Exponenciális eloszlás generálása egyenletes eloszlásból

Eddig folytonos esetben mindig **egyenletes eloszlású** értékeket generáltunk. Kérdés, hogyan adhatunk meg egyenletes eloszlás felhasználásával **exponenciális eloszlást**?

Ahhoz, hogy megkülönböztessük az **egyenletes eloszlást** és az **exponenciális eloszlást**, ahhoz jelöléseket kell bevezetni.

Jelölje U a $[0, 1]$ intervallumon egyenletes eloszlást és ennek valószínűségi mezőjét az (Ω_1, S_1, P_1) . Erre az a jellemző, hogy $\Omega_1 = [0, 1]$ és az eloszlásfüggvénye

$$F_1(x) = P_1(0 \leq U \leq x) = x, \text{ minden } 0 \leq x < 1 \text{ esetén.}$$

Jelölje E a $\lambda > 0$ paraméterű exponenciális eloszlást és ennek valószínűségi mezőjét az (Ω_2, S_2, P_2) . Erre az a jellemző, hogy $\Omega_2 = [0, \infty)$ és az eloszlásfüggvénye

$$F_2(x) = P_2(0 \leq E \leq x) = 1 - e^{-\lambda \cdot x}, \text{ tetszőleges } x \geq 0 \text{ esetén.}$$

Mutassuk meg, hogy az

$$E = -\frac{1}{\lambda} \cdot \ln(U)$$

transzformációs képlettel értelmezett E eloszlás λ -paraméterű exponenciális eloszlású, amennyiben U egyenletes eloszlású a $[0, 1]$ intervallumban.

Nézzük E értékei milyen határok közé esnek. Mivel $0 < U \leq 1$, ezért $-\infty < \ln(U) \leq 0$. Tehát

$$0 \leq -\frac{1}{\lambda} \cdot \ln(U) = E.$$

Nézzük meg mi lesz E eloszlásfüggvénye tetszőleges $x \geq 0$ esetén

$$P_2(0 \leq E \leq x) = P_1\left(0 \leq -\frac{1}{\lambda} \cdot \ln(U) \leq x\right) = P_1\left(e^{-\lambda \cdot x} \leq U \leq 1\right) = 1 - P_1(0 \leq U \leq e^{-\lambda \cdot x}).$$

Mivel U egyenletes eloszlású a $[0, 1]$ intervallumon, ezért a jobb oldal

$$1 - P_1(0 \leq U \leq e^{-\lambda \cdot x}) = 1 - e^{-\lambda \cdot x} = F_2(x)$$

az exponenciális eloszlás eloszlásfüggvénye.

Próbáljuk ki a transzformációt $\lambda = 2$ paraméterrel!

> restart :

```
> with(Statistics) :
> U := RandomVariable(Uniform(0, 1))
> λ := 2; E := - 1/λ · ln(U)
```

$$U := _R$$

$$\lambda := 2$$

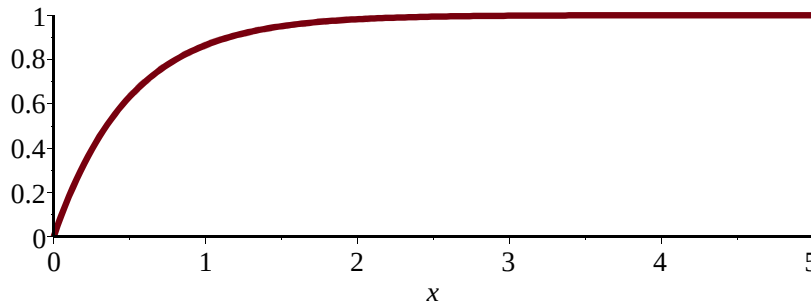
$$E := -\frac{1}{2} \ln(_R)$$

(1.8.7)

[A kapott E eloszlás eloszlásfüggvényét a *Statistics* csomag *CDF* eljárása meg tudja határozni!

```
> CDF(E, x)
> plot(CDF(E, x), x = 0 ..5, thickness = 3, caption
      = Exponenciális eloszlás, amelyet egyenletes eloszlásból kaptunk)
```

$$\begin{cases} 0 & x \leq 0 \\ -e^{-2x} + 1 & 0 < x \end{cases}$$

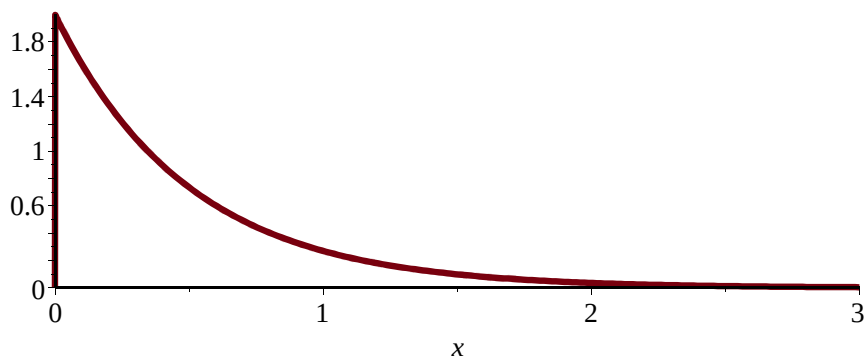


Exponenciális eloszlás, amelyet egyenletes eloszlásból kaptunk

[Az E eloszlás sűrűségfüggvényét az x helyen a *Statistics* csomag *PDF(E, x)* eljárása adja meg.

```
> PDF(E, x)
> suruseg := plot(PDF(E, x), x = 0 ..3, thickness = 3, caption
      = Exponenciális eloszlás, amelyet egyenletes eloszlásból kaptunk) : suruseg
```

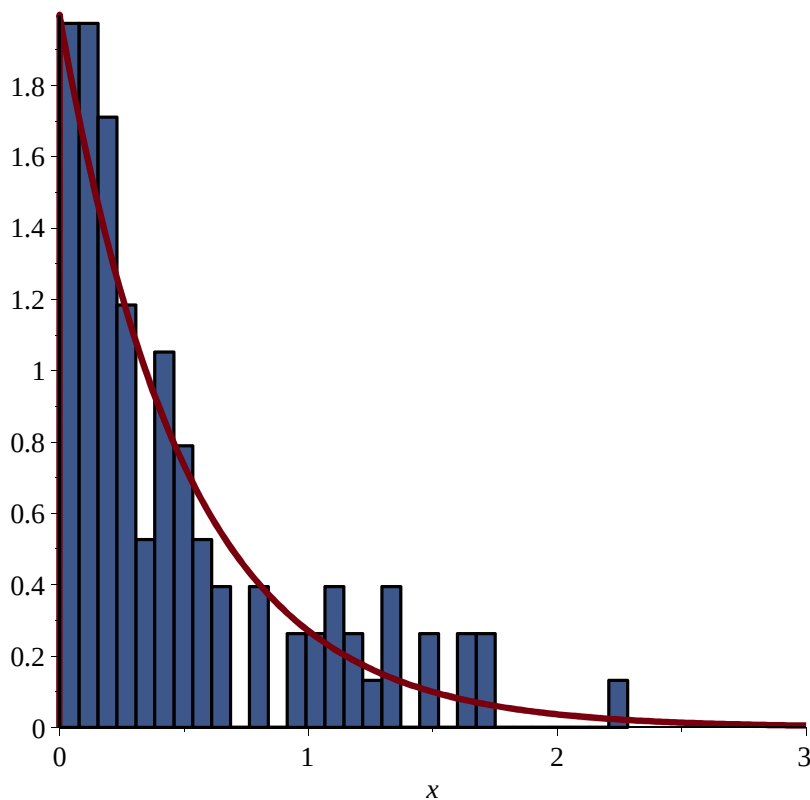
$$\begin{cases} 0 & x \leq 0 \\ 2 e^{-2x} & 0 < x \end{cases}$$



Exponenciális eloszlás, amelyet egyenletes eloszlásból kaptunk

Generáljunk 100 elemű mintát az U egyenletes eloszlásra, majd a minta minden elemére alkalmazzuk az $E : u \rightarrow -\frac{1}{\lambda} \cdot \ln(u)$ transzformációt. A kapott véletlen számok gyakoriság hisztogramját rajzoljuk fel és tegyük egybe a fenti sűrűségfüggvénnyel!

```
> randomize( ) : mintaU := Sample(U, 100) :  
  mintaE := map( - 1/λ · ln, mintaU ) :  
  plots[display]( [Histogram(mintaE), suruseg] )
```



Exponenciális eloszlás, amelyet egyenletes eloszlásból kaptunk

A transzformált minta adatok hisztogramjára az exponenciális eloszlás sűrűségfüggvénye jól illeszkedik. Tehát egy $[0, 1]$ intervallumon egyenletes U eloszlásra az $E = -\frac{1}{\lambda} \cdot \ln(U)$

transzformációt alkalmazva λ -paraméterű E exponenciális eloszláshoz jutunk. Ez a módszer azt sugallja, hogy ha egy számítógépes programba beépítjük a **véletlenszámok egyenletes generálását**, akkor ezek megfelelő függvényei segítségével más eloszlású adatsort is megkaphatunk.

2.8.10. PÉLDA.

Hogyan kapunk exponenciális eloszlást geometriai eloszlásból? Várakozás egy esemény bekövetkezésére.

A geometriai eloszlás és az exponenciális eloszlás sok tekintetben hasonlítanak egymásra. Az X geometriai eloszlásra

$$\lim_{k \rightarrow \infty} P(X = k) = \lim_{k \rightarrow \infty} p \cdot (1 - p)^k = 0 \quad (0 < p < 1)$$

szigorúan monoton csökkenve. Hasonlóan az Y exponenciális eloszlás $f_Y(X) = \lambda \cdot e^{-\lambda \cdot x}$ sűrűség függvényére

$$\lim_{x \rightarrow \infty} f_Y(x) = \lim_{x \rightarrow \infty} \lambda \cdot e^{-\lambda \cdot x} = 0 \quad (\lambda > 0)$$

és itt is a nullához tartás szigorúan monoton csökkenő.

Ha az X geometriai eloszlás $E(X) = \frac{1}{p}$ várható értéke és az Y exponenciális eloszlás $E(Y) = \frac{1}{\lambda}$ várható értéke egyenlő, akkor X eloszlása és Y sűrűségfüggvénye jól közelítik egymást. Az $E(Y) = E(X)$ akkor teljesül, ha $p = \lambda$. Mivel Maple-ben az exponenciális eloszlás λ paramétere helyett a paraméter $\frac{1}{\lambda}$ reciprokát kell megadni, ezért Maple-ben a $\lambda = \frac{1}{p}$ egyenlőséget kell használni.

```
> restart : with(Statistics) :
p := 0.2; X := Distribution(Geometric(p)) : fX := k → ProbabilityFunction(X, k) :
'fX'(k) = fX(k) ;
λ := 1/p; Y := Distribution(Exponential(λ)) : fY := x → PDF(Y, x) :
'fY'(x) = fY(x);
```

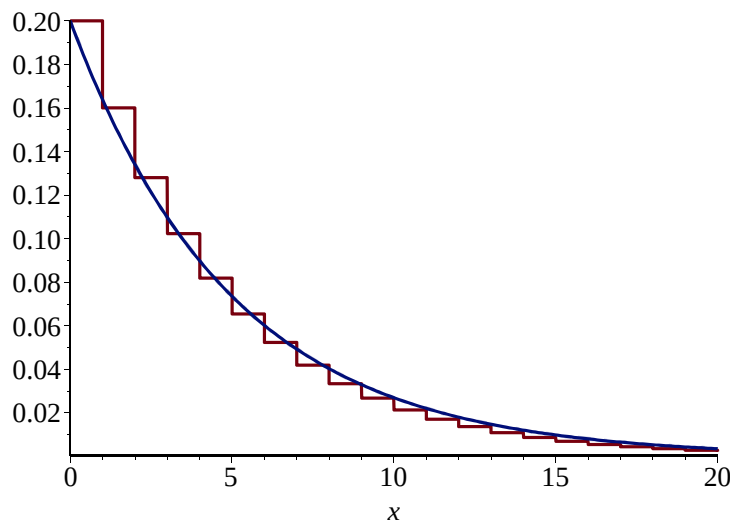
$$p := 0.2$$

$$fX(k) = \begin{cases} 0 & k < 0 \\ 0.2 \cdot 0.8^k & \text{otherwise} \end{cases}$$

$$\lambda := 5.000000000$$

$$fY(x) = \begin{cases} 0 & x < 0 \\ 0.2000000000 e^{-0.2000000000x} & \text{otherwise} \end{cases} \quad (1.8.8)$$

```
> plot( [ fX(floor(x)), fY(x) ], x = 0 .. 20)
```



Látható, hogy az exponenciális eloszlás sűrűség függvénye és a geometriai eloszlás jól illeszkednek egymáshoz. Diszkrét eloszlásból határátmenettel tudunk folytonos eloszlást gyártani. Nézzük a konstrukciót!

Dobjunk fel egy pénz érmét $\tau, 2 \cdot \tau, 3 \cdot \tau, \dots$ időpontokban és legyen $X^{(\tau)}$ a **várakozási idő**, ameddig az első fej érték be nem következik. Vagyis az első fej esemény bekövetkezésére várunk. Ekkor $\frac{1}{\tau} \cdot X^{(\tau)}$ eloszlása **geometriai**

$$P(X^{(\tau)} > k \cdot \tau) = (1 - p)^k,$$

mert

$$P(X^{(\tau)} > k \cdot \tau) = p \cdot (q^k + q^{k+1} + \dots) = p \cdot q^k \cdot (1 + q + q^2 + \dots) = p \cdot q^k \cdot \frac{1}{1 - q} = q^k$$

ahol p a fej dobás valószínűsége minden dobásnál és $q = 1 - p$. (ld. 2.8.4. pontot). Mivel a várható érték $E\left(\frac{X^{(\tau)}}{\tau}\right) = \frac{1}{p}$, ezért $E(X^{(\tau)}) = \frac{\tau}{p}$.

Rögzítsünk egy $t > 0$ időt tetszőlegesen és tartson τ értéke nullához $\tau \searrow 0$ csökkenőleg. Ha azt szeretnénk, hogy p értéke állandó maradjon ezen határátmenet mellett, akkor a

kísérletszámnak végtelenhez kell tartani a $[0, t]$ időtartam alatt, hiszen a kísérletszám a $\left\lfloor \frac{t}{\tau} \right\rfloor$ hányados alsó egész része. Annak érdekében, hogy kompenzáljuk a kísérletszám végtelenhez tartását a $[0, t]$ időintervallumon válasszuk $p = p^{(\tau)}$ értékét a $\frac{p^{(\tau)}}{\tau} = \lambda$ egyenlőségnek

megfelelően, ahol $\lambda > 0$ állandó. Ekkor

$$P(X^{(\tau)} > t) = (1 - p^{(\tau)})^{\left\lfloor \frac{t}{\tau} \right\rfloor} = (1 - \lambda \cdot \tau)^{\left\lfloor \frac{t}{\tau} \right\rfloor} = \left(1 - (\lambda \cdot t) \cdot \frac{\tau}{t}\right)^{\left\lfloor \frac{t}{\tau} \right\rfloor} \rightarrow e^{-\lambda \cdot t} \quad (\forall t > 0)$$

ha $\tau \rightarrow 0$. Tehát $X^{(\tau)}$ tart "eloszlásban" egy λ -paraméterű exponenciális eloszláshoz. (a definíciót lásd 13. fejezetben)

2.9.* Miért szükséges a σ -algebra? A Vitali-féle V halmaz.

Mi indokolja a σ -algebra fogalmának bevezetését a valószínűségelméletbe? Ez a kérdés első megfontolásra azért **nem meglepő**, mert a valószínűségi mező (Ω, S, P) rendezett hármasában szereplő S minden esetben lehetne az Ω eseménytér összes részhalmazából álló halmaz $S = 2^\Omega$. Véges vagy megszámlálhatóan végtelen Ω eseménytérre a 2^Ω hatványhalmaz valóban megfelelő választás az S eseményalgebrahoz. Kiderült azonban, hogy a valós számok bármely $\Omega = [a, b]$ részintervallumára az összes részhalmazok halmazán nem értelmezhető olyan P **valószínűségi mérték**, amely "használható" lenne. Ilyen komplikált részhalmaz konstrukció érdekében a $[0, 1]$ intervallumban egymástól racionális távolságra levő pontokat egy osztályba sorolunk és mindegyik osztályból kiválasztunk egy számot. Ez a konstrukció megfelelően bonyolult halmazt eredményez ahhoz, hogy kapjunk **nem mérhető halmazt**.

2.9.1. DEFINÍCIÓ. Egy ekvivalencia reláció a valós számok halmazán

Azt mondjuk, hogy az $x, y \in \mathbb{R}$ valós számok között fenáll az $x \sim y$ (ejtsd x hullámvonal y) **reláció** (vagy kapcsolat), ha az $x - y = r$ **különbség racionális szám**, vagyis $r \in \mathbb{Q}$.

2.9.2. TÉTEL. Az $x \sim y$ reláció ekvivalencia reláció

Mutassuk meg, hogy a 2.9.1. reláció rendelkezik az alábbi 3 nevezetes tulajdonsággal
(i) $x \sim x$, azaz **reflexív**

- (ii) ha $x \sim y$, akkor $y \sim x$, azaz **szimmetrikus**
 (iii) ha $x \sim y$ és $y \sim z$, akkor $x \sim z$, azaz **transzitiv**

BIZONYÍTÁS

Mivel $x - x = 0$ és a 0 racionális szám, ezért az $x \sim x$ reflexív (i) tulajdonság teljesül tetszőleges valós x számra.

Ha az $x - y = r$ különbség racionális, akkor az $y - x = -r$ is racionális szám. Ez igazolja az (ii) szimmetria tulajdonságot.

A transzitiv (iii) tulajdonság feltételei alapján az $x - y = r_1$ és az $y - z = r_2$ különbségek racionális számok. Meg kell mutatni, hogy ekkor az $x - z = r$ különbség is racionális. Ez következik abból, hogy

$$x - z = (x - y) + (y - z) = r_1 + r_2 = r$$

két racionális szám összege is racionális. ♦

Általában az (i) – (iii) tulajdonságokkal rendelkező valamely relációt *ekvivalencia relációnak* nevezzük. Minden ekvivalencia reláció az alaphalmaz egy **osztályozását** adja, amely olyan mint a teljes eseményrendszer. Tehát a valós számok halmazát diszjunkt osztályokra bontja fel a fenti reláció. Az egy osztályba sorolt valós számok egymással relációban vannak, azaz közöttük a különbség racionális szám.

Így például a $\left\{ \sqrt{2}, \sqrt{2} - 1, \sqrt{2} + \frac{1}{2}, \sqrt{2} - \frac{1}{3} \dots \right\}$ számok ugyanahhoz az osztályhoz

tartoznak. Az osztályozás során az összes racionális szám $\mathbb{Q} = \left\{ \frac{p}{q} \mid p, q \in \mathbb{Z}, q \neq 0 \right\}$ egyetlen osztályt alkot. Két különböző osztálynak nincs közös eleme, továbbá az összes osztály uniója kiadja a teljes valós számok $\mathbb{R}$ halmazát.

2.9.3. Vitali-féle V halmaz konstrukciója és alaptulajdonságai

A $[0,1]$ intervallumba eső valós számok alábbi részhalmazát Giuseppe Vitali (1875-1932) olasz matematikus konstruálta a 2.9.1. reláció segítségével.

Tekintsük a $[0, 1]$ intervallum osztályozását diszjunkt részhalmazokra, amelyet a 2.9.1. ekvivalencia reláció indukál. Vegyünk ki mindegyik osztályból **pontosan egy** számot és ezek összességéből képezzük a Vitali-féle V halmazt.

(a) A V halmaz létezését garantálja a halmazelmélet ún. *kiválasztási axiómája*.

(b) A Vitali halmaz a $[0, 1]$ intervallum részhalmaza: $V \subseteq [0, 1]$.

(c) A V halmaz bármely két u és v nem azonos elemének különbsége irracionális szám: $u - v \in \mathbb{Q}^* = \mathbb{R} \setminus \mathbb{Q}$. Ugyanis, ha $(u - v)$ racionális szám lenne, akkor ugyanahhoz az osztályhoz tartozna. Amely ellentmondana annak, hogy V minden osztályból pontosan egy elemet tartalmaz.

(d) Megmutatható, hogy a V halmaz számossága nem-megszámlálható.

(e) Mivel nincs utasítás arra, hogy az egy osztályba tartozó elemek közül melyiket válasszuk ki, ezért végtelen sok Vitali-féle halmazt konstruálhatunk a $[0, 1]$ intervallumban. Például a $[0, 1]$ intervallumhoz tartozó összes racionális szám $\left\{ 0, 1, \frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{3}{4}, \frac{1}{5}, \dots \right\}$ ugyanahhoz az osztályhoz tartozik. Ezért ezek közül bármelyik tartozhat a Vitali-halmazhoz, de csak egy, például lehet ez a 0. Ekkor más racionális szám már nem lehet eleme V -nek. Ha például az $\frac{1}{2}$ -t választottuk a V halmazba, akkor más racionális szám már nem tartozhat V -hez.

2.9.4. DEFINÍCIÓ. A Vitali halmaz eltoltjai a $[-1, 1]$ intervallumba eső racionális számokkal

Tekintsük a $[-1, 1]$ intervallumba eső összes racionális számok halmazát, vagyis a $\mathbb{Q} \cap [-1, 1]$ halmazt. Ez a halmaz megszámlálhatóan végtelen, ezért sorba lehet szedni elemeit. Például növekvő nevezők szerint az alábbi módon

$$q_0 = 0, q_1 = 1, q_2 = -1, q_3 = \frac{1}{2}, q_4 = -\frac{1}{2}, q_5 = \frac{1}{3}, q_6 = -\frac{1}{3}, q_7 = \frac{2}{3}, q_8 = -\frac{2}{3}, q_9 = \frac{1}{4}, \dots$$

Tehát $\mathbb{Q} \cap [-1, 1] = \{q_0, q_1, q_2, \dots\} = Q_1$ jelöli a $[-1, 1]$ intervallumba eső összes racionális számot.

Jelölje $V_k = V + q_k = \{v + q_k \mid v \in V\}$ a Vitali-féle V halmaz eltoltját a $[-1, 1]$ intervallumba eső valamely $q_k \in Q_1$ racionális számmal, ahol $k = 0, 1, 2, 3, \dots$ lehet.

A halmazok között ott van a Vitali-féle $V = V_0$ halmaz is.

2.9.5. ÁLLÍTÁS. A Vitali-halmaz eltoltjainak tulajdonságai

Az alábbi állítások teljesülnek a Vitali-féle halmaz V_k eltoltjaira

(i) a különböző eltoltak metszete üres: $V_k \cap V_j = \emptyset$, ha $k \neq j$;

(ii) az összes eltolt uniójára teljesülnek a

$$[0, 1] \subseteq A = \bigcup_{k=0}^{\infty} V_k \subseteq [-1, 2] \text{ tartalmazások;}$$

(iii) nem értelmezhető az eltolt V_k halmazok $m(V_k)$ hosszúság mértéke $k = 0, 1, 2, \dots$ esetén.

BIZONYÍTÁS.

Az (i) tulajdonság bizonyításához tegyük fel, hogy $x \in V_k \cap V_j$. Ekkor az eltoltak 2.9.4. definíciója szerint van olyan v és u elem a Vitali-féle V halmazban, melyekre

$$x = u + q_k = v + q_j$$

teljesül. De ekkor az $u - v = q_j - q_k$ különbség racionális szám lenne, amely a 2.9.3. (c) tulajdonság miatt lehetetlen. Ezzel igazoltuk, hogy nem lehet közös eleme két különböző V_k és V_j Vitali eltolt halmaznak ($k \neq j$).

Az (ii) tartalmazási tulajdonságok közül az $A = \bigcup_{k=0}^{\infty} V_k \subseteq [-1, 2]$ tartalmazás könnyen igazolható.

Ugyanis V_k tetszőleges x eleme $x = v + q_k$ alakú, ahol $v \in V$ tetszőleges és $q_k \in [-1, 1]$ rögzített.

A

$$0 \leq v \leq 1 \text{ és } -1 \leq q_k \leq 1,$$

egyenlőtlenségeket összeadva

$$-1 \leq x = v + q_k \leq 2.$$

Tehát mindegyik Vitali eltolt halmaz a $[-1, 2]$ intervallumban van, ennél fogva ez igaz az uniójukra is.

Mutassuk most meg a $[0, 1] \subseteq A = \bigcup_{k=0}^{\infty} V_k$ tartalmazást. Legyen x tetszőleges eleme a $[0, 1]$ intervallumnak. Ekkor vagy $x \in V = V_0$ teljesül és így $x \in A$ is következik. Vagy ha

$x \notin V = V_0$, akkor van olyan $v \in V$, amelyre $x \sim v$. Hiszen a 2.9.1. reláció osztályozás a $[0,$

$1]$ intervallumon, ezért x hozzátartozik valamelyik osztályhoz. Ez azt jelenti, hogy $r = x - v$ egy racionális szám és mivel x is és v is a $[0, 1]$ intervallumban van, ezért $r \in [-1, 1]$. A $q_1, q_2, q_3, \dots$ sorozat tartalmazza a $[-1, 1]$ intervallumba eső összes racionális számot, ezért r megegyezik valamelyik q_k ($k \geq 1$) racionális számmal: $r = q_k$. Ennél fogva

$$x = v + q_k \in V_k$$

és ezért teljesül az $x \in A = \bigcup_{i=0}^{\infty} V_i$ tartalmazás is. Ezzel igazoltuk, hogy az összes Vitali eltolt halmaz uniója lefedi a teljes $[0, 1]$ intervallumot.

Az (iii) igazolásához tegyük fel az állítással ellentétben, hogy a Vitali-féle V halmaznak van

hosszúság mértéke, amelyet jelöljön $m(V)$. A hosszúság egyrészt nem-negatív szám lehet, másrészt mivel V végtelen sok pontból áll, ezért elképzelhető, hogy a hosszúsága lehet végtelen: $0 \leq m(V) \leq \infty$.

Mivel mindegyik $V_k = V + q_k$ halmaz a V halmaz adott racionális számmal vett eltolta és a hosszúságot nem változtatja meg az eltolás, ezért az eltoltak mértéke megegyezik V mértékével:

$$m(V) = m(V_1) = m(V_2) = m(V_3) = \dots$$

Mint láttuk a $V, V_1, V_2, \dots$ halmazok metszetei páronként üresek, ezért az $\bigcup_{k=0}^{\infty} V_k$ unió teljes hossza megegyezik az egyes részek hosszának összegével (σ -additivitás)

$$(2.9.1) \quad m\left(\bigcup_{k=0}^{\infty} V_k\right) = m(V) + m(V_1) + m(V_2) + \dots = m(V) + m(V) + m(V) + \dots = \sum_{k=0}^{\infty} m(V).$$

A (2.9.1) egyenlet jobb oldalán végtelen sokszor összeadjuk ugyanazt a nem-negatív számot önmagával. Az egyenlőség csak úgy teljesülhet, ha $m\left(\bigcup_{k=0}^{\infty} V_k\right) = 0$ vagy $m\left(\bigcup_{k=0}^{\infty} V_k\right) = \infty$. Azonban a 2.9.5. állítás (ii) tartalmazási része miatt

$$m([0, 1]) = 1 \leq m\left(\bigcup_{k=0}^{\infty} V_k\right) \leq m([-1, 2]) = 3.$$

Hiszen a $[0, 1]$ intervallum hossza 1 és a $[-1, 2]$ intervallum hossza 3. Tehát a (2.9.1) egyenlet bal oldala nem lehet se nulla se végtelen. A kapott ellentmondás azt bizonyítja, hogy a Vitali-féle V halmazra és az eltoltsaira nem értelmezhető a hosszúság. ♦

2.9.6. Ha $\Omega = [-1, 2]$ az eseménytér, akkor az $S = 2^{\Omega}$ hatványhalmazon nincs valószínűségi mérték.

Tekintsük a következő (Ω, S, P) valószínűségi mezőt, melyben $\Omega = [-1, 2]$ intervallum, az S eseményalgebra a $[-1, 2]$ eseménytér összes részhalmazából álló halmaz. A $P: S \rightarrow [0, 1]$

valószínűségi mérték legyen az intervallumok hosszából származtatott $P(A) = \frac{m(A)}{m(\Omega)} = \frac{m(A)}{3}$

mérték, ahol $A \subseteq \Omega$ és $m(A)$ az A részhalmaz mértéke. Ha $A = [a, b] \subseteq [-1, 2]$ egy részintervallum, akkor ennek mértéke a hossza

$$m([a, b]) = b - a.$$

Ha megszámlálhatóan végtelen sok **egymásba nem nyúló** $[a_k, b_k]$ intervallumrendszert tekintünk a $[-1, 2]$ intervallumban, akkor ezek uniójának mértékét **természetes módon** az intervallumok összhosszával mérjük

$$m\left(\bigcup_{k=1}^{\infty} [a_k, b_k]\right) = \sum_{k=1}^{\infty} (b_k - a_k).$$

A felírt végtelen összeg konvergens, hiszen az összeg nem lehet nagyobb 3-nál, vagyis a $[-1, 2]$ intervallum hosszánál.

Azonban a $[-1, 2]$ intervallum összes részhalmazai halmazában ott van a 2.9.5. pontban

konstruált bonyolult $A = \bigcup_{k=0}^{\infty} V_k$ részhalmaz is. Láttuk azonban, hogy a fenti távolság mértéket nem

lehet úgy kiterjeszteni, hogy $m(A)$ mértéke létezzen. Ez azt mutatja, hogy egy intervallum összes részhalmaza olyan gazdag, hogy minden részhalmaz hosszúságát lehetetlen értelmezni. Ezt az ellentmondást feloldani úgy lehet, ha kizárjuk az S esemény algebrából a Vitali-féle konstrukciókat! Az építkezést azonban nem a rossz részhalmazok kizárásával kell kezdeni, hanem a jó részhalmazok felsorolásával. ♦

2.9.7. A σ -algebra tulajdonsággal kizárjuk a Vitali-féle konstrukciókat az eseményalgebrából.

Anélkül, hogy a halmazok Lebesgue-szerinti és Borel-szerinti mérhetőségébe részletesen belemennénk a fentiek alapján megmutatjuk, hogyan lehet folytonos

$\Omega = \mathbb{R}, (-\infty, b], [a, \infty), [a, b]$ eseménytérre megadni a (lehető) legnagyobb $\mathcal{S}$ σ -algebrát és azon az egyenletes eloszlást meghatározó P valószínűségi mértékét. Az elmélet alapjainak kidolgozói Henri Lebesgue (1875-1941) és Émile Borel (1871-1956) francia matematikusok voltak.

2.9.7.1. DEFINÍCIÓ. Halmazok külső Lebesgue-mértéke

Tetszőleges $A \subseteq \mathbb{R}$ részhalmaz Lebesgue **külső mértéke** alatt az

$$m^*(A) = \inf Z_A$$

nem-negatív számot vagy a ∞ -t értjük, ahol

$$Z_A = \left\{ \sum_{n=1}^{\infty} L(I_n) \mid I_n = [a_n, b_n] \text{ intervallum, } A \subseteq \bigcup_{n=1}^{\infty} I_n \right\} \text{ és } L(I_n) = b_n - a_n \text{ az intervallum hossza.}$$

Tehát az A halmazt lefedő intervallumrendszerek összhosszának infimumát (alsó korlátok közül a legnagyobbat) értjük a külső mértéken.

A külső mértékkel az a baj, hogy nem rendelkezik a várt additivitási tulajdonsággal, hanem megmutatható, hogy csak szubadditív, azaz $m^*(A \cup B) \leq m^*(A) + m^*(B)$. Ahhoz, hogy additív tulajdonságot kapjunk a komplementer halmaz külső mértékéről is kell tudni valamit.

2.9.7.2. DEFINÍCIÓ. Halmazok Lebesgue mértéke

Valamely $A \subseteq \mathbb{R}$ részhalmazról azt mondjuk, hogy **Lebesgue szerint mérhető**, ha tetszőleges $B \subseteq \mathbb{R}$ halmazra teljesül az

$$m^*(B) = m^*(A \cap B) + m^*(A^c \cap B)$$

egyenlőség, ahol $A^c = \mathbb{R} \setminus A$ az A komplementer halmaza. Jelölje $\mathcal{L}$ az összes ilyen A halmazt, mely a Lebesgue-szerinti mérhető halmazok összessége.

2.9.7.3. TÉTEL. A Lebesgue-mérték σ -additív

Megmutatható, hogy az $m^* : \mathcal{L} \rightarrow [0, \infty)$ külső mérték az $\mathcal{L}$ halmazon σ -additív.

(i) $\emptyset \in \mathcal{L}$

(ii) ha $A \in \mathcal{L}$, akkor $A^c \in \mathcal{L}$.

(iii) ha $A_n \in \mathcal{L}$ minden $n = 1, 2, 3, \dots$ esetén és $A_k \cap A_i = \emptyset$ ($k \neq i$), akkor

$$m^*\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} m^*(A_n)$$

Ha az $m = m^* : \mathcal{L} \rightarrow [0, \infty)$ rendelkezik az (iii) tulajdonsággal, akkor azt mondjuk, hogy m mérték az $\mathcal{L}$ halmazrendszeren és az $(\mathbb{R}, \mathcal{L}, m)$ egy mértéktér.

Az $\mathbb{R}$ valós számokról áttérni tetszőleges Ω mérhető részhalmazra következőképpen lehet. Ehhez tekintsük az $\mathcal{L}_{\Omega} = \{A \cap \Omega \mid A \in \mathcal{L}\}$ σ -algebrát és ezen az eredeti $\mathbb{R}$ -n értelmezett m Lebesgue-

mértéket $m_{\Omega}(A) = m(A)$, ahol $A \in \mathcal{L}_{\Omega}$. Az így kapott $(\Omega, \mathcal{L}_{\Omega}, m_{\Omega})$ mértéktér az Ω mérhető

részalgebrára korlátozott, mely lehet végtelen félegyenes vagy egy véges intervallum.

A mértékelmélet pontosan fordított sorrendben tárgyalja ezeket az eredményeket. Tekint egy tetszőleges nem-üres Ω halmazt és a részhalmazok $\mathcal{S}$ rendszerét, mely σ -algebra és

$m : \mathcal{S} \rightarrow [0, \infty)$ egy mérték $\mathcal{S}$ -n. Ekkor az $(\Omega, \mathcal{S}, m)$ hármas egy mértéktér. A valószínűségi mező olyan speciális mértéktér, melyben az Ω halmaz mértéke 1, vagyis $m(\Omega) = 1$.

2.9.7.4. DEFINÍCIÓ. Borel halmazok

A valós számhalmazok külső mértékéből kiindulva a 2.9.7.2. definícióval sikerült olyan halmazokat kapni, amelyből a Vitali-féle halmazok kívül rekedtek. De vajon milyen halmazok maradtak belül? Erre ad választ a Borel-halmazok rendszere, mely az intervallumok által generált σ -algebra.

A valós számok

$\mathcal{B} = \bigcap \{ \mathcal{G} \mid \mathcal{G} \text{ egy } \sigma\text{-algebra, mely tartalmazza } \mathbb{R} \text{ összes intervallumát} \}$
halmazrendszerének elemeit **Borel-halmazoknak** nevezzük.

Korlátozhatjuk az összes intervallumot csak a nyitott (a, b) intervallumokra vagy csak a zárt $[a, b]$, $[a, b)$, $(a, b]$ intervallumokra, akkor is a fenti $\mathcal{B}$ - σ -algebrát kapjuk.

Mivel a Lebesgue-mérhető halmazok $\mathcal{L}$ rendszere is σ -algebra és tartalmazza az összes intervallumot, valamint a Borel-féle σ -algebra a legszűkebb ilyen, ezért teljesül a

$$\mathcal{B} \subset \mathcal{L}$$

tartalmazás. A tartalmazás szigorú, mert van olyan halmaz mely Lebesgue-mérhető viszont nem Borel-halmaz. Megmutatható továbbá, hogy bármely $B \in \mathcal{B}$ Borel-halmazhoz van olyan $A \in \mathcal{L}$ Lebesgue mérhető halmaz, hogy a két halmaz Lebesgue-mértéke egyenlő $m(B) = m(A)$.

Azt mondhatjuk, hogy a Lebesgue-mérhető halmazok rendszere folytonos lezártja a Borel-halmazok rendszerének ugyanúgy mint, ahogyan a valós számok halmaza folytonos lezártja a racionális számok halmazának. Az előbbi esetben a mérést a Lebesgue-mérték adja, míg az utóbbinál a valós számok távolsága.

2.10. Példák események valószínűségeinek számítására

Ha ismerjük az eseménytér néhány eseményének valószínűségét, akkor a műveletekkel kapott események valószínűségeit meg tudjuk határozni a levezetett műveleti szabályokat használva. Ugyanez a számítás fordítva is kivitelezhető.

2.10.1. PÉLDA

Legyenek adottak a $P(A^c) = 0.6$, $P(B) = 0.3$ és $P(A \cdot B) = 0.2$ valószínűségek.

Határozzuk meg a $P(A + B)$ és $P(A^c \cdot B^c)$ valószínűségeket!

Megoldás

Adjuk át változóknak a valószínűségeket a Maple többszörös értékadó utasítását használva!

> restart :

$$P(Ac), P(B), P(AB) := 0.6, 0.3, 0.2$$

$$P(Ac), P(B), P(AB) := 0.6, 0.3, 0.2 \quad (1.10.1)$$

Mivel az összeg esemény valószínűségére vonatkozó

$P(A + B) = P(A) + P(B) - P(A \cdot B)$ azonosságban szükség van a $P(A)$ esemény valószínűségére, ezért előbb a tagadás esemény valószínűségéről szóló tételt kell alkalmazni.

> $(P(A)) := 1 - P(Ac)$

$$P(A) := 0.4 \quad (1.10.2)$$

> $\text{összeg} := P(A + B) = P(A) + P(B) - P(AB); \text{assign}(\text{összeg}) :$

$$\text{összeg} := P(A + B) = 0.5 \quad (1.10.3)$$

A Maple a kifejezések kiértékelését a **teljes kiértékelési láncnak** megfelelően végzi. Vagyis azokra a változó nevekre, amelyekre talál értékadást, azt a kifejezésbe behelyettesíti.

Amennyiben a behelyettesítés után újabb változó neveket talál, akkor ezt a folyamatot addig végzi, amíg van behelyettesítendő érték. Mivel az összegben szereplő $P(A)$, $P(B)$ és $P(AB)$ változó nevek kaptak értéket korábban, ezért ezeket behelyettesíti. Így kaptuk az összeg esemény valószínűségére a $P(A + B) = 0.5$ értéket!

A $P(A^c \cdot B^c)$ valószínűséget a tagadás esemény valószínűségét és a DeMorgan azonosságot alkalmazva kapjuk. Ugyanis a DeMorgan - azonosság szerint $(A + B)^c = A^c \cdot B^c$. Ahonnan

$$P(A^c \cdot B^c) = P((A + B)^c) = 1 - P(A + B)$$

a tagadás esemény azonosságát alkalmazva.

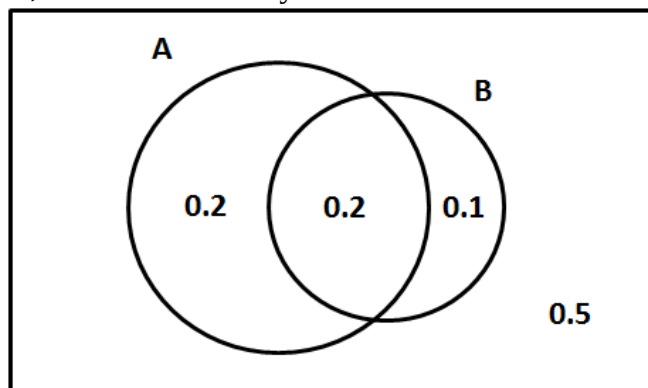
$$> (P(AcBc)) := 1 - P(A + B)$$

$$P(AcBc) := 0.5$$

(1.10.4)

Tehát $P(A + B) = 0.5$ és $P(A^c \cdot B^c) = 0.5$.

Az alábbi Venn-diagram az egyes diszjunkt részek valószínűségeit mutatja. A valószínűségekkel úgy számolhatunk, mintha a tartományok területe lenne.



A valószínűségek kitöltési sorrendje a következő: A középen álló 3 valószínűség közül a középsőt tudjuk, mert ez a metszet és valószínűsége 0.2. Mivel $P(B) = 0.3$ adott és ebből a metszetre 0.2 jut, ezért a jobb oldali részbe $0.1 = 0.3 - 0.2$ valószínűség marad. Miután kiszámoltuk, hogy az A esemény valószínűsége 0.4 és ebből a metszetre 0.2 jut, ezért a bal oldali részre marad $0.2 = 0.4 - 0.2$. Végül, ha összeadjuk, az összes eddig beírt 3 valószínűség (terület) értéket $0.2 + 0.2 + 0.1 = 0.5$ értéket kapunk, mely a keresett $A + B$ összeg valószínűsége. Marad $A + B$ unión kívüli területre is $0.5 = 1 - 0.5$, amely a $P(A^c \cdot B^c)$ valószínűsége.

2.10.2. PÉLDA

Legyenek adottak a $P(A + B) = \frac{3}{4}$, $P(A^c) = \frac{2}{3}$ és a $P(A \cdot B) = \frac{1}{4}$ valószínűségek.

Határozzuk meg a $P(A)$, $P(B)$ és $P(A \cdot B^c)$ valószínűségeket!

Megoldás

Kezdjük most is a megadott valószínűségek Maple változó nevekhez rendelésével!

$$> P(A + B), P(Ac), P(AB) := \frac{3}{4}, \frac{2}{3}, \frac{1}{4}$$

$$P(A + B), P(Ac), P(AB) := \frac{3}{4}, \frac{2}{3}, \frac{1}{4}$$

(1.10.5)

A tagadás esemény azonossága szerint $P(A) = 1 - P(A^c)$, ezért azonnal megkapjuk $P(A)$ valószínűségét!

$$> (P(A)) := 1 - P(Ac)$$

$$P(A) := \frac{1}{3}$$

(1.10.6)

Ha felírjuk az $A + B$ összeg esemény valószínűségére az azonosságot, akkor abban csak a $P(B)$ valószínűség lesz az ismeretlen. Ezért az egyenletből kifejezve, megkapjuk $P(B)$ értékét!

$$> \text{összeg} := P(A + B) = P(A) + P(B) - P(AB);$$

$$\text{isolate}(\text{összeg}, P(B))$$

$$\text{összeg} := \frac{3}{4} = \frac{1}{12} + P(B)$$

$$P(B) = \frac{2}{3} \quad (1.10.7)$$

Az *isolate* eljárás kifejezte, izolálta az egyenletből a $P(B)$ változót. Ezzel még nem kapott értéket a $P(B)$ név. Alkalmazzuk az *assign(%)* eljárást ahhoz, hogy az előzőekben kapott egyenlőség bal oldala vegye fel a jobb oldalon szereplő értéket!

> *assign(%)*; $P(B)$

$$\frac{2}{3} \quad (1.10.8)$$

A $P(A \cdot B^c)$ valószínűség kiszámításához keresnünk kell olyan azonosságot, melyben az ismeretlen valószínűség más ismert valószínűségekkel együtt szerepel. Nem sokáig kell keresgélni, mert alkalmazzuk a következő ismert halmaz azonosságot

$$A = A \cdot \Omega = A \cdot (B + B^c) = A \cdot B + A \cdot B^c$$

Ebben a jobb oldalon szereplő két esemény egymást kizáró $(A \cdot B) \cap (A \cdot B^c) = \emptyset$, mert az $A \cdot B \subseteq B$ valamint $A \cdot B^c \subseteq B^c$. Kolmogorov (A.3.3) axiómája alapján az összeg valószínűségét ilyenkor tagonként vehetjük. Ezért $P(A) = P(AB) + P(AB^c)$ egyenletet kapjuk, melyben ismeretlen csak a keresett $P(A \cdot B^c)$ valószínűség!

> *azonosság* := $P(A) = P(AB) + P(AB^c)$

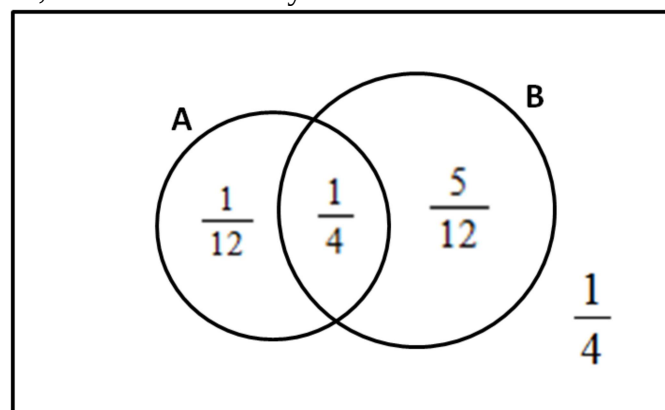
$$\text{azonosság} := \frac{1}{3} = \frac{1}{4} + P(AB^c) \quad (1.10.9)$$

> *isolate(azonosság, P(AB^c))*

$$P(AB^c) = \frac{1}{12} \quad (1.10.10)$$

Tehát a keresett valószínűségek rendre $P(A) = \frac{1}{3}$, $P(B) = \frac{2}{3}$ és $P(A \cdot B^c) = \frac{1}{12}$.

Az alábbi Venn-diagram az egyes diszjunkt részek valószínűségeit mutatja. A valószínűségekkel úgy számolhatunk, mintha a tartományok területe lenne.



A területekkel való számolás sorrendje itt eltér a példa megoldási lépéseitől, de a végeredményeket az ábráról leolvashatjuk. Írjuk be a metszet $\frac{1}{4}$ területét középre, amely adott.

Mivel A valószínűsége

$\frac{1}{3}$, ezért a baloldali részbe $\frac{1}{3} - \frac{1}{4} = \frac{1}{12}$ kerül. Mivel $A + B$ valószínűsége $\frac{3}{4}$ adott, ezért a jobboldali részbe $\frac{3}{4} - \frac{1}{4} - \frac{1}{12} = \frac{5}{12}$ kerül. Végül az $(A + B)$ unión kívüli negyedek tartományba az $1 - \frac{3}{4} = \frac{1}{4}$ kerül.

A választ az ábra alapján adhatjuk meg, mert B valószínűségére $\frac{1}{4} + \frac{5}{12} = \frac{3+5}{12} = \frac{8}{12} = \frac{2}{3}$ adódik. Az $A \cdot B^c$ tartományt pedig fel kell ismerni, hogy az ábrán a bal oldali része. Ezért valószínűsége $\frac{1}{12}$.

2.11. Elméleti ellenőrző kérdések

- 2.11.1. Mit értünk egy A esemény gyakorisága és relatív gyakorisága alatt? Mondjon példát!
- 2.11.2. Milyen alapvető tulajdonságait tudja felsorolni a relatív gyakoriságoknak?
- 2.11.3. Hogyan közelítik a relatív gyakoriságok az A esemény elméleti valószínűségét? Mondjon példát!
- 2.11.4. Milyen részekre oszthatók a valószínűségelmélet Kolmogorov-axiómái?
- 2.11.5. Mit értünk valószínűségi mező alatt? Mondjon példákat!
- 2.11.6. Milyen valószínűségi mezőt nevezünk klasszikusnak? Mondjon példákat!
- 2.11.7. Milyen szabállyal kapjuk egy klasszikus valószínűségi mezőben az A esemény valószínűségét? Mondjon példát!
- 2.11.8. Mit mond ki a $P(A^c)$ tagadás vagy komplementer esemény kiszámításáról szóló szabály?
- 2.11.9. Mit mond ki a $P(A + B)$ összeg esemény kiszámításáról szóló szabály?
- 2.11.10. Mit mond ki a $P(A - B)$ különbség esemény kiszámításáról szóló szabály?
- 2.11.11. Mit mond ki az $A \subseteq B$ részesemények valószínűségeinek összehasonlításáról szóló tétel?
- 2.11.12. Hogyan adható meg a geometriai- eloszlás valószínűségi mezője? Milyen kísérlettel kapunk ilyen mezőt?
- 2.11.13. Hogyan adható meg egy $\lambda > 0$ paraméterű Poisson- eloszlás valószínűségi mezője? Milyen kísérlettel kapunk ilyen mezőt?
- 2.11.14. Jellemezze az $[a, b]$ intervallumon folytonos egyenletes eloszlás valószínűségi mezőjét! Mondjon példát!
- 2.11.15. Adja meg az $[a, b]$ intervallumon folytonos egyenletes eloszlás sűrűségfüggvényét és eloszlásfüggvényét! Milyen kapcsolat van a két függvény között?
- 2.11.16. Jellemezze a $\lambda > 0$ paraméterű exponenciális eloszlás valószínűségi mezőjét! Mondjon példát!
- 2.11.17. Adja meg a $\lambda > 0$ paraméterű exponenciális eloszlás sűrűségfüggvényét és eloszlásfüggvényét! Milyen kapcsolat van a két függvény között?
- 2.11.18. Milyen módon tudunk egy $[0,1]$ intervallumon egyenletes eloszlásból előállítani $\lambda > 0$ paraméterű exponenciális eloszlást?
- 2.11.19. Milyen kapcsolat van egy folytonos eloszlás sűrűségfüggvénye és az adatok relatív gyakoriságainak hisztogramja között? Mondjon példát!
- 2.11.20. Miért szükséges a σ -algebra alkalmazása a valószínűségelméletben? Mondjon példát!

2.12. Gyakorló feladatok

2.12.1.feladat

Kérdezzük meg diákjainkat, hogy árulják el melyik hónapban születtek.

(i) Készítsünk az adatokból gyakoriság és relatív gyakoriság adatokat és ábrázoljuk ezeket!

(ii) Adjuk meg a **szubjektív valószínűségi adatokhoz** az (Ω, S, P) valószínűségi mezőt!

(iii) Levonható-e valamilyen **következtetés** az adatok eloszlására?

2.12.2. feladat

Szennyvíz tisztítási folyamatból kikerülő megtisztított víz analízise során a mikroszkóp tárgyasztalára helyezett folyadékban levő **veszélyes sejtek számát** vizsgálják. Elvégeztek 100 független mérést egy $1\text{ mm} \times 1\text{ mm} = 1\text{ mm}^2$ -es négyzet alakú területen található veszélyes sejtek számára. Az alábbi táblázat a 100 mérési eredményből az egyes káros sejtek számának gyakoriságait mutatja

A káros sejtek száma	0	1	2	3	4	5	6	7	8	9	10	11
Előfordulási gyakorisága	1	3	8	14	17	19	14	12	6	2	2	2

(i) Készítsünk az adatokból relatív gyakoriság adatokat és ábrázoljuk!

(ii) Adjuk meg az **empirikus valószínűségi adatokhoz** az (Ω, S, P) valószínűségi mezőt!

(iii) Levonható-e valamilyen következtetés az adatok eloszlására?

2.12.3. feladat

Véletlenszerűen választunk egy számot az egymást követő $\{1, 2, 3, 4, \dots, 1000\}$ természetes számok közül.

(i) Mekkora a valószínűsége, hogy a választott szám **osztható hárommal**?

(ii) Végezzünk szimulációs kísérletet 200 véletlenszám generálásával és rajzoljuk fel a relatív gyakoriságokat!

(iii) Vessük össze a valószínűséget és a relatív gyakoriságok változását!

2.12.4. feladat

Véletlenszerűen választunk egy számot az egymást követő $\{1, 2, 3, 4, \dots, 1000\}$ természetes szám közül.

(i) Mekkora a valószínűsége, hogy a szám osztható **ötten**?

(ii) Mekkora a valószínűsége, hogy a szám osztható **hárommal is és ötten is**?

(iii) Mekkora a valószínűsége, hogy a szám osztható **hárommal vagy ötten**?

2.12.5. feladat

Hány olyan valódi négyjegyű szám (a nulla jegy nem lehet elöl) van, melynek számjegyei között

(a) nincs páros számjegy?

(b) legalább egy páros számjegy szerepel?

(c) Mekkora a valószínűsége, hogy az összes szám közül kiválasztva egyet az a (b) tulajdonsággal rendelkezik?

(d) Készítsünk szimulációt a (c) feladatra!

2.12.6. feladat

Adja meg a két független kocka dobás kísérletének (Ω, S, P) valószínűségi mezőjét!

2.12.7. feladat

Kör alakú céltáblára lövésnél milyen sugarú koncentrikus köröket válasszunk, hogy az esélyek az $n = 5$ tartományba esésnél egyformák legyenek? Adja meg a kísérlet (Ω, S, P) esemény mezőjét!

2.12.8. feladat

Egy véletlen kísérlet eseménytere $\Omega = \{1, 2, 3, 4\}$ és a megfelelő valószínűségek

$$p_1 = \frac{1}{2}, p_2 = \frac{1}{4}, p_3 = \frac{1}{8}, p_4 = \frac{1}{8}.$$

(a) Mutassa meg, hogy megadható olyan kísérlet egy szabályos pénzérme feldobásával, amely a fenti valószínűségi mezőt eredményezi!

(b) Mutassa meg, hogy megadható olyan kísérlet egy urna modell segítségével, amely a fenti valószínűségi mezőt eredményezi!

(c) Mutassa meg, hogy megadható olyan kísérlet egy 52 kártyát tartalmazó pakli

segítségével, amely a fenti valószínűségi mezőt eredményezi!

2.12.9. feladat

Egy véletlen kísérlet abból áll, hogy két golyót kihúzunk egymás után egy olyan urnából, amelyben 3 fekete és 4 piros golyó van.

- (a) Adja meg a kísérlet Ω eseményterét!
- (b) A kísérletet módosítsuk úgy, hogy az elsőre kivett golyót visszatesszük. Mi lesz ekkor az eseménytér?
- (c) Végezzon Maple szimulációt a (piros, piros) kimenetel relatív gyakoriságainak felrajzolására az (a) modell esetén! A (b) modell esetén!
- (d) Mennyi lesz a valószínűsége a (piros, piros) kimenetelnek az (a) modell esetén? A (b) modell esetén?

2.12.10. feladat

Tegyük fel, hogy a $V(t) = 2 \cdot \cos(2 \cdot \pi \cdot t)$ feszültségű elektromos jelből veszünk a t időben véletlenszerű mintákat.

- (a) Határozza meg a $V_+ = \{a \text{ feszültség pozitív} \}$ és $V_- = \{a \text{ feszültség kisebb, mint } (-2) \}$ események relatív gyakoriságainak hosszútávú viselkedését
- (c) Változik-e a válasz akkor, ha a mintavételt $t = \tau$ időközönként vesszük?

2.12.11. feladat

Annak érdekében, hogy a véletlenszámok generálása mennyire véletlen, generáljunk 7 számjegyű telefon számokat. Legyen a kísérlet kimenetele $X = 0$, ha a telefonszám utolsó számjegye páros és legyen $X = 1$, ha az utolsó számjegy páratlan.

- (a) Mit várunk el a kísérlet kimeneteleitől, ha a kapott sorozat valóban "véletlen"?
- (b) Teszteljük a véletlenszerűséget az elemi események relatív gyakoriságainak felrajzolásával!
- (c) Teszteljük a véletlenszerűséget az elemi eseményekből képezett rendezett párok relatív gyakoriságainak felrajzolásával!

2.12.12. feladat

Egy liftbe 2 utas száll be a földszinten, akik bármelyike ugyanolyan eséllyel száll ki az első, második, harmadik vagy negyedik emeleten.

- (i) Mekkora a valószínűsége, hogy a két utas különböző emeleten száll ki?
- (ii) A fenti eseményre készítsünk szimulációs programot a relatív gyakoriságok felrajzolásával!
- (iii) Hasonlítsuk össze a relatív gyakoriságok ingadozását és az elméleti valószínűséget!

2.12.13. feladat

Bizonyítsa be, hogy ha B részeseménye az A eseménynek $B \subset A$, akkor

$$P(A - B) = P(A) - P(B)$$

2.12.14. feladat

Legyenek adottak a $P(A) = \frac{5}{6}$, $P(A + B) = \frac{8}{9}$ és $P(A \cdot B^c) = \frac{5}{9}$ valószínűségek.

Határozzuk meg a $P(B)$ és $P(A \cdot B)$ valószínűségeket!

2.12.15. feladat

Legyenek adottak a $P(B) = \frac{1}{3}$, $P(A + B) = \frac{8}{9}$ és $P(A^c \cdot B) = \frac{1}{18}$ valószínűségek.

Határozzuk meg a $P(A)$ és $P(A \cdot B)$ valószínűségeket!

2.12.16. feladat

Legyenek adottak a $P(A + B) = \frac{7}{8}$, $P(A \cdot B) = \frac{1}{10}$ és $P(B - A) = \frac{3}{4}$ valószínűségek.

Határozzuk meg a $P(A)$ és $P(B)$ valószínűségeket!

2.12.17. feladat

- (a) Mennyi a valószínűsége, hogy valamely, taláломra kiválasztott, az egységénél rövidebb éllel

rendelkező

téglalap átlója kisebb az egységnél?

(b) Mennyi a valószínűsége, hogy valamely, találmányra kiválasztott, az egységnél rövidebb élekkel rendelkező

téglalatest testátlója kisebb az egységnél?

2.12.18. feladat

Legyen m végesen additív halmazfüggvény az S algebrán és $m(\Omega)$ véges. Mutassuk meg, hogy az alábbi állítások ekvivalensek

(i) m σ -additív (vagyis, ha S σ -algebra és $m(\Omega) = 1$, akkor m valószínűségi mérték)

(ii) Ha az $A_n \subseteq A_{n+1}$ monoton növekvő halmaz sorozat S -ben, akkor $m\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n \rightarrow \infty} m(A_n)$

(iii) Ha az $A_{n+1} \subseteq A_n$ monoton csökkenő halmaz sorozat S -ben, akkor

$$m\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n \rightarrow \infty} m(A_n).$$

2.12.19. feladat

Legyen adott az Ω eseménytéren két valószínűségi mérték, P_1 és P_2 . Ha a λ tetszőleges 0 és 1 közötti szám $0 \leq \lambda \leq 1$, akkor mutassuk meg, hogy a

$$P(A) := \lambda \cdot P_1(A) + (1 - \lambda) \cdot P_2(A)$$

képlettel értelmezett P függvény kielégíti a valószínűségi axióma összes követelményét.

2.12.20. feladat

Legyen az eseménytér $\Omega = [0, 1]$ intervallum és $A \subseteq [0, 1]$ mérhető halmazra a valószínűségi mérték

$$P(A) = \int_A 1 \, dm. \text{ Speciálisan teljesül a } P([a, b]) = (b - a). \text{ Tekintsük a következő halmaz}$$

sorozatot:

$A_0 = [0, 1]$, $A_1 \subset A_0$ az A_0 halmaz középső harmadának eltávolításával keletkezik. Így

$$A_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right].$$

Legyen $A_2 \subset A_1$ az A_1 halmaz mindegyik intervalluma középső harmadának eltávolításával keletkező intervallumrendszer. Könnyen látható, hogy

$$A_2 = \left(\left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{3}{9}\right]\right) \cup \left(\left[\frac{6}{9}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, \frac{9}{9}\right]\right).$$

Hasonló módon az A_2 intervallum rendszer minden részintervallumának középső harmadát távolítsuk el és így kapjuk az A_3 intervallumrendszert.

Folytassuk az konstrukciót az $A_4, A_5, \dots$ halmazok képzésével!

(a) Számoljuk ki a $P(A_0)$, $P(A_1)$, $P(A_2)$ és $P(A_3)$ mértékeket!

(b) Keressünk formulát a $P(A_n)$ mértékre!

(c) A Cantor-halmazt az $A = \bigcap_{n=0}^{\infty} A_n$ metszet definiálja. Keressük meg a $P(A)$

mértéket! Magyarazzuk meg a kapott eredményt!