Paradoxonok a véletlen matematikájában

Csató Lehel

Matematika és Informatika Tanszék,
Babeş–Bolyai Tudományegyetem, Kolozsvár
http://www.cs.ubbcluj.ro/~csatol

Bolyai Nyári Egyetem 2007

Klasszikus az, amit mindenki szeretett volna már elolvasni, de amit olvasni senki sem szeretne.

Célok

- érdekes feladatok felsorolása;
- alapfogalmak tisztázása;
- Matlab illusztrációk (ahol szükséges)

Tartalom

Tartalom

Definíciók

Paradoxon

Olyan meglepő állítás, mely a **„józan paraszti észnek”** ellentmondani látszik.

Paradoxonok szerepe:

- megvilágítanak egy problémát, melyre megoldást kell javasolni.

(tudománytörténet)

- bemutatják „paraszti” (lásd fentebb) példákon keresztül a *formális gondolkodás* szükségességét.

(oktatás)

Kérdések, melyekre kereshetjük a választ

- van nyerő kombináció a lottón?
- jó, ha kölcsönös ajándékozásnál sorsoljuk az ajándékozó személyt?
- ha ismételten játszunk egy játékot, hogyan kell azt „optimálisan játszani”?
- helyes az a táblázat, melyben az átlagos életkor 26 év; ugyanakkor 50% a valószínűsége annak, hogy valaki a 8. évet *ne* érje meg?
- véletlen a számsorozat, amit látunk?
- mit jelent a val. változók függetlensége?
- Mekkora valószínűséggel lesz egy húr adott hossznál nagyobb?

Dr. Tel

(Pólya György)

A valószínűség nem garancia!

Dr. Tel

Betegvizsgálat után, hosszas fejrázás közben:

- Önnek nagyon súlyos betegsége van; tíz közül csak egy beteg éli túl.

Dr. Tel

(Pólya György)

A valószínűség nem garancia!

Dr. Tel

Betegvizsgálat után, hosszas fejrázás közben:

- Önnek nagyon súlyos betegsége van; tíz közül csak egy beteg éli túl.

majd nyugtatólag:

- De nagy szerencséje van, hogy hozzám fordult! Eddig kilenc hasonló betegem volt és eddig mindenki belehalt ...

Dr. Tel

(Pólya György)

A valószínűség nem garancia!

Dr. Tel

Betegvizsgálat után, hosszas fejrázás közben:

- Önnek nagyon súlyos betegsége van; tíz közül csak egy beteg éli túl.

majd nyugtatólag:

- De nagy szerencséje van, hogy hozzám fordult! Eddig kilenc hasonló betegem volt és eddig mindenki belehalt . . .

Nem biztos, hogy helyesen értelmezett valószínűség!

Tartalom

Diszkrét valószínűségi változók

- Adott egy eseményhalmaz $\Omega = \{A_1, \ldots, A_d\}$
- Adottak az eseményekhez rendelt $p(A_i) \geq 0$ valószínűségek:

$$\sum_i p(A_i) = 1$$

- Numerikus esetben számíthatunk átlagot:

$$\sum_i x_i \cdot p(x_i) = 1$$

Matlab segédfüggvények

A **MATLAB** programnyelv támogatja a

- mátrix-jelölést: ciklusok helyett „matematikusan”:
  $$C = A * B$$
- gyors prototípuskészítést:
  pl. nincsenek változó-kijelentések.
- numerikus szimulációk írását,

Valószínűségszámítási függvények:

- `rand(m,n)` egy $m \times n$-es mátrixot feltölt pszeudorandom számokkal;
- `hist(X)` egy mátrix(vektor) értékeinek gyakoriságát rajzolja ki;

Tartalom

Lottózás

Felmerülő kérdések:

- létezik nyerő stratégia a lottón?
- létezik jó kombináció amit játszani érdemes?

ha ***igen****, melyek ezek*

- természetesen nincs nyerő stratégia,
- Jó kombináció ≡ ha nyerünk, mások ne legyenek a listán.

Túl szabályos tippek: pl. „két egymás utáni szám”, egy ötszámjegyű egymásutáni számsorozat, ...

Lottózás

Felmerülő kérdések:

- létezik nyerő stratégia a lottón?
- létezik jó kombináció amit játszani érdemes?

ha ***igen****, melyek ezek*

- természetesen nincs nyerő stratégia,
- Jó kombináció $\equiv$ ha nyerünk, mások ne legyenek a listán.

Túl szabályos tippek: pl. „két egymás utáni szám”, egy ötszámjegyű egymásutáni számsorozat, ...

Lottózás

Felmerülő kérdések:

- létezik nyerő stratégia a lottón?
- létezik jó kombináció amit játszani érdemes?

ha ***igen***, *melyek ezek*

- természetesen nincs nyerő stratégia,
- Jó kombináció $\equiv$ ha nyerünk, mások ne legyenek a listán.

Túl szabályos tippek: pl. „két egymás utáni szám”, egy ötszámjegyű egymásutáni számsorozat, ...

Lottózás

Felmerülő kérdések:

- létezik nyerő stratégia a lottón?
- létezik jó kombináció amit játszani érdemes?

ha ***igen****, melyek ezek*

- természetesen nincs nyerő stratégia,
- Jó kombináció $\equiv$ ha nyerünk, mások ne legyenek a listán.

Túl szabályos tippek: pl. „két egymás utáni szám”, egy ötszámjegyű egymásutáni számsorozat, ...

0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0001
0000	0000	0000	0000	0000	0000	1010	0000	0000	0000	0101
0000	0100	0000	0000	0000	0000	0000	0000	0000	0001	0000
1100	0010	0000	0000	1000	0000	0000	0000	0000	0000	0000
0000	0001	0000	0000	0000	0000	0000	0000	0001	0000	0000
0000	0000	0000	0000	0000	0000	0001	0000	0000	0000	0000
0000	0000	0000	0000	0100	1000	0000	1000	0000	0000	0100
0000	0000	0000	0000	0000	0000	0000	0000	0010	0000	0000
0000	0000	1000	0100	0000	0000	0010	0000	0000	0000	0000
0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000
0000	0010	0000	0000	0000	0000	0000	0000	0000	0000	0000
0000	0001	0000	0000	0000	0000	0000	0000	0000	0000	0000
0000	0000	0000	0000	0000	0000	0001	0000	0000	0000	0100
0000	0000	0000	0000	0000	0000	0010	0000	0001	0000	0000
0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000
0000	0000	0000	0000	0000	0010	0000	0000	0001	0000	0000
0000	0001	0000	0000	0000	0000	0000	0000	0000	0100	0001
0001	1000	0000	0001	0000	0000	0000	0000	0000	0000	0100
0001	0000	0000	0001	0000	0000	0000	0000	0000	0000	0000
0000	0000	0000	0000	0010	0000	0000	0000	0100	0000	0000
0000	0000	0000	0000	0000	0000	0000	0100	0000	0000	0000
0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000
0000	0000	0000	0000	0000	0000	0000	0000	0010	1000	1000
0000	0000	0000	0000	0000	0001	0000	0000	0000	0000	0000
0000	0000	0000	0001	0000	0000	0000	0000	0000	0000	0001
0000	0010	0000	0000	0000	0000	0000	0000	0000	0000	0000
0000	0000	0000	0000	0000	0000	0000	0001	0000	0000	0000
0000	0000	0000	0000	0000	1000	0000	0000	0000	0000	0000
0000	0000	0000	0000	0000	0000	0000	1000	0000	0000	0000
0000	0000	0000	0000	0000	0000	0100	0000	0000	0000	1001
0000	0000	0000	0000	0001	0000	0000	0000	0000	0000	0000
0000	0010	0001	0000	0000	0000	0000	0000	1000	0000	0000
0000	0010	0000	1000	0100	0000	0001	0000	0000	0000	1000

Egymásutáni számok ritkák:

- két egymásutáni szám:

$$P_{2eu} \approx 89 \cdot \frac{5 \cdot 4}{90 \cdot 89} = 0.2222$$

- három egymásutáni szám:

$$P_{3eu} \approx 88 \cdot \frac{5 \cdot 4 \cdot 3}{88 \cdot 89 \cdot 90} = 0.0075$$

- $$P_{4eu} \approx 0.00017$$

Nagy nyereségre számíthatunk

ha nyerünk és „ritka” kombinációkat választunk.

Egymásutáni számok ritkák:

- két egymásutáni szám:

$$P_{2eu} \approx 89 \cdot \frac{5 \cdot 4}{90 \cdot 89} = 0.2222$$

- három egymásutáni szám:

$$P_{3eu} \approx 88 \cdot \frac{5 \cdot 4 \cdot 3}{88 \cdot 89 \cdot 90} = 0.0075$$

-

$$P_{4eu} \approx 0.00017$$

Nagy nyereségre számíthatunk

ha nyerünk és „ritka” kombinációkat választunk.

MATLAB program

```
% lotto paradoxon szimulacio
% parameterek
nL   = 90;          % az osszes szam
nO   = 5;           % hanyat
nH   = 100000;      % hany jatekot szimulalunk
% generaljuk
lOsszes = ceil(rand(nH,nO)*90);
% rendezzuk NOVEKVO sorrendbe
lOsszes = sort(lOsszes,2);
% keressuk az egymasutani szamokat
fMat   = [1 -1];     % szuro matrix
lSucc = filter(fMat,1,lOsszes,-100,2);
lSucc = lSucc(:,2:end);
% keressuk azon elemeket, ahol 1 a kulonbseg
iTwo = find(lSucc==1);
fprintf('Ket elemhossz: %5.4f',length(iTwo)/nH);

% nullazunk minden indexet, ami NEM egy.
lInd=zeros(size(lSucc));
lInd(iTwo)  = 1;
% ebben - mivel CSAK nulla es egy volt - az eredmeny 0,1,2
fMat   = [1 1];     % szuro matrix
lThree = filter(fMat,1,lInd,-100,2);
iThree = find(lThree==2);

fprintf('Harom elemhossz: %5.4f',length(iThree)/nH);
```

Eredmények (100e):

Teszt#	Kettő	Három
1	0.2238	0.0077
2	0.2226	0.0075
3	0.2203	0.0073
4	0.2232	0.0074
...		

Függetlenség paradoxona

Megfogalmazás:
Két szabályos érme dobásánál jelölje A az „első dobás fej”; B a „második dobás fej”; C pedig azt, hogy a két dobás közül egy (és csak egy) fej.

Ekkor:

- bármely két esemény független, **ugyanakkor**
- bármely kettő meghatározza a harmadikat.

Megjegyzés: A és B függetlenek, ha $p(A,B) = p(A)p(B)$.

Függetlenség paradoxona

Megfogalmazás:
Két szabályos érme dobásánál jelölje A az „első dobás fej”; B a „második dobás fej”; C pedig azt, hogy a két dobás közül egy (és csak egy) fej.

Ekkor:

- bármely két esemény független, **ugyanakkor**
- bármely kettő meghatározza a harmadikat.

Megjegyzés: A és B függetlenek, ha $p(A, B) = p(A)p(B)$.

Ajándékozási paradoxon

Megfogalmazás:
Egy társaság tagjai kölcsönösen meg akarják ajándékozni egymást. Az ajándékokat összegyűjtik majd kisorsolják.

Paradoxon:

Lényegesen nagyobb annak az esélye, hogy lesz valaki, aki visszakapja ajándékát, mint annak, hogy nem kapja senki a saját ajándékát vissza.

Összesen $n!$ eset van. Ebből kedvező

$$\binom{n}{0} n! - \binom{n}{1}(n-1)! + \dots (-1)^n \binom{n}{n} 0!,$$

ennek aránya

$$p_n = \frac{1}{2!} - \frac{1}{3!} - \frac{1}{4!} + \dots (-1)^n \frac{1}{n!} \to \frac{1}{e}$$

Sok kicsi sokra mehet!

Ajándékozási paradoxon

Megfogalmazás:
Egy társaság tagjai kölcsönösen meg akarják ajándékozni egymást. Az ajándékokat összegyűjtik majd kisorsolják.

Paradoxon:

Lényegesen nagyobb annak az esélye, hogy lesz valaki, aki visszakapja ajándékát, mint annak, hogy nem kapja senki a saját ajándékát vissza.

Összesen $n!$ eset van. Ebből kedvező

$$\binom{n}{0}n! - \binom{n}{1}(n-1)! + \dots (-1)^n\binom{n}{n}0!,$$

ennek aránya

$$p_n = \frac{1}{2!} - \frac{1}{3!} - \frac{1}{4!} + \dots (-1)^n\frac{1}{n!} \to \frac{1}{e}$$

Sok kicsi sokra mehet!

Ajándékozási paradoxon

Megfogalmazás:
Egy társaság tagjai kölcsönösen meg akarják ajándékozni egymást. Az ajándékokat összegyűjtik majd kisorsolják.

Paradoxon:
Lényegesen nagyobb annak az esélye, hogy lesz valaki, aki visszakapja ajándékát, mint annak, hogy nem kapja senki a saját ajándékát vissza.

Összesen $n!$ eset van. Ebből kedvező

$$\binom{n}{0} n! - \binom{n}{1}(n-1)! + \dots (-1)^n \binom{n}{n} 0!,$$

ennek aránya

$$p_n = \frac{1}{2!} - \frac{1}{3!} - \frac{1}{4!} + \dots (-1)^n \frac{1}{n!} \to \frac{1}{e}$$

Sok kicsi sokra mehet!

Pétervári paradoxon

Megfogalmazás:
Egy szabályos pénzérmével addig dobunk, amíg fejet nem kapunk. Ha az első fej, akkor kapunk 2 forintot, ha a második, akkor 4-et, ..., ha az n-edik, akkor 2^n forintot.
Mennyi a játék értéke?

∞

Bármennyit fizetünk, várható értékben nyerünk!

Az ismételt játék várható értéke

$$\frac{1}{2}2 + \cdots + \frac{1}{2^n}2^n + \cdots$$

sok kicsi sokra megy! (nem kicsik a mennyiségek ...)

Pétervári paradoxon

Megfogalmazás:
Egy szabályos pénzérmével addig dobunk, amíg fejet nem kapunk. Ha az első fej, akkor kapunk 2 forintot, ha a második, akkor 4-et, . . . , ha az n-edik, akkor 2^n forintot.
Mennyi a játék értéke?

∞

Bármennyit fizetünk, várható értékben nyerünk!

Az ismételt játék várható értéke

$$\frac{1}{2}2 + \cdots + \frac{1}{2^n}2^n + \cdots$$

sok kicsi sokra megy! (nem kicsik a mennyiségek ...)

Pétervári paradoxon

Megfogalmazás:
Egy szabályos pénzérmével addig dobunk, amíg fejet nem kapunk. Ha az első fej, akkor kapunk 2 forintot, ha a második, akkor 4-et, ..., ha az n-edik, akkor 2^n forintot.
Mennyi a játék értéke?

∞

Bármennyit fizetünk, várható értékben nyerünk!

Az ismételt játék várható értéke

$$\frac{1}{2}2 + \cdots + \frac{1}{2^n}2^n + \cdots$$

sok kicsi sokra megy! (nem kicsik a mennyiségek ...)

Modellezési kérdések:

- milyen gyakran mekkorát nyerünk?
- felső korlát esetén mennyi a játék értéke?

Modellezés:

- kiválasztjuk a sorozat hosszát;
- a szimulációk számát;

```
% szentpetervari paradoxon
jHossz = 10000;
mSzam  = 10000;
nyerek = zeros(mSzam,1);
for ii=1:mSzam;
  % dobasok
 dobas = round(rand(jHossz,1));
  % FEJ keresese
 indF  = find(dobas==1);
  % NEM FEJ sorozatok hossza
 diff = [indF;1+jHossz]- [0;indF];
  nyer = 2.^diff;
  % szamitjuk a nyereseget
  nyerek(ii)=sum(nyer);
end;
% hisztogramot iratunk
hist(log10(nyerek),200)
```

Modellezési kérdések:

- milyen gyakran mekkorát nyerünk?
- felső korlát esetén mennyi a játék értéke?

Modellezés:

- kiválasztjuk a sorozat hosszát;
- a szimulációk számát;

```
% szentpetervari paradoxon
jHossz = 10000;
mSzam  = 10000;
nyerek = zeros(mSzam,1);
for ii=1:mSzam;
  % dobasok
 dobas = round(rand(jHossz,1));
  % FEJ keresese
 indF  = find(dobas==1);
  % NEM FEJ sorozatok hossza
 diff = [indF;1+jHossz]- [0;indF];
  nyer = 2.^diff;
  % szamitjuk a nyereseget
  nyerek(ii)=sum(nyer);
end;
% hisztogramot iratunk
hist(log10(nyerek),200)
```

Valószínűségi paradoxonok

Csató Lehel

Bevezető

Matematikai fogalmak
Valószínűségi változók
MATLAB

Paradoxonok
Lotto
Függetlenség
Ajándékozás
Pétervár
Játékelmélet
Halandóság
Bernoulli
Első előfordulás
De Moivre

Modellezési kérdések:

- milyen gyakran mekkorát nyerünk?
- felső korlát esetén mennyi a játék értéke?

Modellezés:

- kiválasztjuk a sorozat hosszát;
- a szimulációk számát;

```
% szentpetervari paradoxon
jHossz = 10000;
mSzam  = 10000;
nyerek = zeros(mSzam,1);
for ii=1:mSzam;
  % dobasok
 dobas = round(rand(jHossz,1));
  % FEJ keresese
 indF  = find(dobas==1);
  % NEM FEJ sorozatok hossza
 diff = [indF;1+jHossz]- [0;indF];
  nyer = 2.^diff;
  % szamitjuk a nyereseget
  nyerek(ii)=sum(nyer);
end;
% hisztogramot iratunk
hist(log10(nyerek),200)
```

MATLAB program I

Modellezési kérdések:

- milyen gyakran mekkorát nyerünk?
- felső korlát esetén mennyi a játék értéke?

Modellezés:

- kiválasztjuk a sorozat hosszát;
- a szimulációk számát;

```
% szentpetervari paradoxon
jHossz = 10000;
mSzam  = 10000;
nyerek = zeros(mSzam,1);
for ii=1:mSzam;
  % dobasok
 dobas = round(rand(jHossz,1));
  % FEJ keresese
 indF  = find(dobas==1);
  % NEM FEJ sorozatok hossza
diff = [indF;1+jHossz]- [0;indF];
  nyer = 2.^diff;
  % szamitjuk a nyereseget
  nyerek(ii)=sum(nyer);
end;
% hisztogramot iratunk
hist(log10(nyerek),200)
```

MATLAB program I

Modellezési kérdések:

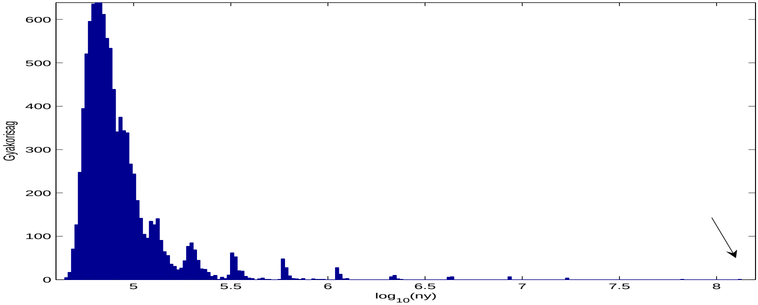
- milyen gyakran mekkorát nyerünk?
- felső korlát esetén mennyi a játék értéke?

Modellezés:

- kiválasztjuk a sorozat hosszát;
- a szimulációk számát;

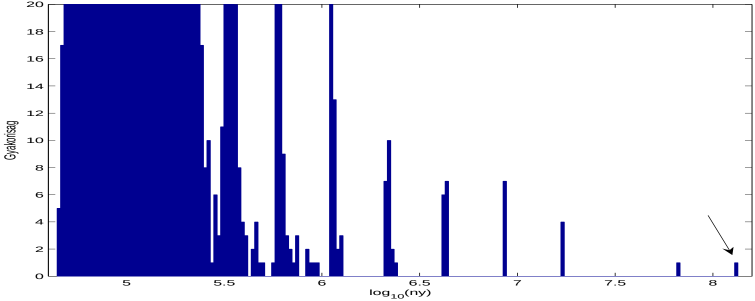
```
% szentpetervari paradoxon
jHossz = 10000;
mSzam  = 10000;
nyerek = zeros(mSzam,1);
for ii=1:mSzam;
  % dobasok
 dobas = round(rand(jHossz,1));
  % FEJ keresese
 indF  = find(dobas==1);
  % NEM FEJ sorozatok hossza
 diff = [indF;1+jHossz]- [0;indF];
  nyer = 2.^diff;
  % szamitjuk a nyereseget
  nyerek(ii)=sum(nyer);
end;
% hisztogramot iratunk
hist(log10(nyerek),200)
```

Teszteredmények I



Hossz	Átlag	Átlagnyereség (Ny/dobás)	Max
100	800	8	524.288
1.000	11.000	11	2.097.152
5.000	91.000	18.2	16.777.216
$\Rightarrow$ 10.000	168.000	16.8	134.217.728
100.000	2.000.000	20.0	1.073.741.824
500.000	23.000.000	46.0	137.000.000.000

Teszteredmények I



Hossz	Átlag	Átlagnyereség (Ny/dobás)	Max
100	800	8	524.288
1.000	11.000	11	2.097.152
5.000	91.000	18.2	16.777.216
$\Rightarrow$ 10.000	168.000	16.8	134.217.728
100.000	2.000.000	20.0	1.073.741.824
500.000	23.000.000	46.0	137.000.000.000

Teszteredmények II

A gyakorlatban: a nyereménynek van egy felső határa, ez maximálja a lehetséges nyereséget. (alább $2^{23} \approx 8m$ a felső határ).

Felső korlát esetén:

Hossz	Átlag	Ny/d.	Korláttal	Ny/d.	Vágás #
100	800	8	800	8	0
1.000	11.000	11	11.000	11	1
5.000	91.000	18.2	60.000	12	2
10.000	168.000	16.8	117.000	11.7	5
100.000	2.000.000	20	1.200.000	12	60
500.000	23.000.000	46	6.000.000	12	295

Játékelméleti paradoxon I

Ketten játszanak: Q és R.

- Egy vagy két ujjat mutatnak fel;
- Páros ujj-szám esetén Q fizet R-nek, ellenkező esetben fordítva;
- annyit nyernek/veszítenek, ahány ujjat felmutattak.

Van véletlennél jobb stratégia

IGEN, egyik játékos veszít!

Melyik?

Játékelméleti paradoxon I

Ketten játszanak: Q és R.

- Egy vagy két ujjat mutatnak fel;
- Páros ujj-szám esetén Q fizet R-nek, ellenkező esetben fordítva;
- annyit nyernek/veszítenek, ahány ujjat felmutattak.

Van véletlennél jobb stratégia

IGEN, egyik játékos veszít!

Melyik?

Játékelméleti paradoxon II

Ismételt játékok esetén mi a nyerő stragégia?

Nyereség/veszteség mátrix:

	Q.1	Q.2
$R.1$	2	-3
$R.2$	-3	4

$$\mathbf{r}^T\ \mathbf{M}\ \mathbf{q} = 2r_1q_1 - 3r_1q_2 - 3r_2q_1 + 4r_2q_2$$

Mindegyik játékos úgy játszik, hogy az eredmény **ne** függjön a másik játékostól (és tudva, hogy $p_2 = 1 - p_1,\ q_2 = 1 - q_1$):

$$(12r_1 - 7)q_1 + (4 - 7r_1).$$

R választása: $12r_1 - 7 = 0$.

A megoldás:

$$r_1 = 7/12,\ r_2 = 5/12,\ q_1 = 7/12,\ q_2 = 5/12$$

Átlegnyereség-veszteség: $-1/12$

Halandósági paradoxon

- az élettartam matematikai vizsgálata *Halley 1693*;
- biztosításelmélet XVII. században indul jelentős fejlődésnek;
- életkor–táblázatok összeállítása;
- *d'Alembert* nevéhez kötődik:

Életkor

Az átlagos életkor 26 év, mégis ugyanakkora az esélye annak, hogy valaki túléli a nyolc évet, mint annak, hogy nem.

- van az átlagon kívül más jellemző, ami hasznos?
- ha **csak** az átlagot ismerjük, mennyit **tudunk** az adatokról?

Halandósági paradoxon

- az élettartam matematikai vizsgálata *Halley 1693*;
- biztosításelmélet XVII. században indul jelentős fejlődésnek;
- életkor–táblázatok összeállítása;
- *d'Alembert* nevéhez kötődik:

Életkor

Az átlagos életkor 26 év, mégis ugyanakkora az esélye annak, hogy valaki túléli a nyolc évet, mint annak, hogy nem.

- van az átlagon kívül más jellemző, ami hasznos?
- ha **csak** az átlagot ismerjük, mennyit **tudunk** az adatokról?

Halandósági paradoxon

- az élettartam matematikai vizsgálata *Halley 1693*;
- biztosításelmélet XVII. században indul jelentős fejlődésnek;
- életkor–táblázatok összeállítása;
- *d'Alembert* nevéhez kötődik:

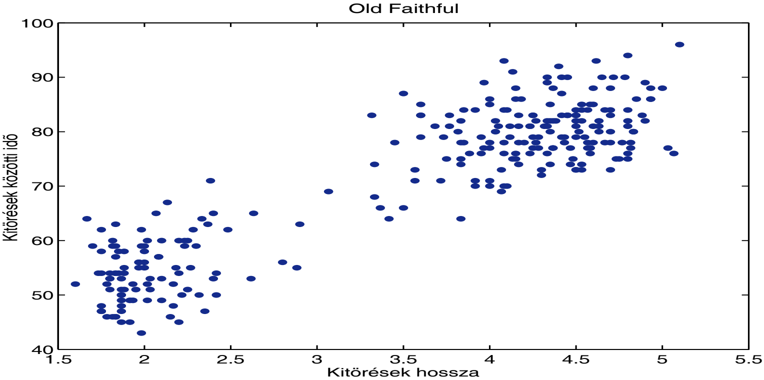
Életkor

Az átlagos életkor 26 év, mégis ugyanakkora az esélye annak, hogy valaki túléli a nyolc évet, mint annak, hogy nem.

- van az átlagon kivül más jellemző, ami hasznos?
- ha **csak** az átlagot ismerjük, mennyit **tudunk** az adatokról?

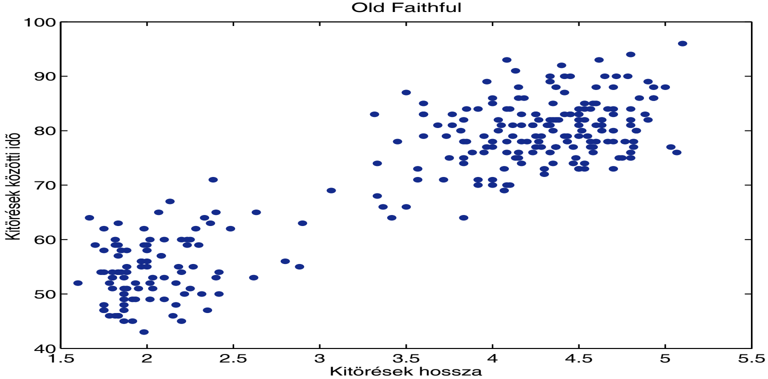
Old Faithful gejzír

Old Faithful gejzír



Statisztikák

Old Faithful gejzír

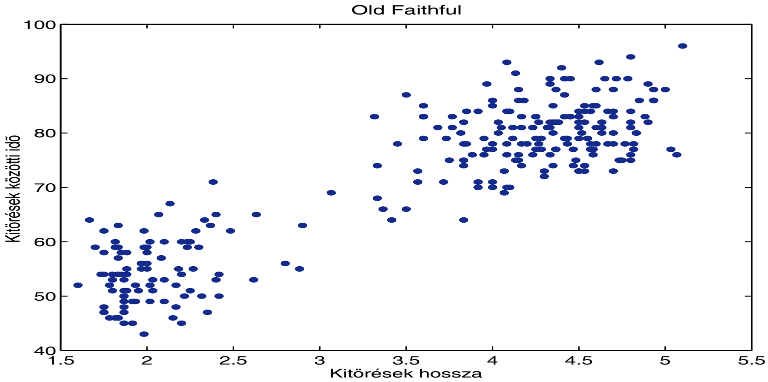


A grafikon leírására

- az átlag jó jellemző?
- más statisztika építése?

Statisztikák

Old Faithful gejzír



A grafikon leírására

- az átlag jó jellemző?
- más statisztika építése?

Más statisztikai modell felállítása I

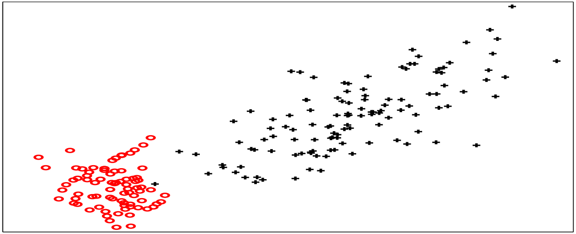
A következő modell:

$$p(x) = \pi_1 \mathcal{N}(x|m_1, \Sigma_1) + \pi_2 \mathcal{N}(x|m_2, \Sigma_2)$$

A következő modell:

$$p(x) = \pi_1 \mathcal{N}(x|m_1, \Sigma_1) + \pi_2 \mathcal{N}(x|m_2, \Sigma_2)$$

A modell **magyaráz** csoportosulást



melyben

- a különböző csoportokhoz tartozó pontok arányai π_1 illetve π_2,
- a csoportok leírhatók egy-egy Gaussz-eloszlással,
- ellenben nem tudjuk, melyik csoportba is tartozik egy-egy pont.

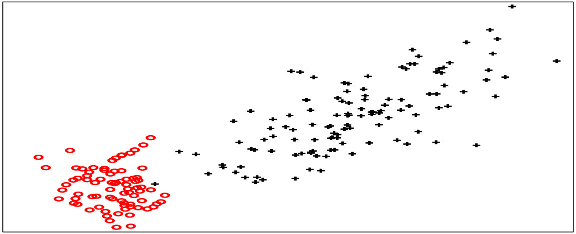
Más statisztikai modell felállítása

I

A következő modell:

$$p(x) = \pi_1 \mathcal{N}(x|m_1, \Sigma_1) + \pi_2 \mathcal{N}(x|m_2, \Sigma_2)$$

A modell **magyaráz** csoportosulást



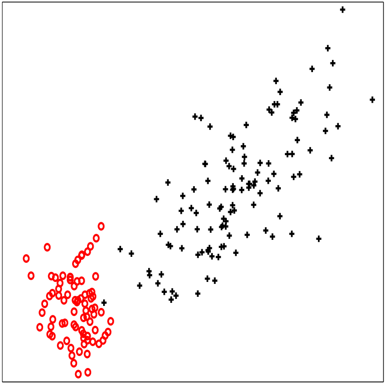
melyben

- a különböző csoportokhoz tartozó pontok arányai π_1 illetve π_2,
- a csoportok leírhatók egy-egy Gaussz-eloszlással,
- ellenben nem tudjuk, melyik csoportba is tartozik egy-egy pont.

Más statisztika felállítása II

Statisztikák:

- π_1, π_2 a csoportok aránya;
- m_1, m_2 a csoportok középpontja;
- Σ_1, Σ_2 a csoportok szórásmátrixa;



Nehéz feladat

a pontok hovatartozásának a megállapítása.

„Expectation-Maximisation”

Nagy számok paradoxona I

- *Bernoulli – Ars conjectandi, 1713*

- sokszori dobás esetén, az írás/fej arány egyhez tart szabályos érme esetén;
- a dobások függetlenek;

100 „írás” dobása után is 50% az írás valószínűsége

Bármennyi sok „fej” dobás esetén, a következő dobás $1/2$ valószínűséggel lesz „fej”.

Nagy számok paradoxona I

- *Bernoulli – Ars conjectandi, 1713*

- sokszori dobás esetén, az írás/fej arány egyhez tart szabályos érme esetén;
- a dobások függetlenek;

100 „írás” dobása után is 50% az írás valószínűsége

Bármennyi sok „fej” dobás esetén, a következő dobás $1/2$ valószínűséggel lesz „fej”.

- nem valószínű a hosszú sorozat, ellenben
- a pénzérmének nincs emlékezete, tehát nem a „múlt” alapján dönt.

Nagy számok paradoxona I

- *Bernoulli – Ars conjectandi, 1713*

- sokszori dobás esetén, az írás/fej arány egyhez tart szabályos érme esetén;
- a dobások függetlenek;

100 „írás” dobása után is 50% az írás valószínűsége

Bármennyi sok „fej” dobás esetén, a következő dobás $1/2$ valószínűséggel lesz „fej”.

- nem (?) valószínű a hosszú sorozat, ellenben
- a pénzérmének nincs emlékezete, tehát nem a „múlt” alapján dönt.

Sorozatok hossza

MATLAB

```
%% BINARIS SZAMSOROZATOK HOSSZA
nRun = 500000;
nN   = [10 100 1000];
%% futasi valtozok
nBin = floor(5+ 2.5*log2(nN(end)));
vBins = zeros(length(nN),nBin);
bins  = 1:nBin;
bb    = 1;
%% kiserletek
for iSize=nN;
 for ii=1:nRun;
  ossSor = 2*round(rand(1,iSize))-1;
  ossSor(:,iSize+1) = - ossSor(:,iSize);
  valt   = ossSor(:,2:end)-ossSor(:,1:(end-1));
  valt   = abs(valt)/2;
  indF     = find(valt==1);
  hossz    = indF - [0 indF(1:end-1)];
  [bBin,bins] = hist(hossz,bins);
  vBins(bb,:) = vBins(bb,:) + bBin;
 end;
 bb=bb+1;
end;
```

```
%% kirajzolas
clf; hold on; box on;
semilogy(bins,vBins');
xlabel('Hossz');
ylabel('Hanyszor');
ylim([0 100]);
```

```
%% BINARIS SZAMSOROZATOK HOSSZA
nRun = 500000;
nN   = [10 100 1000];
%% futasi valtozok
nBin = floor(5+ 2.5*log2(nN(end)));
vBins = zeros(length(nN),nBin);
bins  = 1:nBin;
bb    = 1;
%% kiserletek
for iSize=nN;
 for ii=1:nRun;
  ossSor = 2*round(rand(1,iSize))-1;
  ossSor(:,iSize+1) = - ossSor(:,iSize);
 valt  = ossSor(:,2:end)-ossSor(:,1:(end-1));
  valt   = abs(valt)/2;
  indF     = find(valt==1);
  hossz    = indF - [0 indF(1:end-1)];
  [bBin,bins] = hist(hossz,bins);
  vBins(bb,:) = vBins(bb,:) + bBin;
 end;
 bb=bb+1;
end;
```

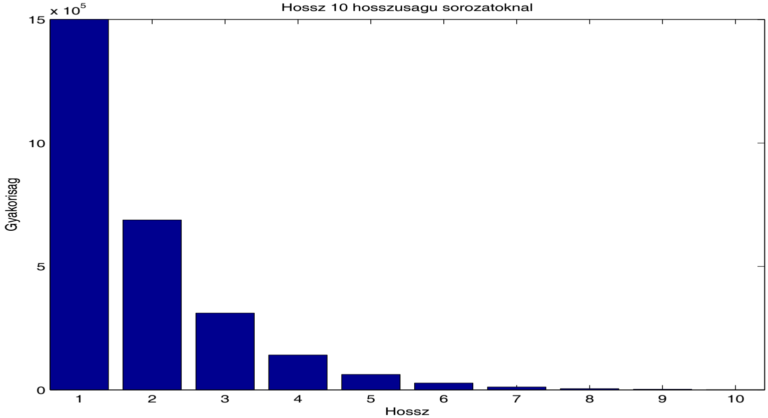
```
%% kirajzolas
clf; hold on; box on;
semilogy(bins,vBins');
xlabel('Hossz');
ylabel('Hanyszor');
ylim([0 100]);
```

```
%% BINARIS SZAMSOROZATOK HOSSZA
nRun = 500000;
nN   = [10 100 1000];
%% futasi valtozok
nBin = floor(5+ 2.5*log2(nN(end)));
vBins = zeros(length(nN),nBin);
bins   = 1:nBin;
bb     = 1;
%% kiserletek
for iSize=nN;
 for ii=1:nRun;
  ossSor = 2*round(rand(1,iSize))-1;
  ossSor(:,iSize+1) = - ossSor(:,iSize);
  valt   = ossSor(:,2:end)-ossSor(:,1:(end-1));
  valt   = abs(valt)/2;
 indF    = find(valt==1);
 hossz   = indF - [0 indF(1:end-1)];
  [bBin,bins] = hist(hossz,bins);
  vBins(bb,:) = vBins(bb,:) + bBin;
 end;
 bb=bb+1;
end;
```

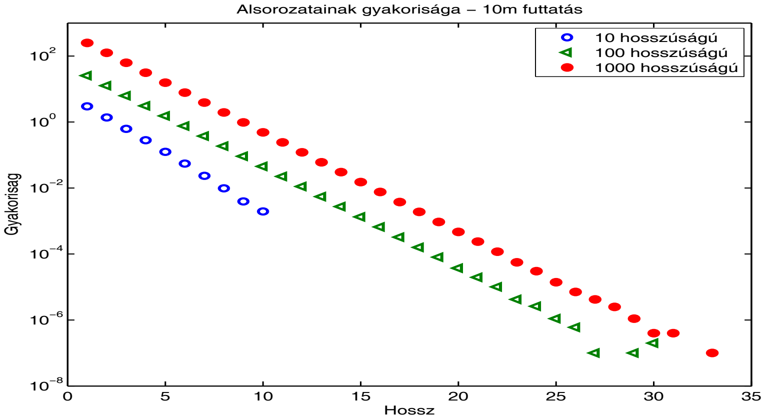
```
%% kirajzolas
clf; hold on; box on;
semilogy(bins,vBins');
xlabel('Hossz');
ylabel('Hanyszor');
ylim([0 100]);
```

```
%% BINARIS SZAMSOROZATOK HOSSZA
nRun = 500000;
nN   = [10 100 1000];
%% futasi valtozok
nBin = floor(5+ 2.5*log2(nN(end)));
vBins = zeros(length(nN),nBin);
bins  = 1:nBin;
bb    = 1;
%% kiserletek
for iSize=nN;
 for ii=1:nRun;
  ossSor = 2*round(rand(1,iSize))-1;
  ossSor(:,iSize+1) = - ossSor(:,iSize);
  valt   = ossSor(:,2:end)-ossSor(:,1:(end-1));
  valt   = abs(valt)/2;
  indF     = find(valt==1);
  hossz    = indF - [0 indF(1:end-1)];
 [bBin,bins] = hist(hossz,bins);
 vBins(bb,:) = vBins(bb,:) + bBin;
 end;
 bb=bb+1;
end;
```

```
%% kirajzolas
clf; hold on; box on;
semilogy(bins,vBins');
xlabel('Hossz');
ylabel('Hanyszor');
ylim([0 100]);
```



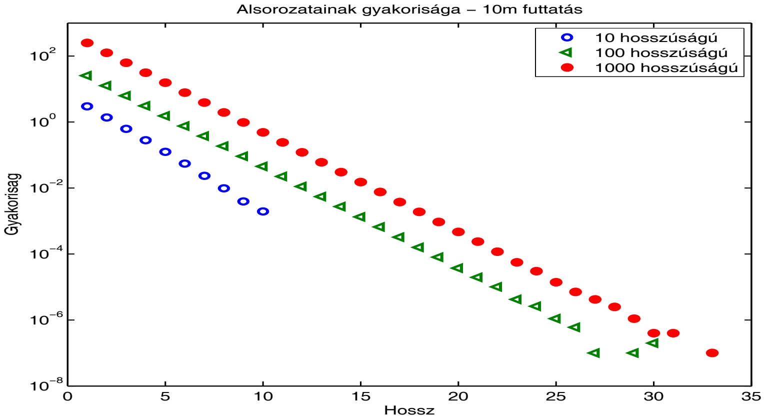
Gyakoriság és hossz közötti kapcsolat



Gyakoriság és hossz közötti kapcsolat

Sorozatok hossza

Szimuláció



Gyakoriság és hossz közötti kapcsolat

Egy adott hossz valószínűsége: $p(h) \approx c \exp\left(-k\ h\right)$

Megfogalmazás:

- Szabályos pénzérmét addig dobunk, amíg *FF* vagy *FI* nem jön ki.
- A sorozatok ugyanolyan valószínűek.

Megfogalmazás:

- Szabályos pénzérmét addig dobunk, amíg *FF* vagy *FI* nem jön ki.
- A sorozatok ugyanolyan valószínűek.

Annak a valószínűsége, hogy az *FF* hamarabb jön ki mint az *FI* egyenlő 50%-kal, ugyanis

ha fejet dobtunk, akkor 50% a valószínűsége annak, hogy a következő dobás fej lesz. Ez a valószínűség megegyezik az írásével.

Megfogalmazás:

- Szabályos pénzérmét addig dobunk, amíg *FF* vagy *FI* nem jön ki.
- A sorozatok ugyanolyan valószínűek.

Annak a valószínűsége, hogy az *FF* hamarabb jön ki mint az *FI* egyenlő 50%-kal, ugyanis

ha fejet dobtunk, akkor 50% a valószínűsége annak, hogy a következő dobás fej lesz. Ez a valószínűség megegyezik az írásével.

Mégis

Az **első előfordulás**ig tartó sorozat hossza nem ugyanaz.

Megfogalmazás:

- Szabályos pénzérmét addig dobunk, amíg *FF* vagy *FI* nem jön ki.
- A sorozatok ugyanolyan valószínűek.

Annak a valószínűsége, hogy az *FF* hamarabb jön ki mint az *FI* egyenlő 50%-kal, ugyanis

ha fejet dobtunk, akkor 50% a valószínűsége annak, hogy a következő dobás fej lesz. Ez a valószínűség megegyezik az írásével.

Mégis

Az **első előfordulás**ig tartó sorozat hossza nem ugyanaz. Az *FF*-hez átlagosan 6 dobás kell, az *FI*-hez 4.

Hogyan?

NEM azonos eseményeket keresünk:

FI esetén

$$\begin{pmatrix} +- & -+- & --+- & ---+- \\ & ++- & -++- & --++- \\ & & +++- & -+++- \\ & & & ++++- \\ 2 & 3 & 4 & 5 \\ \frac{1}{2^2} & \frac{2}{2^3} & \frac{3}{2^4} & \frac{4}{2^5} \end{pmatrix}$$

FF esetén

$$\begin{pmatrix} ++ & -++ & --++ & ---++ & ----++ \\ & & +-++ & -+-++ & --+-++ \\ & & & +--++ & -+--++ \\ & & & & +---++ \\ & & & & +-+-++ \\ 2 & 3 & 4 & 5 & 6 \end{pmatrix}$$

Az első előfordulás átlagos hossza:

$$\sum_{n=2}^{\infty} n \frac{n-1}{2^n} = 4$$

Kiszámítása:

$$\frac{1}{1-x} = \sum_{n=1}^{\infty} \frac{1}{x^n}$$

$$\left(\frac{1}{1-x}\right)''_{xx} = \sum_{n=2}^{\infty} \frac{n(n-1)}{x^{n-2}}$$

Rekurrencia relációk felállítása:

M_+ - a fejjel kezdődő sorozatok átlaghossza;

M_- - az írással kezdődő sorozatok átlaghossza;

$$M_- = 1 + \frac{M_- + M_+}{2}$$

$$M_+ = 1 + \frac{1 + M_-}{2}$$

NEM azonos eseményeket keresünk:

FI esetén

$$\begin{pmatrix} +- & -+- & --+- & ---+- \\ & ++- & -++- & --++- \\ & & +++- & -+++- \\ & & & ++++- \\ & & & \\ 2 & 3 & 4 & 5 \\ \frac{1}{2^2} & \frac{2}{2^3} & \frac{3}{2^4} & \frac{4}{2^5} \end{pmatrix}$$

Az első előfordulás átlagos hossza:

$$\sum_{n=2}^{\infty} n\,\frac{n-1}{2^n} = 4$$

Kiszámítása:

$$\frac{1}{1-x} = \sum_{n=1}^{\infty} \frac{1}{x^n}$$

$$\left(\frac{1}{1-x}\right)''_{xx} = \sum_{n=2}^{\infty} \frac{n(n-1)}{x^{n-2}}$$

FF esetén

$$\begin{pmatrix} ++ & -++ & --++ & ---++ & ----++ \\ & & +-++ & -+-++ & --+-++ \\ & & & +--++ & -+--++ \\ & & & & +---++ \\ & & & & +-+-++ \\ 2 & 3 & 4 & 5 & 6 \end{pmatrix}$$

Rekurrencia relációk felállítása:

M_+ - a fejjel kezdődő sorozatok átlaghossza;

M_- - az írással kezdődő sorozatok átlaghossza;

$$M_- = 1 + \frac{M_- + M_+}{2}$$

$$M_+ = 1 + \frac{1 + M_-}{2}$$

Az egyenletrendszer megoldása: $M_+ = 5$ és $M_- = 7$. Ahonnan:

$$\frac{M_+ + M_-}{2} = 6$$

NEM azonos eseményeket keresünk:

FI esetén

$$\left(\begin{array}{cccc} +- & -+- & --+- & ---+- \\ & ++- & -++- & --++- \\ & & +++- & -+++- \\ & & & ++++- \\ 2 & 3 & 4 & 5 \\ \frac{1}{2^2} & \frac{2}{2^3} & \frac{3}{2^4} & \frac{4}{2^5} \end{array}\right)$$

Az első előfordulás átlagos hossza:

$$\sum_{n=2}^{\infty} n\,\frac{n-1}{2^n} = 4$$

Kiszámítása:

$$\frac{1}{1-x} = \sum_{n=1}^{\infty} \frac{1}{x^n}$$

$$\left(\frac{1}{1-x}\right)''_{xx} = \sum_{n=2}^{\infty} \frac{n(n-1)}{x^{n-2}}$$

FF esetén

$$\left(\begin{array}{ccccc} ++ & -++ & --++ & ---++ & ----++ \\ & & +-++ & -+-++ & --+-++ \\ & & & +--++ & -+--++ \\ & & & & +---++ \\ & & & & +-+-++ \\ 2 & 3 & 4 & 5 & 6 \end{array}\right)$$

Rekurrencia relációk felállítása:

M_+ - a fejjel kezdődő sorozatok átlaghossza;

M_- - az írással kezdődő sorozatok átlaghossza;

$$M_- = 1 + \frac{M_- + M_+}{2}$$

$$M_+ = 1 + \frac{1 + M_-}{2}$$

Az egyenletrendszer megoldása: $M_+ = 5$ és $M_- = 7$. Ahonnan:

$$\frac{M_+ + M_-}{2} = 6$$

NEM azonos eseményeket keresünk:

FI esetén

$$\begin{pmatrix} +- & -+- & --+- & ---+- \\ & ++- & -++- & --++- \\ & & +++- & -+++- \\ & & & ++++- \\ 2 & 3 & 4 & 5 \\ \frac{1}{2^2} & \frac{2}{2^3} & \frac{3}{2^4} & \frac{4}{2^5} \end{pmatrix}$$

Az első előfordulás átlagos hossza:

$$\sum_{n=2}^{\infty} n\,\frac{n-1}{2^n} = 4$$

Kiszámítása:

$$\begin{aligned} \frac{1}{1-x} &= \sum_{n=1}^{\infty} \frac{1}{x^n} \\ \left(\frac{1}{1-x}\right)''_{xx} &= \sum_{n=2}^{\infty} \frac{n(n-1)}{x^{n-2}} \end{aligned}$$

FF esetén

$$\begin{pmatrix} ++ & -++ & --++ & ---++ & ----++ \\ & & +-++ & -+-++ & --+-++ \\ & & & +--++ & -+--++ \\ & & & & +---++ \\ & & & & +-+-++ \\ 2 & 3 & 4 & 5 & 6 \end{pmatrix}$$

Rekurrencia relációk felállítása:

M_+ - a fejjel kezdődő sorozatok átlaghossza;
M_- - az írással kezdődő sorozatok átlaghossza;

$$\begin{aligned} M_- &= 1 + \frac{M_- + M_+}{2} \\ M_+ &= 1 + \frac{1 + M_-}{2} \end{aligned}$$

NEM azonos eseményeket keresünk:

FI esetén

$$\begin{pmatrix} +- & \begin{array}{c} -+- \\ ++- \end{array} & \begin{array}{c} --+- \\ -++- \\ +++- \end{array} & \begin{array}{c} ---+- \\ --++- \\ -+++- \\ ++++- \end{array} \\ 2 & 3 & 4 & 5 \\ \frac{1}{2^2} & \frac{2}{2^3} & \frac{3}{2^4} & \frac{4}{2^5} \end{pmatrix}$$

Az első előfordulás átlagos hossza:

$$\sum_{n=2}^{\infty} n \frac{n-1}{2^n} = 4$$

Kiszámítása:

$$\frac{1}{1-x} = \sum_{n=1}^{\infty} \frac{1}{x^n}$$

$$\left(\frac{1}{1-x}\right)''_{xx} = \sum_{n=2}^{\infty} \frac{n(n-1)}{x^{n-2}}$$

FF esetén

$$\begin{pmatrix} ++ & -++ & \begin{array}{c} --++ \\ +-++ \end{array} & \begin{array}{c} ---++ \\ -+-++ \\ +--++ \end{array} & \begin{array}{c} ----++ \\ --+-++ \\ -+--++ \\ +---++ \\ +-+-++ \end{array} \\ 2 & 3 & 4 & 5 & 6 \end{pmatrix}$$

Rekurrencia relációk felállítása:

M_+ - a fejjel kezdődő sorozatok átlaghossza;
M_- - az írással kezdődő sorozatok átlaghossza;

$$M_- = 1 + \frac{M_- + M_+}{2}$$
$$M_+ = 1 + \frac{1 + M_-}{2}$$

Az egyenletrendszer megoldása: $M_+ = 5$ és $M_- = 7$. Ahonnan:

$$\frac{M_+ + M_-}{2} = 6$$

Első előfordulás II

Valószínűségi paradoxonok
Csató Lehel
Bevezető
Matematikai fogalmak
Valószínűségi változók
MATLAB
Paradoxonok
Lotto
Függetlenség
Ajándékozás
Pétervár
Játékelmélet
Halandóság
Bernoulli
Első előfordulás
De Moivre

NEM azonos eseményeket keresünk:

FI esetén

$$\begin{pmatrix} +- & -+- & --+- & ---+- \\ & ++- & -++- & --++- \\ & & +++- & -+++- \\ & & & ++++- \\ & & & \\ 2 & 3 & 4 & 5 \\ & & & \\ \frac{1}{2^2} & \frac{2}{2^3} & \frac{3}{2^4} & \frac{4}{2^5} \end{pmatrix}$$

Az első előfordulás átlagos hossza:

$$\sum_{n=2}^{\infty} n\,\frac{n-1}{2^n} = 4$$

Kiszámítása:

$$\begin{aligned} \frac{1}{1-x} &= \sum_{n=1}^{\infty} \frac{1}{x^n} \\ \left(\frac{1}{1-x}\right)''_{xx} &= \sum_{n=2}^{\infty} \frac{n(n-1)}{x^{n-2}} \end{aligned}$$

FF esetén

$$\begin{pmatrix} ++ & -++ & --++ & ---++ & ----++ \\ & & +-++ & -+-++ & --+-++ \\ & & & +--++ & -+--++ \\ & & & & +---++ \\ & & & & +-+-++ \\ & & & & \\ 2 & 3 & 4 & 5 & 6 \end{pmatrix}$$

Rekurrencia relációk felállítása:

M_+ - a fejjel kezdődő sorozatok átlaghossza;
M_- - az írással kezdődő sorozatok átlaghossza;

$$\begin{aligned} M_- &= 1 + \frac{M_- + M_+}{2} \\ M_+ &= 1 + \frac{1 + M_-}{2} \end{aligned}$$

Az egyenletrendszer megoldása: $M_+ = 5$ és $M_- = 7$. Ahonnan:

$$\frac{M_+ + M_-}{2} = 6$$

NEM azonos eseményeket keresünk:

FI esetén

$$\begin{pmatrix} +- & -+- & --+- & ---+- \\ & ++- & -++- & --++- \\ & & +++- & -+++- \\ & & & ++++- \\ 2 & 3 & 4 & 5 \\ \frac{1}{2^2} & \frac{2}{2^3} & \frac{3}{2^4} & \frac{4}{2^5} \end{pmatrix}$$

Az első előfordulás átlagos hossza:

$$\sum_{n=2}^{\infty} n\,\frac{n-1}{2^n} = 4$$

Kiszámítása:

$$\frac{1}{1-x} = \sum_{n=1}^{\infty} \frac{1}{x^n}$$

$$\left(\frac{1}{1-x}\right)''_{xx} = \sum_{n=2}^{\infty} \frac{n(n-1)}{x^{n-2}}$$

FF esetén

$$\begin{pmatrix} ++ & -++ & --++ & ---++ & ----++ \\ & & +-++ & -+-++ & --+-++ \\ & & & +--++ & -+--++ \\ & & & & +---++ \\ & & & & +-+-++ \\ 2 & 3 & 4 & 5 & 6 \end{pmatrix}$$

Rekurrencia relációk felállítása:

M_+ - a fejjel kezdődő sorozatok átlaghossza;
M_- - az írással kezdődő sorozatok átlaghossza;

$$M_- = 1 + \frac{M_- + M_+}{2}$$
$$M_+ = 1 + \frac{1 + M_-}{2}$$

Az egyenletrendszer megoldása: $M_+ = 5$ és $M_- = 7$. Ahonnan:

$$\frac{M_+ + M_-}{2} = 6$$

NEM azonos eseményeket keresünk:

FI esetén

$$\begin{pmatrix}
+- & -+- & --+- & ---+- \\
 & ++- & -++- & --++- \\
 & & +++- & -+++- \\
 & & & ++++- \\
2 & 3 & 4 & 5 \\
\frac{1}{2^2} & \frac{2}{2^3} & \frac{3}{2^4} & \frac{4}{2^5}
\end{pmatrix}$$

Az első előfordulás átlagos hossza:

$$\sum_{n=2}^{\infty} n\,\frac{n-1}{2^n} = 4$$

Kiszámítása:

$$\frac{1}{1-x} = \sum_{n=1}^{\infty} \frac{1}{x^n}$$

$$\left(\frac{1}{1-x}\right)''_{xx} = \sum_{n=2}^{\infty} \frac{n(n-1)}{x^{n-2}}$$

FF esetén

$$\begin{pmatrix}
++ & -++ & --++ & ---++ & ----++ \\
 & & +-++ & -+-++ & --+-++ \\
 & & & +--++ & -+--++ \\
 & & & & +---++ \\
 & & & & +-+-++ \\
2 & 3 & 4 & 5 & 6
\end{pmatrix}$$

Rekurrencia relációk felállítása:

M_+ - a fejjel kezdődő sorozatok átlaghossza;

M_- - az írással kezdődő sorozatok átlaghossza;

$$\begin{aligned}
M_- &= 1 + \frac{M_- + M_+}{2} \\
M_+ &= 1 + \frac{1 + M_-}{2}
\end{aligned}$$

Az egyenletrendszer megoldása: $M_+ = 5$ és $M_- = 7$. Ahonnan:

$$\frac{M_+ + M_-}{2} = 6$$

Megfogalmazás:
Szabályos pénzérmét átlagosan legalább 8-szor dobunk ahhoz, hogy az $+,-$ sorozat valamely háromtagú permutációját kapjuk.

Az egyéni valószínűségek nem egyformák

Az $++-, -++, --+, +--$ kombinációk a legkedvezőbbek.

Megfogalmazás:
Szabályos pénzérmét átlagosan legalább 8-szor dobunk ahhoz, hogy az $+,-$ sorozat valamely háromtagú permutációját kapjuk.

Az egyéni valószínűségek nem egyformák

Az $++-$, $-++$, $--+$, $+--$ kombinációk a legkedvezőbbek.

Példa: vizsgáljuk meg, hogy az $++-$ és az $-++$ sorozatok közül melyik előfordulása valószínűbb korábban.

$$\begin{pmatrix} ++- & +++- & ++++- & +++++- & \dots \\ \frac{1}{2^3} & \frac{1}{2^4} & \frac{1}{2^5} & \frac{1}{2^6} & \dots \end{pmatrix}$$

Az összes többi esetben a $-++$ fordul elő előbb (egy $-$ mindig van; a $++-$-t pedig mindenképp megelőzi egy $-++$).

Megfogalmazás:
Szabályos pénzérmét átlagosan legalább 8-szor dobunk ahhoz, hogy az $+,-$ sorozat valamely háromtagú permutációját kapjuk.

Az egyéni valószínűségek nem egyformák

Az $++-, -++, --+, +--$ kombinációk a legkedvezőbbek.

Példa: vizsgáljuk meg, hogy az $++-$ és az $-++$ sorozatok közül melyik előfordulása valószínűbb korábban. Mi az eseménytér?

$$\begin{pmatrix} ++- & +++- & ++++- & +++++- & \dots \\ \frac{1}{2^3} & \frac{1}{2^4} & \frac{1}{2^5} & \frac{1}{2^6} & \dots \end{pmatrix}$$

Az összes többi esetben a $-++$ fordul elő előbb (egy $-$ mindig van; a $++-$-t pedig mindenképp megelőzi egy $-++$).

Megfogalmazás:
Szabályos pénzérmét átlagosan legalább 8-szor dobunk ahhoz, hogy az $+,-$ sorozat valamely háromtagú permutációját kapjuk.

Az egyéni valószínűségek nem egyformák

Az $++-$, $-++$, $--+$, $+--$ kombinációk a legkedvezőbbek.

Példa: vizsgáljuk meg, hogy az $++-$ és az $-++$ sorozatok közül melyik előfordulása valószínűbb korábban. Mi az eseménytér?

$$\begin{pmatrix} ++- & +++- & ++++- & +++++- & \dots \\ \frac{1}{2^3} & \frac{1}{2^4} & \frac{1}{2^5} & \frac{1}{2^6} & \dots \end{pmatrix}$$

Az összes többi esetben a $-++$ fordul elő előbb (egy $-$ mindig van; a $++-$-t pedig mindenképp megelőzi egy $-++$).

Megfogalmazás:
Szabályos pénzérmét átlagosan legalább 8-szor dobunk ahhoz, hogy az $+,-$ sorozat valamely háromtagú permutációját kapjuk.

Az egyéni valószínűségek nem egyformák

Az $++-$, $-++$, $--+$, $+--$ kombinációk a legkedvezőbbek.

Példa: vizsgáljuk meg, hogy az $++-$ és az $-++$ sorozatok közül melyik előfordulása valószínűbb korábban. Mi az eseménytér?

$$\begin{pmatrix} ++- & +++- & ++++- & +++++- & \cdots \\ \frac{1}{2^3} & \frac{1}{2^4} & \frac{1}{2^5} & \frac{1}{2^6} & \cdots \end{pmatrix}$$

Az összes többi esetben a $-++$ fordul elő előbb (egy $-$ mindig van; a $++-$-t pedig mindenképp megelőzi egy $-++$).

Megfogalmazás:
Szabályos pénzérmét átlagosan legalább 8-szor dobunk ahhoz, hogy az $+,-$ sorozat valamely háromtagú permutációját kapjuk.

Az egyéni valószínűségek nem egyformák

Az $++-$, $-++$, $--+$, $+--$ kombinációk a legkedvezőbbek.

Példa: vizsgáljuk meg, hogy az $++-$ és az $-++$ sorozatok közül melyik előfordulása valószínűbb korábban. Mi az eseménytér?

$$\begin{pmatrix} ++- & +++- & ++++- & +++++- & \dots \\ \frac{1}{2^3} & \frac{1}{2^4} & \frac{1}{2^5} & \frac{1}{2^6} & \dots \end{pmatrix}$$

Az összes többi esetben a $-++$ fordul elő előbb (egy $-$ mindig van; a $++-$-t pedig mindenképp megelőzi egy $-++$).

Megfogalmazás:
Szabályos pénzérmét átlagosan legalább 8-szor dobunk ahhoz, hogy az $+,-$ sorozat valamely háromtagú permutációját kapjuk.

Az egyéni valószínűségek nem egyformák

Az $++-$, $-++$, $--+$, $+--$ kombinációk a legkedvezőbbek.

Példa: vizsgáljuk meg, hogy az $++-$ és az $-++$ sorozatok közül melyik előfordulása valószínűbb korábban. Mi az eseménytér?

$$\begin{pmatrix} ++- & +++- & ++++- & +++++- & \dots \\ \frac{1}{2^3} & \frac{1}{2^4} & \frac{1}{2^5} & \frac{1}{2^6} & \dots \end{pmatrix}$$

Az összes többi esetben a $-++$ fordul elő előbb (egy $-$ mindig van; a $++-$-t pedig mindenképp megelőzi egy $-++$).
A valószínűség:

$$\sum_{n=3}^{\infty} \frac{1}{2^n} = \frac{1}{8} \sum_{n=0}^{\infty} \frac{1}{2^n} = \frac{1}{8}\, 2 = \frac{1}{4}$$

- **De Moivre** *1718: The doctrine of chances.*

Nagy számok ellentmondása

- a fej/írás arány konvergál 1-hez – a közelítő egyenlőség;
- a „fej”=”írás” esemény valószínűsége konvergál nullához – a pontos egyenlőség.

$$\lim_{n\to\infty} P\left(\left\|\frac{n_{Fej}}{n_{Iras}} - 1\right\| < \epsilon\right) = 1$$

illetve

$$\lim_{n\to\infty} P\left(n_{Fej} = n_{Iras}\right) = 0$$

Normális eloszlás

- **De Moivre** *1718: The doctrine of chances.*

Nagy számok ellentmondása

- a fej/írás arány konvergál 1-hez – a közelítő egyenlőség;
- a „fej”=”írás” esemény valószínűsége konvergál nullához – a pontos egyenlőség.

$$\lim_{n\to\infty} P\left(\left\|\frac{n_{Fej}}{n_{Iras}} - 1\right\| < \epsilon\right) = 1$$

illetve

$$\lim_{n\to\infty} P\left(n_{Fej} = n_{Iras}\right) = 0$$

Normális eloszlás II

De Moivre (1738): „Merem állítani, hogy ez a legnehezebb probléma, amit fel lehet vetni a Véletlen Tudományában..."
Legyen az érme: $x_n \in \{-1, 1\}$ úgy, hogy

$$p(x_n = 1) = p(x_n = -1) = 0.5.$$

Ekkor a dobás-sorozatok átlaga 0 és szórása:

$$\sigma_B = E\left[(x_n - \bar{x}_n)^2\right] = 1$$

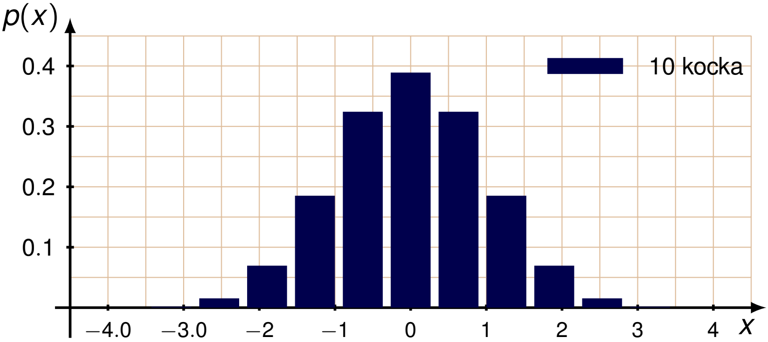
Ha 6 dobás esetét vizsgáljuk, a lehetséges (fej,érme) párok, illetve összeg:

$$\begin{bmatrix} (6,0) & (5,1) & \dots & (0,6) \\ -6 & -4 & \dots & 6 \end{bmatrix}$$

Normális eloszlás III

A dobások összege és azok valószínűsége:

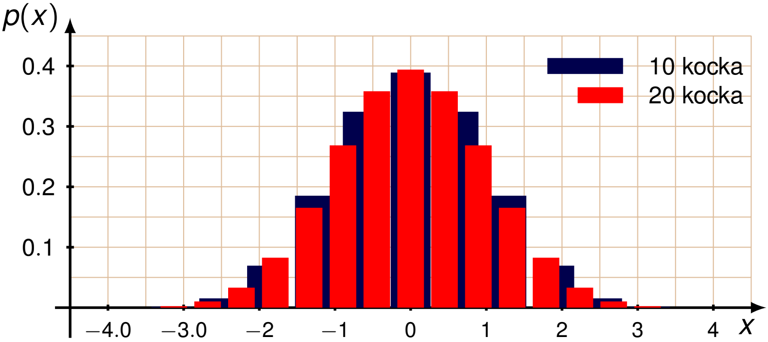
$$\begin{bmatrix} -10 & -8 & -6 & \dots & 10 \\ \frac{\binom{10}{0}}{2^{10}} & \frac{\binom{10}{1}}{2^{10}} & \frac{\binom{10}{2}}{2^{10}} & \dots & \frac{\binom{10}{10}}{2^{10}} \end{bmatrix}$$



Normális eloszlás III

A dobások összege és azok valószínűsége:

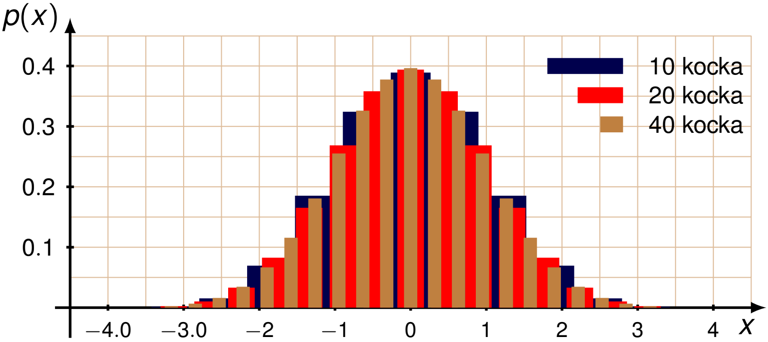
$$\begin{bmatrix} -10 & -8 & -6 & \dots & 10 \\ \frac{\binom{10}{0}}{2^{10}} & \frac{\binom{10}{1}}{2^{10}} & \frac{\binom{10}{2}}{2^{10}} & \dots & \frac{\binom{10}{10}}{2^{10}} \end{bmatrix}$$



Normális eloszlás III

A dobások összege és azok valószínűsége:

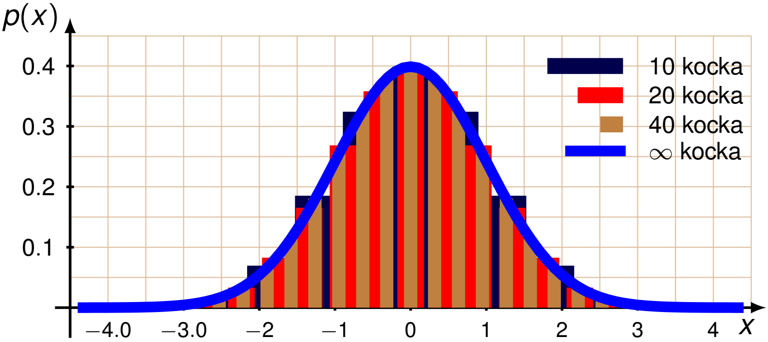
$$\begin{bmatrix} -10 & -8 & -6 & \dots & 10 \\ \frac{\binom{10}{0}}{2^{10}} & \frac{\binom{10}{1}}{2^{10}} & \frac{\binom{10}{2}}{2^{10}} & \dots & \frac{\binom{10}{10}}{2^{10}} \end{bmatrix}$$



Normális eloszlás III

A dobások összege és azok valószínűsége:

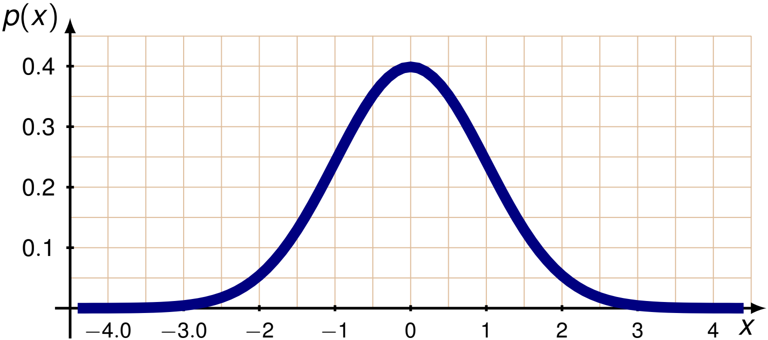
$$\begin{bmatrix} -10 & -8 & -6 & \dots & 10 \\ \frac{\binom{10}{0}}{2^{10}} & \frac{\binom{10}{1}}{2^{10}} & \frac{\binom{10}{2}}{2^{10}} & \dots & \frac{\binom{10}{10}}{2^{10}} \end{bmatrix}$$



Normális eloszlás IV

Az eloszlás képlete:

$$p(x) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{x^2}{2}\right]$$



```
function [x,y] = moivre(N,style)
%% valasztunk parametereket
if nargin==0;
  N = 6;
elseif nargin==1;
  style='k';
end;
%% X- es Y- tengely
x = 0:N;
y = zeros(size(x));
for ii=x;
  y(ii+1) = nchoosek(N,ii);
end;
y = y ./ (2^N);
x = 2*x - N;
%% normalissa valtoztatni
x = x/sqrt(N);
y = y/2*sqrt(N);
%% kirajzolni
box on; hold on;
bar(x,y,style);
xlim([-4.5,4.5]);
%% tul kicsi ertekek
ii = find(y>0.001);
x  = x(ii);
y  = y(ii);
```

```
[x,y]=moivre(40,'k');
h = fopen('moivre_40.dat','wt');
fprintf(h,'%5.4f %5.4f/n',[x' y']');
fclose(h);
```

Irodalom I

Székely J. Gábor
Paradoxonok a véletlen matematikájában.
Typotex, 2004.