A41ea

2021. szeptember 9., csütörtök 8:02

A41ea

A4 Valószínűségszámítás — I. EA

dr. Keszthelyi Gabriella
Sztochasztika és Dinamikai rendszerek tanszék

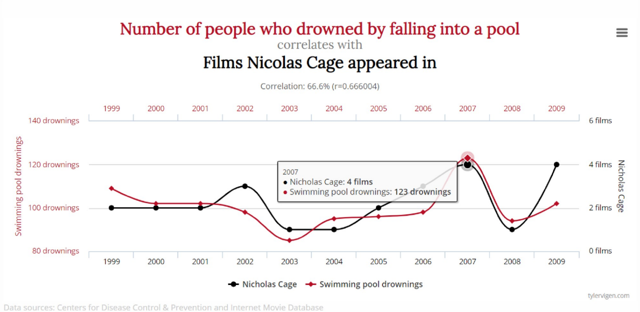
2021. szeptember 9.

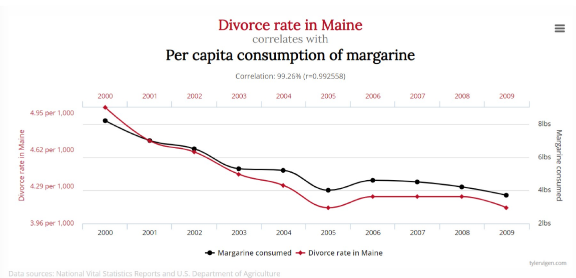
Követelmények

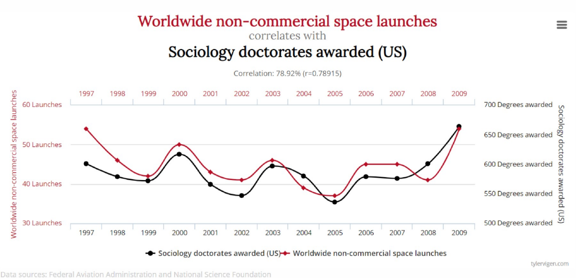
- 3 kis zh és 1 nagy zh. A kis zh-k közül egyet (a legrosszabbat) figyelmen kívül hagyunk.
- Az aláíráshoz a zh-kból min. 40%-ot kell teljesíteni (a nagy zh-ból külön, a kis zh-kból együtt).
- A zh-k 40%-os, a vizsga 60%-os súllyal számít be a jegybe. A vizsgán is el kell érni min. 40%-ot.
- Amennyiben félév közben online-ra váltunk, akkor a zh-kat házik fogják kiváltani, ugyanazzal a súlyozással.

"Éppen mivel a bizonytalanságról van szó, várható, hogy felületes elmék azt fogják hinni, hogy e téren nem kell törekedni a gondolkodás lehető legteljesebb tisztaságára. Nem lepne meg, hogy olyanok, akik a szabatos matematikai gondolkodástól amúgy is idegenkednek, azt hinnék, hogy - mivel a véletlen jelenségeket úgysem lehet teljes pontossággal előrelátni, legfeljebb nagy vonalakban – e jelenségek matematikai vizsgálatánál megengedhetik maguknak a pongyolaságot, a csak félig átgondolt és csak negyedrészt megemésztett fogalmakkal való dobálózást."

Mire használjuk a valószínűségszámítást?







Mi a valószínűség?

"60% az esélye, hogy tudunk olajsejkséget alapítani az Alföldön"

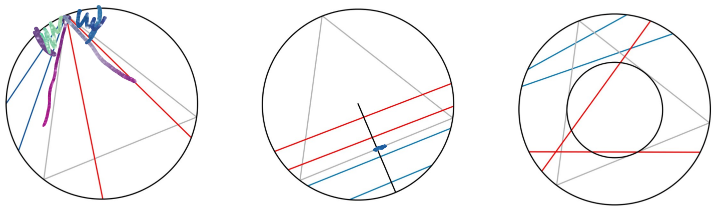
Iterpretációk:

1. Az Alföldhöz hasonló környezeti tulajdonságokkal rendelkező régiók 60%-ban van olaj. OBJEKTÍV (a kimenetel tulajdonsága, a megfigyelt esetek hányadrészében következik be az esemény?)

2. A geológus személyes intuíciója szerint 60% az esélye, hogy ha most elkezdünk fúrni, akkor találunk olajat. SZUBJEKTÍV (milyen arányú fogadást kötnénk rá?)

Bertrand paradoxon

Mi a valószínűsége, hogy egy random húr hosszabb lesz mint a háromszög oldala?



random végpont $\frac{1}{3}$ random sugárpont $\frac{1}{2}$ random belső pont $\frac{1}{4}$

Eseménytér

"Tételezzük fel, hogy valamely véletlen folyamatnak sok, egyformán valószínű végkimenetele van, ezek közül egyesek kedvezőek (vagyis nyerők), mások kedvezőtlenek (vagyis veszteséget jelentenek) számunkra. Ekkor a nyerő végeredmény elérésének valószínűsége egyenlő a számunka kedvező eredmények részarányával. A lehetséges eredmények összességét eseménytérnek nevezzük."
Cardano *A szerencsejátékról*

Eseménytér

Kísérletek — elemi események

- Kocka: {1,2,3,4,5,6}
- Érme: {F,I}
- 2 érme: {FF,II,FI,IF}
- Érmedobálás, amíg fej nem lesz: $\{1, 2, 3, \ldots \infty\}$
- A4-re ráfordított tanulás ideje: $(0, \infty)$

Események mint halmazok - De Morgan azonosságok

Eseményalgebra

Eseményalgebra

- Lehetetlen esemény $\emptyset$
- Biztos esemény S
- Komplementer/kizáró események $A, \bar{A}$
- Események metszete $A \cap B$
- Események uniója $A \cup B$
- Egyik esemény maga után vonja a másikat $A \subset B$

STANDARD KOCKA: 7. [illegible]

$\{1,2,3,4,5,6\}$

$A = \{1;2\}$ $\bar{A} = \{3,4,5,6\}$

$A = \{1;2\}$ $B = \{2;4;6\}$

PROBABILITY

$A \cap B = \{2\}$ $P(A \cap B) = \frac{1}{6}$

EGYÜTTES BEKÖVETKEZÉS

$B = \{1 \text{ vagy } 3 \text{ vagy } 5\}$

$C = \{1\}$

$C = \{2;4\}$

$C \subset B$

$A \cup B = \{1,2,4,6\}$

LEGALÁBB AZ EGYIK BEKÖVETKEZIK

Információelmélet

161-an vették fel az órát, ha kiválasztok valakit a névsorról, legalább hány kérdésből lehet biztosan kitalálni, hogy kiről van szó?

$H = \{x_1, x_2, \dots x_N\},\ p_1 = p_2 = \dots = p_N = 1/N$
N db szimbólum, mind azonos valószínűséggel szerepel. Hány biten tudom lekódolni őket?

$161 = 10101101$

$\underbrace{\log_2 128}_{7} < \underbrace{\log_2 161}_{7,\dots} < \underbrace{\log_2 256}_{8}$

Információelmélet

Két pénzérmével dobunk, jelentse ξ, hogy hány fej lesz közöttük. Ekkor $S = \{0, 1, 2\}$ a valószínűségek pedig
$P(\xi = 0) = 1/4, P(\xi = 1) = 1/2, P(\xi = 2) = 1/4$
Hogyan kódoljuk el (hossz)optimálisan a dobássorozatot?

- $\xi = 1 \to 1$
- $\xi = 0 \to 00$
- $\xi = 2 \to 01$

Várható kódhossz

$$\frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 2 + \frac{1}{4} \cdot 2 = 1,5$$

Információelmélet

$H(\xi)$ Információs függvény

(Hartley)

$$H(\xi) = \log_2 N$$

Ha p_i-k különbözőek, akkor

(Shannon)

$$H(\xi) = p_1 \log_2(1/p_1) + p_2 \log_2(1/p_2) + \dots + p_N \log_2(1/p_N)$$

Információelmélet

$H(\xi)$ vs minimum bitszám?

$H(\xi)$ általában nem egész és egy átlagos kódhosszt ad meg, a szükséges bitek számára egy alsó korlátot.

(Boltzmann) Egy nagyszámú molekulából álló gázban az egyes molekulák lehetséges állapotainak valószínűségei rendre $p_1, p_2, \dots p_N$, akkor a rendszer entrópiája

$$H = c(p_1 \log_2(1/p_1) + p_2 \log_2(1/p_2) + \dots p_N \log_2(1/p_N))$$

Azaz a rendszer entrópiája a molekulái állapotára vonatkozó *bizonytalanság* mértékszáma.

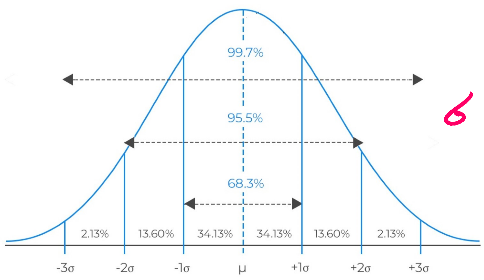
Statisztika, mérési hibák

A XVIII. század végén Franciaországban a fizikában két irányzat uralkodott: a *matematikusoké*, akik Newton mozgás- és gravitációs törvényeit tanulmányozták és a *kísérleti filozófusoké*, akik tapasztalati úton vizsgálták az elektromosság, mágnesesség, fény és hő jelenségeit. Ez utóbbiból nőtte ki magát a kísérleti fizika és később a frekventista statisztika.

Statisztika, mérési hibák

A következő problémát a mérések különbözősége jelentette: Bessel 11 csoportba sorolta a különböző mérési hibákat. A csillagászoknak ki kellett követeztetniük, hogy mérések sorozatából hogyan tudják kiszámítani egy égitest valódi helyét. A hibák eloszlása egyetemes törvényszerűséget követ, ezt a hibatörvényt írja le a haranggörbe.

Haranggörbe AKA normális eloszlás



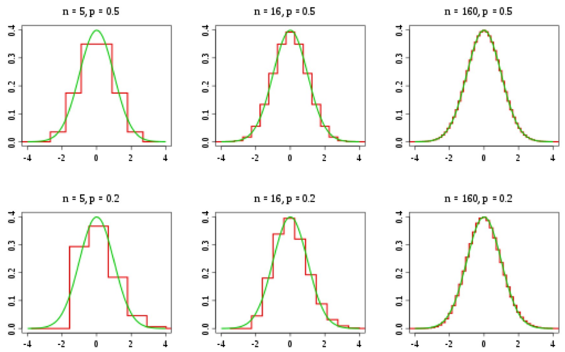
σ = szórás

átlag / várható érték

Centrális Határeloszlás Tétele (CHT)

A mérési hibáim együttesen kioltják vagy felerősítik egymást?

Az egymástól független mérések hibáinak (in fact tetszőleges független valószínűségi változók) összege normális eloszláshoz tart.



2 kocka dobása

2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12

$P(2) = \frac{1}{36}$ $\quad$ $P(12) = \frac{1}{36}$

$P(3) = \frac{2}{36}$ $\quad$ $P(11) = \frac{2}{26}$

$P(4) = \frac{3}{36}$

$\vdots$

És ha nem függetlenek az események?

Annak a valószínűsége, hogy egy random kiválasztott személy **elmebeteg** és annak a valószínűsége, hogy egy random személy **azt hiszi, hogy a felesége tud olvasni a gondolataiban** egyaránt kicsi.
Viszont sokkal nagyobb annak valószínűsége, hogy

egy random személy elmebeteg, ***ha*** *azt hiszi, hogy a felesége tud olvasni a gondolataiban,*

mintsem annak, hogy

egy random személy azt hiszi, hogy a felesége tud olvasni a gondolataiban, ***ha*** *az illető elmebeteg.*

Bayes-i típusú tévedések

A MÁSODIK OLTÁS UTÁNI FERTŐZÉSEK ÉS HALÁLOZÁSOK ADATAI

2020. DECEMBER 26. ÉS 2021. ÁPRILIS 20. KÖZÖTT

Vakcinák	Immunizáció utáni események			
	megbetegedettek		elhunytak	
	száma	százezer oltottra vetített száma	száma	százezer oltottra vetített száma
Szputnyik-V	201	95	2	1
Moderna	108	177	12	20
Sinopharm	1744	356	78	16
Pfizer/BioNTech	3048	555	175	32
AstraZeneca	613	700	6	7
Összesen	**5714**	**408**	**273**	**20**

Bayes-i típusú tévedések

A MÁSODIK OLVASÁS UTÁNI FERTŐZÉSEK ÉS HALÁLOZÁSOK ADATAI

2020. december 26. és 2021. április 20. között

Olvasmányok	Olvasás utáni események			
	megbetegedések		elhunytak	
	száma	százezer olvasóra vetített száma	száma	százezer olvasóra vetített száma
Dörmögő Dömötör	200	100	2	1
Ifjúsági Magazin	300	200	3	2
Nemzeti Sport	3000	500	30	5
Életet az éveknek	20000	2000	200	20
Nyugdíjas Újság	60000	3000	600	30
Összesen	83500	1000	835	10

Bayesiánus statisztika

$$\underbrace{P(A|B)}_{\text{aposteriori}} = \frac{\overbrace{P(B|A)}^{\text{likelihood}}\ \overbrace{P(A)}^{\text{apriori}}}{P(B)}$$

Thomas Bayes
1702 - 1761

Bayes-i módszert akkor alkalmazzuk, amikor van valamilyen előzetes apriori tudásunk vagy feltételezésünk, és ezt belekalkuláljuk a becslésünkbe.

Hogyan változtassuk meg a becslésünket új információk fényében?

Bayesiánus statisztika alkalmazása

- Kódfejtés (Alan Turing Bayes-i modellekel fejtette meg a német hajók jeleit)
- Statisztikai mintavételezés (Monte Carlo módszerek)
- Természetes Nyelv Feldolgozás (NLP)

Pascal háromszög

```
              1
            1   1
          1   2   1
        1   3   3   1
      1   4   6   4   1
    1   5  10  10   5   1
  1   6  15  20  15   6   1
1   7  21  35  35  21   7   1
```

$$\binom{0}{0}$$
$$\binom{1}{0}\binom{1}{1}$$
$$\binom{2}{0}\binom{2}{1}\binom{2}{2}$$
$$\binom{3}{0}\binom{3}{1}\binom{3}{2}\binom{3}{3}$$
$$\binom{4}{0}\binom{4}{1}\binom{4}{2}\binom{4}{3}\binom{4}{4}$$
$$\binom{5}{0}\binom{5}{1}\binom{5}{2}\binom{5}{3}\binom{5}{4}\binom{5}{5}$$
$$\binom{6}{0}\binom{6}{1}\binom{6}{2}\binom{6}{3}\binom{6}{4}\binom{6}{5}\binom{6}{6}$$
$$\binom{7}{0}\binom{7}{1}\binom{7}{2}\binom{7}{3}\binom{7}{4}\binom{7}{5}\binom{7}{6}\binom{7}{7}$$

$$(a+b)^7 = a^7 + 7a^6b + 21a^5b^2 + 35a^4b^3 + 35a^3b^4 + 21a^2b^5 + 7ab^6 + b^7$$

Kombinatorikai alapképletek

	ismétlés nélküli	ismétléses
permutáció	$n!$ n futó beérkezésének sorrendje	$\frac{n!}{k_1!k_2!\ldots k_r!}$ n golyót ennyiféleképpen állíthatunk sorba, ha $k_1, k_2, \ldots k_r$ db külön-külön egyszínű
variáció	$\frac{n!}{(n-k)!}$ n futó beérkezésének sorrendje ha csak az első k helyet tekintjük	l^k l darab betűből készíthető k hosszú szavak száma (egy-egy betű többször is felhasználható)
kombináció	$\binom{n}{k} = \frac{n!}{k!(n-k)!}$ n golyóból kiválasztunk k darabot (nem számít a kiválasztás sorrendje)	$\binom{k+l-1}{l}$ k féle sütiből (sok van belőlük) hazaviszünk l -et, ennyiféleképpen tehetjük meg

A,B,C,D,E

$5! = 5\cdot4\cdot3\cdot2\cdot1$

$\underbrace{}_{5}\ \underbrace{}_{4}\ \underbrace{}_{3}\ \underbrace{}_{2}\ \underbrace{}_{1}$

AAA,B,B

$\frac{5!}{3!\,2!} = \binom{5}{2} = \binom{5}{3}$

kombináció

$\binom{n}{k} = \frac{n!}{(n-k)!\,k!}$

$\binom{k+l-1}{l}$

variáció

1,2,3,4,5,6

6!

Irodalomjegyzék

Irodalomjegyzék

- Vetier András—Valószínűségszámítás
- Rényi Alfréd—Ars Mathematica
- Leonard Mlodinow—Drunkard's Walk: How Randomness Rules Our Lives
- Székely J. Gábor —Paradoxonok a véletlen matematikájában

Köszönöm a figyelmet!

1, 2, 3, 4, 5, 6

$\underbrace{}_{6}\underbrace{}_{5}\underbrace{}_{4}\underbrace{}_{3} = \frac{6!}{(6-4)!}$

$\frac{n!}{(n-k)!}$

ISMÉTLÉSES VARIÁCIÓ

1, 2, 3, 4, 5, 6

akárhányszor kihúzhatom a számjegyet

$\underbrace{}_{6}\underbrace{}_{6}\underbrace{}_{6}\underbrace{}_{6}$ 6^4

n^k