Valószínűségszámítás
2. előadás

Arató Miklós

Esély

Várható

Kockázat

Véletlen

Hiba

Előrejelzés

ELTE TTK Valószínűségelméleti és Statisztika Tanszék

Feltételes valószínűség

- Amennyiben $P(B) > 0$, akkor az A esemény B feltétel melletti feltételes valószínűsége:
$$P(A|B) := \frac{P(AB)}{P(B)}$$
- Kombinatorikus valószínűségi mező esetén:

$$P(A|B) = \frac{\frac{|AB|}{|\Omega|}}{\frac{|B|}{|\Omega|}} = \frac{|AB|}{|B|}$$

Feltételes valószínűség (folyt.)

- Amennyiben $P(B) > 0$, akkor $(\Omega, \mathcal{A}, P(.|B))$ Kolmogorov-féle valószínűségi mező lesz.

$$\Longrightarrow$$

- A „"normál" valószínűségre bizonyítottak a feltételes valószínűségre is teljesülnek.

Az ügyvédnek vagy ügyésznek volt igaza ?

- Nicole Brown-t 1994-ben gyilkolták meg. Férjét, O.J. Simpsont, gyanúsították meg.
- Az ügyész külön hangsúlyozta, hogy Simpson korábban már bántalmazta feleségét.
- Az ügyvéd válaszában arra hivatkozott, hogy a statisztikák szerint "csak" minden 100-adik bántalmazó férj öli meg feleségét. Valójában a házastársuk/élettársuk által bántalmazottak közül "csak" minden 2500-adikat öli meg házastársuk/élettársuk.

Az ügyvédnek vagy ügyésznek volt igaza ? (folyt.)

- Valójában a helyes kérdés az, hogy milyen valószínűséggel ölte meg a feleségét a férj, ha a feleséget megölték és ő korábban bántalmazta a feleséget?
- 1/20000 annak az esélye, hogy valakit megölnek az USA-ban.
- Csak a bántalmazottak körét tekintjük.
- B: férj öli meg, C: nem férj öli meg
- Kérdés: $P(B|B \cup C) =?$

$$P(B|B \cup C) = \frac{P(B)}{P(B \cup C)} = \frac{\frac{1}{2500}}{\frac{1}{2500} + \frac{1}{20000}} = \frac{8}{9}$$

Bayes-formula

- Tegyük fel, hogy $P(B) > 0, 0 < P(A) < 1$, ekkor

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\bar{A})P(\bar{A})}$$

Bizonyítás:

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{P(BA)}{P(BA)+P(B\bar{A})} =$$

$$= \frac{\frac{P(BA)}{P(A)}P(A)}{\frac{P(BA)}{P(A)}P(A) + \frac{P(B\bar{A})}{P(\bar{A})}P(\bar{A})}$$

Mammográfiás vizsgálat

- 50 éves nő (tünetek nélkül) rutin mammográfiás vizsgálaton vett részt.
- A teszt pozitív lett.
- Mennyi a valószínűsége, hogy emlődaganata van?
- Adatok:
  - 50 éves nőnél emlődaganat valószínűsége 1%
  - Emlődaganat felismerésének valószínűsége 90%
  - Emlődaganat nélkül pozitív teszt valószínűsége 9%

Melyikhez van közelebb a valószínűség?

- 1%
- 10%
- 90%

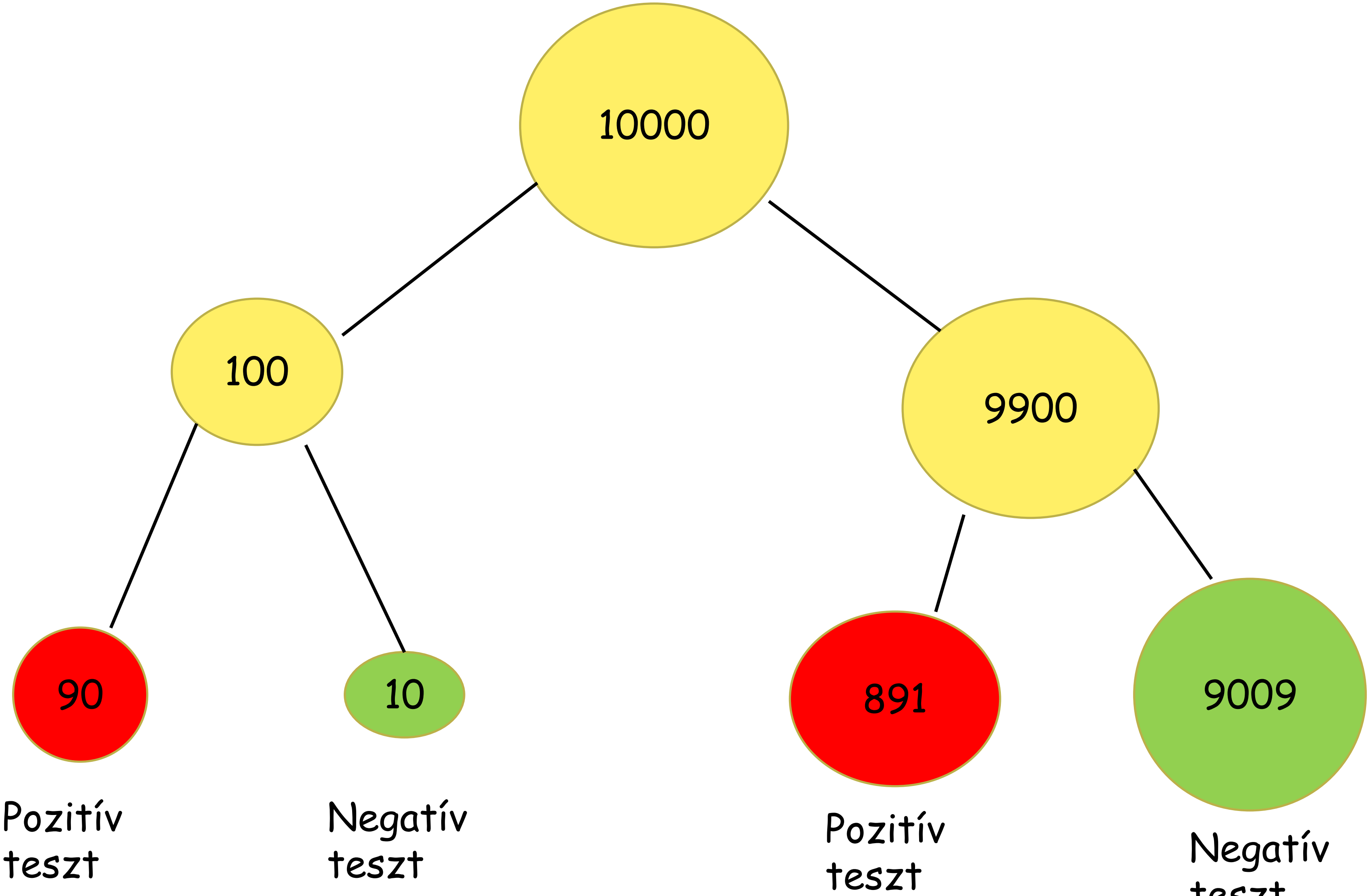
Megoldás

- Legyen A: emlődaganata van, B: pozitív a teszt. Ekkor

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\bar{A})P(\bar{A})} =$$

$$= \frac{0,9 \cdot 0,01}{0,9 \cdot 0,01 + 0,09 \cdot 0,99} = 0,09174312$$

Gyakoriságok



Teljes eseményrendszer

- Definíció: $A_1, A_2, \ldots$ teljes eseményrendszert alkotnak, ha

$$A_i A_j = \emptyset, i \neq j\text{-re és}$$

$$\bigcup_i A_i = \Omega$$

Teljes valószínűség tétele

- Tegyük fel, hogy $A_1, A_2, \ldots$ $(0 < P(A_i))$ teljes eseményrendszer, ekkor

$$P(B) = \sum_i P(B|A_i)P(A_i)$$

Bizonyítás:

$$B = \bigcup_i BA_i \Longrightarrow P(B) = \sum_i P(BA_i) = \sum_i \frac{P(BA_i)}{P(A_i)} P(A_i)$$

Szindbád és a háremhölgyek

- Szindbád a szultánnak tett szolgálataiért cserében 100 háremhölgy közül választhat. Azonban a háremhölgyeket nem egyszerre, hanem sorban egymás után mutatják be neki. Amennyiben egy bemutatott hölgyet nem választ ki azonnal, úgy már az örökké elveszik számára. Milyen stratégiát válasszon Szindbád, hogy a legszebb választásának minél nagyobb legyen a valószínűsége?

Stratégia

- Szindbád azt a stratégiát választja, hogy elengedi az első k hölgyet, és az utána következők közül az addigi legszebbet választja.
- Mennyi az optimális k?

Megoldás

- A_i: $i -$ edik hölgyet választja Szindbád $(i = k + 1, k + 2, \ldots, n = 100)$.
- A_0 egyik hölgyet sem választja ki Szindbád
- $A_0, A_{k+1}, A_{k+2}, \ldots, A_n$ teljes eseményrendszer
- B: a legszebbet választja ki

Megoldás (folyt.)

- $P(B) = \sum_i P(BA_i)$
- $P(BA_0) = 0$
- $P(BA_i) = \frac{1 \cdot 2 \cdot \ldots \cdot k \cdot k \cdot (k+1) \cdot \ldots \cdot (i-2) \cdot 1 \cdot i \cdot \ldots \cdot (n-1)}{1 \cdot 2 \cdot \ldots \cdot n} =$

$$= \frac{k}{(i-1)n}$$

- $P(B) = \sum_{i=k+1}^{n} \frac{k}{(i-1)n} = \frac{k}{n} \sum_{i=k}^{n-1} \frac{1}{i}$
- Milyen k-ra lesz ez maximális?

Megoldás (folyt.)

- k_n^*: az optimális érték
- $\frac{n}{k_n^*} \to e$

Bayes-formula általános alakja

- Tegyük fel, hogy $P(B) > 0, A_1, A_2, \dots$

$(0 < P(A_i))$ teljes eseményrendszer, ekkor

$$P(A_k|B) = \frac{P(B|A_k)P(A_k)}{\sum_i P(B|A_i)P(A_i)}$$

Bizonyítás:

$$P(A_k|B) = \frac{P(BA_k)}{P(B)} = \frac{\frac{P(BA_k)}{P(A_k)}P(A_k)}{\sum_i P(B|A_i)P(A_i)}$$

Monty Hall-paradoxon

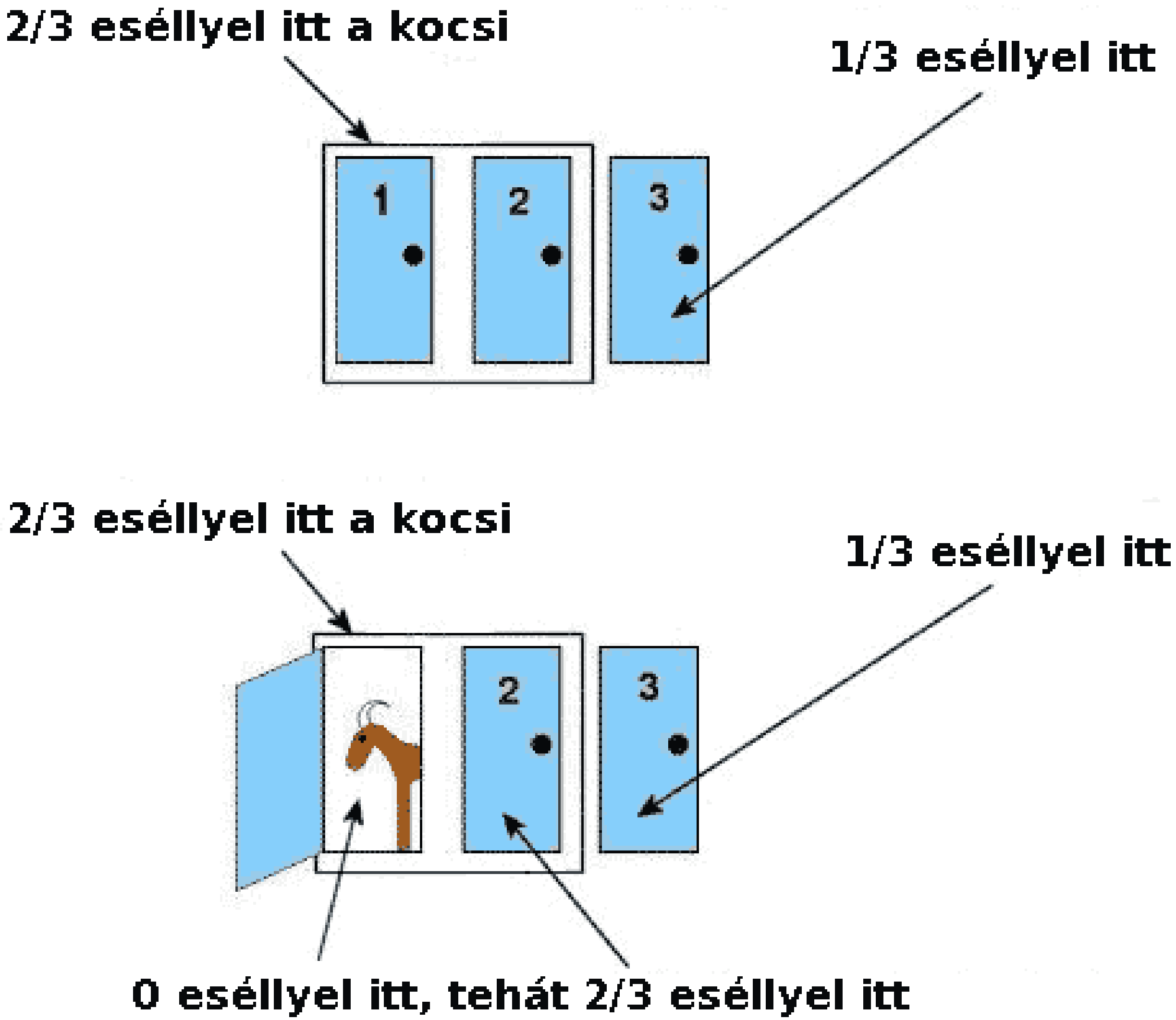
- Képzeljük el, hogy egy vetélkedőben szerepel, és három ajtó közül kell választania. Az egyik mögött kocsi van, a másik kettő mögött viszont kecske. Tegyük fel, hogy maga a 3. ajtót választja, mire a műsorvezető, aki tudja, melyik ajtó mögött mi van, kinyitja az 1. ajtót, megmutatván, hogy amögött kecske van. Ezután önhöz fordul, és megkérdezi: „Nem akarja esetleg mégis a 2. ajtót választani?” Vajon előnyére válik, ha vált?

Megoldás

- A_i: $i -$ edik ajtó mögött van a kocsi
- $P(A_i) = \frac{1}{3}, i = 1, 2, 3$
- B: műsorvezető az első ajtót nyitja ki
- Kérdés: $P(A_3|B) =?$
- $P(A_3|B) =$

$$\frac{P(B|A_3)P(A_3)}{P(B|A_1)P(A_1)+P(B|A_2)P(A_2)+P(B|A_3)P(A_3)} =$$

$$= \frac{\frac{1}{2} \cdot \frac{1}{3}}{0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3}} = \frac{1}{3}$$


2/3 eséllyel itt a kocsi
1/3 eséllyel itt
1
2
3
2/3 eséllyel itt a kocsi
1/3 eséllyel itt
2
3
0 eséllyel itt, tehát 2/3 eséllyel itt

Függetlenség

- Az egyik legfontosabb valószínűségszámítási fogalom.
- Definíció: A és B függetlenek, ha $P(AB) = P(A)P(B)$.
- Ha $P(B) > 0$, akkor

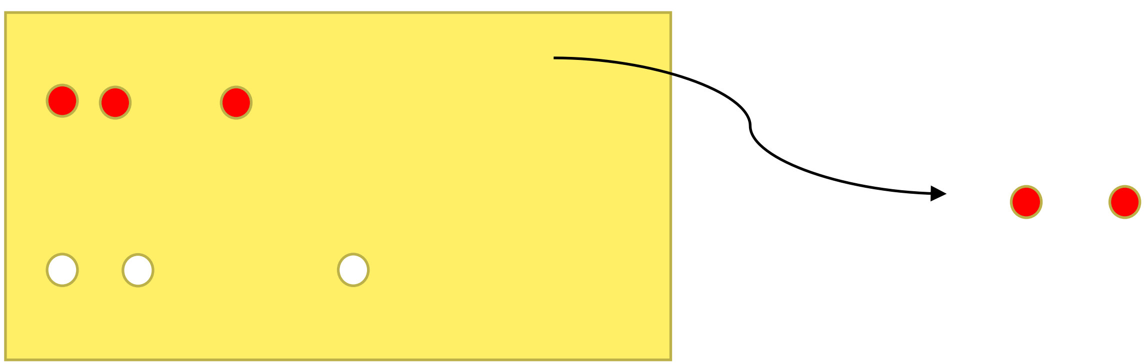
A és B függetlenek $\Leftrightarrow P(A|B) = P(A)$

Bizonyítás:

$$\frac{P(AB)}{P(B)} = P(A) \underset{P(B)>0}{\Longleftrightarrow} P(AB) = P(A)P(B)$$

Visszatevéses és visszatevés nélküli mintavétel

- Egy dobozban M piros és $N - M$ fehér golyó van. 2-szer húzunk.
- A: elsőre pirosat, B: másodikra pirosat húzunk. Függetlenek-e?



Visszatevéses és visszatevés nélküli mintavétel (folyt.)

- $P(A) = P(B) = \frac{M}{N}$.
- Visszatevésesnél:

$$P(AB) = \frac{M^2}{N^2} = P(A)P(B).$$

- Visszatevés nélkül:

$$P(AB) = \frac{M(M-1)}{N(N-1)} \neq P(A)P(B).$$

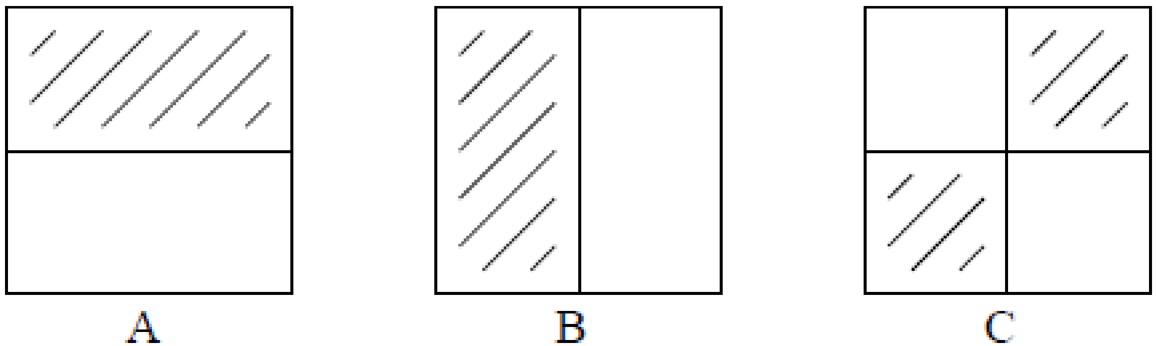
Több esemény függetlensége

- Definíció: $A_1, A_2, \dots, A_n$ függetlenek, ha bárhogy választunk ki közülük k darabot $(2 \leq k \leq n)$ úgy:

$$P(A_{i_1} A_{i_2} \dots A_{i_k}) = P(A_{i_1}) P(A_{i_2}) \dots P(A_{i_k})$$

- Nem elég sem a páronkénti függetlenség, sem az "n-es szorzat"!

Páronkénti, de nem teljes függetlenség



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

$$P(AB) = P(AC) = P(BC) = \frac{1}{4}$$

$$P(ABC) = 0 \neq P(A) \cdot P(B) \cdot P(C)$$

Jogos volt-e az ítélet?

- 1999-ben egy brit bíróság elítélte Sally Clarkot, mert 2 gyermeke is hirtelen csecsemőhalállal hunyt el.
- Az indok az volt, hogy a gyermekorvos szakértő szerint egy csecsemőnél 1/8500 az esélye egy ilyen halálesetnek, ezért a bíróság szerint a 2 eset valószínűsége ~ 1/73 millió.
- Későbbi kutatások kimutatták, hogy az első haláleset után a második esetnek már 1/100 az esélye (nem függetlenek).