

Valószínűségszámítás

2. előadás

Arató Miklós

2024.02.19.

Tartalomjegyzék

- 1 Feltételes valószínűség
- 2 Teljes valószínűség tétele
- 3 Függetlenség

Meghatározás

Definíció: 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:

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

Megjegyzés

Amennyiben $P(B) > 0$, akkor

$$(\Omega, \mathcal{A}, P(\cdot|B))$$

is Kolmogorov féle valószínűségi mező lesz.

\Rightarrow

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. Simpsonsot, 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.

Nicole és O.J.



Az ügyvédnek vagy ügyésznek volt igaza ?

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

Tétel: 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%

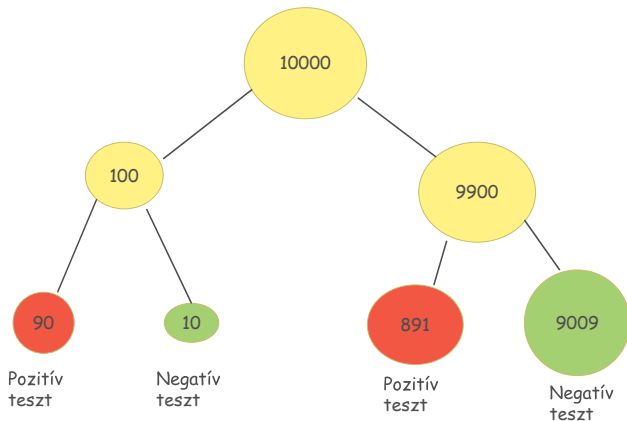
Megoldás

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

Gyakoriságok



Teljes eseményrendszer

Definíció: A_1, A_2, \dots események teljes eseményrendszert alkotnak, ha $A_i A_j = \emptyset, i \neq j$ és $\bigcup_i A_i = \Omega$.

Teljes valószínűség tétele

Állítás. Legyen A_1, \dots, A_i, \dots teljes eseményrendszer, valamint $P(A_i) > 0$ minden i -re. Ekkor $P(B) = \sum_i P(B|A_i) \cdot P(A_i)$.

Teljes valószínűség tétele

Állítás. Legyen A_1, \dots, A_i, \dots teljes eseményrendszer, valamint $P(A_i) > 0$ minden i -re. Ekkor $P(B) = \sum_i P(B|A_i) \cdot P(A_i)$.

Bizonyítás. Tudjuk, hogy $B = \bigcup_i BA_i$, ezért $P(B) = \sum_i P(BA_i)$.

Alkalmazzuk a $P(B|A_i) \cdot P(A_i) = P(BA_i)$ összefüggést.

Példa. Adott két fiók, az egyikben 1 db húszezer és 9 db ötszáz forintos, a másikban 10 db húszezer forintos. Találomra választva egy fiókot, mi a valószínűsége, hogy húszezer forintot húzunk?

Teljes valószínűség tétele

Állítás. Legyen A_1, \dots, A_i, \dots teljes eseményrendszer, valamint $P(A_i) > 0$ minden i -re. Ekkor $P(B) = \sum_i P(B|A_i) \cdot P(A_i)$.

Bizonyítás. Tudjuk, hogy $B = \bigcup_i BA_i$, ezért $P(B) = \sum_i P(BA_i)$.

Alkalmazzuk a $P(B|A_i) \cdot P(A_i) = P(BA_i)$ összefüggést.

Példa. Adott két fiók, az egyikben 1 db húszezer és 9 db ötszáz forintos, a másikban 10 db húszezer forintos. Találomra választva egy fiókot, mi a valószínűsége, hogy húszezer forintot húzunk? A teljes valószínűség tétele szerint $P(A) = P(A|B_1) \cdot P(B_1) + P(A|B_2) \cdot P(B_2) = \frac{1}{10} \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{11}{20}$, ahol B_i az i -edik fiók választását jelöli, A pedig azt, hogy húszezer forintot húzunk.

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?



Szindbád és a háremhölgyek (folyt.)

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

Szindbád és a háremhölgyek (folyt.)

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

Legyen A : a legszebbet választja, B_k : a k . hölgyet választja,
 B_0 : egyiket sem választja.

B_0, B_{m+1}, \dots, B_n teljes eseményrendszer, így

$$P(A) = 0 + \sum_{k=m+1}^n P(AB_k).$$

A hölgyek lehetséges sorrendjei: $|\Omega| = 1 \cdot 2 \cdot 3 \cdot \dots \cdot n = n!$

Mennyi AB_k elemszáma?

Az első m helyen tetszőleges lehet a sorrend. Az $m + 1$ -edik hölgytől kezdve a $k - 1$ -ig a hölgyek nem lehetnek az addigi legszebbek. A k . hölgy a legszebb, a $k + 1$ -edik hölgytől kezdve a hölgyek csúnyábbak a k -adiknál. \Rightarrow

Szindbád és a háremhölgyek (folyt.)

$$|AB_k| = 1 \cdot 2 \cdot \dots \cdot m \cdot m \cdot (m+1) \cdot \dots \cdot (k-2) \cdot 1 \cdot k \cdot \dots \cdot (n-1) \Rightarrow$$

$$P(AB_k) = \frac{|AB_k|}{n!} = \frac{m}{(k-1)n} \Rightarrow$$

$$P(A) = \sum_{k=m+1}^n P(AB_k) = \frac{m}{n} \cdot \sum_{k=m+1}^n \frac{1}{k-1} = \frac{m}{n} \sum_{k=m}^{n-1} \frac{1}{k}.$$

Szindbád és a háremhölgyek (folyt.)

$$|AB_k| = 1 \cdot 2 \cdot \dots \cdot m \cdot m \cdot (m+1) \cdot \dots \cdot (k-2) \cdot 1 \cdot k \cdot \dots \cdot (n-1) \Rightarrow$$

$$P(AB_k) = \frac{|AB_k|}{n!} = \frac{m}{(k-1)n} \Rightarrow$$

$$P(A) = \sum_{k=m+1}^n P(AB_k) = \frac{m}{n} \cdot \sum_{k=m+1}^n \frac{1}{k-1} = \frac{m}{n} \sum_{k=m}^{n-1} \frac{1}{k}. \text{ Mivel}$$

$$\sum_{k=m}^{n-1} \frac{1}{k+1} \leq \int_m^n \frac{1}{x} dx = -\ln \frac{m}{n} \leq \sum_{k=m}^{n-1} \frac{1}{k}, \text{ ezért}$$

$$-\frac{m}{n} \cdot \ln \frac{m}{n} \leq \frac{m}{n} \cdot \sum_{k=m}^{n-1} \frac{1}{k} \leq -\frac{m}{n} \cdot \ln \frac{m}{n} + \frac{1}{n}. \text{ Továbbá ha } m_n^* \text{-ra}$$

$$\text{teljesül, hogy } \frac{m_n^*}{n} \cdot \sum_{k=m_n^*}^{n-1} \frac{1}{k} = \max_m \frac{m}{n} \cdot \sum_{k=m}^{n-1} \frac{1}{k}, \text{ akkor belátható,}$$

$$\text{hogy } \frac{m_n^*}{n/e} \xrightarrow{n \rightarrow \infty} 1.$$

Bayes-formula általános alakja

Állítás. Legyen A_1, \dots, A_i, \dots teljes eseményrendszer, valamint $P(B) > 0, P(A_i) > 0$ minden i -re. Ekkor

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

Bayes-formula általános alakja

Állítás. Legyen A_1, \dots, A_i, \dots teljes eseményrendszer, valamint $P(B) > 0, P(A_i) > 0$ minden i -re. 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?



Monty Hall-paradoxon (megoldás)

A_i : i -edik ajtó mögött van a kocsi

$$P(A_i) = \frac{1}{3}, i = 1, 2, 3$$

B : a 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 esemény függetlensége

Definíció: A és B függetlenek, ha $P(AB) = P(A) \cdot P(B)$.

Megjegyzések

- a) Ha $P(A) = 0$, akkor A minden más eseménytől független.

2 esemény függetlensége

Definíció: A és B függetlenek, ha $P(AB) = P(A) \cdot P(B)$.

Megjegyzések

- Ha $P(A) = 0$, akkor A minden más eseménytől független.
- Ha $P(B) > 0$, akkor A és B pontosan akkor függetlenek, ha $P(A|B) = P(A)$.

2 esemény függetlensége

Definíció: A és B függetlenek, ha $P(AB) = P(A) \cdot P(B)$.

Megjegyzések

- Ha $P(A) = 0$, akkor A minden más eseménytől független.
- Ha $P(B) > 0$, akkor A és B pontosan akkor függetlenek, ha $P(A|B) = P(A)$.
- Az egymást kizáró események nem függetlenek, azaz $P(A) > 0$, $P(B) > 0$, $P(AB) = 0$ esetén A és B nem független.

Több esemény függetlensége

Definíció: Az A_1, A_2, \dots, A_n események **függetlenek**, ha $P(A_{i_1} \dots A_{i_k}) = P(A_{i_1}) \cdot \dots \cdot P(A_{i_k})$ minden $k \leq n$ -re és $1 \leq i_1 < \dots < i_k \leq n$ -re.

Több esemény függetlensége

Definíció: Az A_1, A_2, \dots, A_n események **függetlenek**, ha $P(A_{i_1} \dots A_{i_k}) = P(A_{i_1}) \cdot \dots \cdot P(A_{i_k})$ minden $k \leq n$ -re és $1 \leq i_1 < \dots < i_k \leq n$ -re.

Definíció: Az A_1, A_2, \dots események **függetlenek**, ha minden n -re A_1, \dots, A_n függetlenek.

Visszatevéses és visszatevés nélküli mintavétel

Egy dobozban M piros és $(N - M)$ fehér golyó van, ezek közül húzunk ki kettőt. Legyen A : az első piros, B : a második piros.

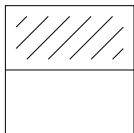
a) Visszatevéssel: $P(A) = \frac{M \cdot N}{N^2} = \frac{M}{N}$, $P(B) = \frac{N \cdot M}{N^2} = \frac{M}{N}$ és

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

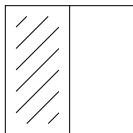
b) Visszatevés nélkül: $P(A) = \frac{M(N-1)}{N(N-1)} = P(B) = \frac{M}{N}$, de

$$P(AB) = \frac{M(M-1)}{N(N-1)} \neq \left(\frac{M}{N}\right)^2 = P(A) \cdot P(B).$$

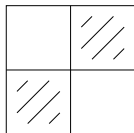
Páronkénti és "többszörös" függetlenség



A



B



C

Itt $P(A) = P(B) = P(C) = \frac{1}{2}$ és $P(AB) = P(AC) = P(BC) = \frac{1}{4}$,
azaz a három esemény páronként független, de
 $P(ABC) = 0 \neq P(A) \cdot P(B) \cdot P(C)$, így együttesen nem
függetlenek.