

Valószínűségszámítás előadás informatika BSC/A szakosoknak és matematika BSC-seknek

2019/2020 1. félév

Zempléni András

zempleni@caesar.elte.hu

<http://zempleni.elte.hu>

Megszámlálható valószínűségi mező

$\Omega = \{\omega_1, \omega_2, \dots, \omega_n, \dots\}$, $\mathcal{A} = \mathcal{P}(\Omega)$.

Jelölés: $p_i = P(\omega_i)$, valószínűségeloszlás:

$p_i \geq 0$, az összegük 1.

A σ -additivitás miatt tetszőleges A eseményre megy a véges esetre látott számítás:

$$P(A) = P(\cup_{i:\omega_i \in A} \omega_i) = \sum_{i:\omega_i \in A} p_i$$

Példa: Hányadikra dobjuk az első fejet egy szabályos érmével? $p_i = 1/2^i$ ($i=1,2,\dots$)

Geometriai valószínűségi mező

- $\Omega = V$ (az n -dimenziós Euklideszi tér véges térfogatú tartománya)
- \mathcal{A} a tengelyekkel párhuzamos oldalú téglákat:

$$T = \prod_{i=1}^n [a_i, b_i)$$

tartalmazó legszűkebb σ -algebra (elnevezés: Borel-halmazok, tartalmazza pl. az összes nyílt halmazt).

- $A \in \mathcal{A}$, $A \subseteq V$ esetén $P(A) = \lambda(A) / \lambda(V)$, ahol λ az n -dimenziós térfogat.

Példa: Buffon tű-kísérlete

- Egymástól egységnyi távolságra levő, párhuzamos vonalak közé ejtünk le ugyancsak egységnyi hosszúságú tűt. Mi a valószínűsége, hogy a tű metsz egy egyenest?

Tű-kísérlet

- Modell: A tű középpontja és az egyenesekkel bezárt szöge egy pont a $V = [0,1] \times [0,\pi]$ tartománynak. Feltesszük, hogy a valószínűség a geometriai módon számolható, azaz egy adott tartományba a tartomány területével arányos valószínűséggel esik.

A valószínűség folytonossága

Állítás. Ha $A_i \in \mathcal{A}$ ($i=1,2,\dots$) és $A_1 \supseteq A_2 \supseteq \dots$

akkor az $A = \bigcap_{i=1}^{\infty} A_i$ jelöléssel $\lim_{n \rightarrow \infty} P(A_n) = P(A)$

Bizonyítás. $A_1 = A \cup (A_1 \setminus A_2) \cup (A_2 \setminus A_3) \cup \dots$

diszjunkt felbontás, tehát a

$$P(A_1 \setminus A_2) + P(A_2 \setminus A_3) + \dots$$

sor konvergens. A fenti felbontást A_n -re alkalmazva:

$$P(A_n) = P(A) + P(A_n \setminus A_{n+1}) + P(A_{n+1} \setminus A_{n+2}) + \dots$$

Események uniójának valószínűsége

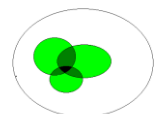
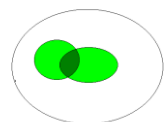
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Példa: Magyar kártyacsomagból kétszer húzunk visszatevéssel. Mi a valószínűsége, hogy húzunk pirosat?

A : első piros, B : második piros
 $P(A) = P(B) = 1/4$, $P(A \cap B) = 1/16$
 Tehát $P(A \cup B) = 7/16$

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) -$$

$$- P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$



Szita (Poincaré) formula

Képlet az általános esetre:

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = \sum_{i=1}^n (-1)^{i+1} S_i^{(n)}$$

ahol

$$S_i^{(n)} = \sum_{1 \leq j_1 < j_2 < \dots < j_i \leq n} P(A_{j_1} \cap A_{j_2} \cap \dots \cap A_{j_i})$$

az i tényezős metszetek valószínűségeinek összege.

Alkalmazások

- Ha az egyes események és metszeteik is egyformán valószínűek, akkor

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = \sum_{i=1}^n (-1)^{i+1} \binom{n}{i} P(A_1 \cap A_2 \cap \dots \cap A_i)$$

- Átfogalmazás metszetekre:
 $P(\bar{A}_1 \cap \bar{A}_2 \cap \dots \cap \bar{A}_n) =$

$$= 1 - P(A_1 \cup A_2 \cup \dots \cup A_n) = \sum_{i=0}^n (-1)^i S_i^{(n)}$$

(Megállapodás: $S_0 = 1$.)

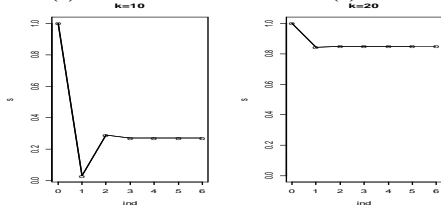
- Példa: Mi a valószínűsége, hogy adott (k) számú kockadobásból minden számot legalább egyszer megkaptunk?

Megoldás

- A_i : az i számot nem dobtuk

$$P(\bar{A}_1 \cap \bar{A}_2 \cap \dots \cap \bar{A}_6) =$$

$$= \sum_{i=0}^6 (-1)^i \binom{6}{i} P(A_1 \cap A_2 \cap \dots \cap A_i) = \sum_{i=0}^6 (-1)^i \binom{6}{i} \left(\frac{6-i}{6}\right)^k$$



Bonferroni egyenlőtlenség

- A megközelítés mindig ilyen (a váltakozó előjelű tagok összege közrefogja a pontos értéket), azaz tetszőleges l -re:

$$\sum_{i=0}^{2l-1} (-1)^i S_i^{(n)} \leq P(\bar{A}_1 \cap \bar{A}_2 \cap \dots \cap \bar{A}_n) \leq \sum_{i=0}^{2l} (-1)^i S_i^{(n)}$$

Jordán formula

- Annak a B_k eseménynek a valószínűségét keressük, hogy pontosan k db esemény következik be.

$$p_{n,k} = P(B_k) = \sum_{i=0}^{n-k} (-1)^i \binom{k+i}{i} S_{k+i}$$

- Spec.: $k=0$ -ra éppen a komplementerek metszetére vonatkozó képletet kapjuk vissza.
- Érdekes kérdés: hányadik kísérletnél lesz meg az adott számú esemény. A valószínűség kiszámítása:

$$p_{n,k} - p_{n-1,k}$$

- Szimuláció (szelvénygyűjtés)

Feltételes valószínűség 1.

- Az A esemény valószínűségét keressük.
- Tudjuk, hogy B esemény bekövetkezett.
- A relatív gyakoriságokkal: csak azokat a kísérleteket nézzük, amelyekben B bekövetkezett. Ezen részsorozatban az A relatív gyakorisága:

$$r_{A|B} / r_B$$

Feltételes valószínűség 2.

- Megfelelője a valószínűségekre:

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

az A esemény B-re vonatkozó feltételes valószínűsége (feltétel: $P(B) > 0$).

- Példa: kockadobás. $A = \{\text{páros számot dobunk}\}$
 $B = \{\text{3-nál nagyobbat dobunk}\}$
 $P(A|B) = 2/3$.
- Monty Hall játék:** melyik ajtót válasszuk? (Kettő mögött kecske, a harmadik mögött autó van.)
- Ugyanez szimulációval: [Monty Hall szimuláció](#)

Teljes eseményrendszer

- Definíció.** Események A_1, A_2, \dots , sorozata *teljes eseményrendszer*, ha egymást páronként kizárják és egyesítésük Ω .
- Tulajdonság: $P(A_1) + P(A_2) + \dots = 1$
- Legtöbbször véges sok elemből álló teljes eseményrendszereket vizsgálunk.

Teljes valószínűség tétele.

- Legyen B_1, B_2, \dots , pozitív valószínűségű eseményekből álló teljes eseményrendszer, $A \in \mathcal{A}$ tetszőleges. Ekkor

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \dots$$

- Bizonyítás.** $A = (A \cap B_1) \cup (A \cap B_2) \cup \dots$ diszjunkt tagokra bontás, tehát

$$P(A) = P(A \cap B_1) + P(A \cap B_2) + \dots$$

és $P(A \cap B_i) = P(A|B_i)P(B_i)$ adja a tételt.

Példák

- Összetett modellek (pl. nemtől függő valószínűségek): a színvakság valószínűsége a férfiaknál 0.05, a nőknél 0.0025 (Tegyük fel, hogy ugyanannyi a férfi, mint a nő.) Mi a valószínűsége, hogy egy találmásra választott ember színvak?
- A teljes eseményrendszer: {férfi} {nő}.
 $p = 0.05/2 + 0.0025/2 = 0.02625$

Bayes tétele

Legyen B_1, B_2, \dots , pozitív valószínűségű eseményekből álló teljes eseményrendszer, $A \in \mathcal{A}$ pozitív valószínűségű. Ekkor

$$P(B_k|A) = \frac{P(A|B_k)P(B_k)}{\sum P(A|B_i)P(B_i)}$$

(Visszakövetkeztetés az első lépés eredményére.)

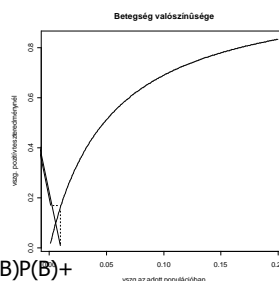
Bizonyítás. A nevező éppen $P(A)$ a teljes valószínűség tétele miatt.

A számláló pedig $P(A \cap B_k)$, definíció szerint.

Példa

Ha egy, az egészségesekre 5% eséllyel téves diagnózist adó szűrővizsgálatnál betegnek tűnünk, akkor a betegség tényleges valószínűsége (p a betegség vszge, $\{B = \text{beteg}, E = \text{egészséges}\}$ a teljes eseményrendszer):

$$P(B|\text{poz}) = P(\text{poz}|B)P(B) / (P(\text{poz}|B)P(B) + P(\text{poz}|E)P(E)) = p / (p + 0.05(1-p))$$



Események függetlensége

- Ha a B esemény bekövetkezése nem befolyásolja az A valószínűségét, azaz $P(A|B)=P(A)$, akkor azt mondjuk, hogy az A és B függetlenek. Ez így nem ideális definíció (nem szimmetrikus, $P(B)>0$ kell hozzá), ezért
- Definíció. Az A és B események függetlenek, ha $P(A \cap B) = P(A)P(B)$.

Példák

- Húzzunk egy lapot egy magyarkártya-csomagból. A : piros B : ász.
 $P(A)=1/4$, $P(B)=1/8$, $P(A \cap B)=1/32$, tehát függetlenek.
- A függetlenség nagyon ritka azonos kísérletből meghatározott eseményeknél!
- Tipikus eset függetlenségre: A az első, B a második kísérlet eredménye.

Tulajdonságok

- Ha A és B diszjunktak, akkor csak triviális ($P(A)=0$ vagy $P(B)=0$) esetben függetlenek.
- Ha A és B függetlenek, akkor komplementereik is függetlenek.
- Önmaguktól csak a triviális események függetlenek.
- $A \subset B$ esetén csak akkor függetlenek, ha legalább az egyik triviális.

Független kísérletek

- $A \in (\Omega_1, \mathcal{A}_1, P_1)$ az egyik kísérlethez kapcsolódik,
- $B \in (\Omega_2, \mathcal{A}_2, P_2)$ pedig a másik kísérlethez kapcsolódik.
- Függetlenségük értelmezéséhez kell a valószínűségi mezők szorzata:
 $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \times \mathcal{A}_2, P_1 \times P_2)$ ami szintén valószínűségi mező.
 $\mathcal{A}_1 \times \mathcal{A}_2$ elemei az $A \times B$ alakú események ($A \in \mathcal{A}_1$ megfelelője az $A \times \Omega_2$, így már értelmezhető $A \cap B = A \times B$).
- $P_1 \times P_2 (A \times B) = P_1(A)P_2(B)$, ami éppen a függetlenséget jelenti.
- Véges esetben $|\Omega_1|=n$, $|\Omega_2|=m$ mellett $|\Omega_1 \times \Omega_2|=nm$, amit már sokszor használtunk is.

Általánosítás

- Két eseményrendszer független, ha az első tetszőleges eleme független a második tetszőleges elemétől.
- n esemény független, ha
 $P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = P(A_{i_1})P(A_{i_2}) \dots P(A_{i_k})$
teljesül tetszőleges $1 \leq i_1 < i_2 < \dots < i_k \leq n$ indexsorozatra és minden $2 \leq k \leq n$ számra.

Megjegyzések

- Nem elég a fenti szorzat-tulajdonságot $k=2$ -re megkövetelni. Ha csak ez teljesül: páronkénti függetlenségről beszélünk.
- A szorzat-előállítás segítségével n független kísérlethez tartozó valószínűségi mező is értelmezhető. Ha A_i az i -edik kísérlethez tartozik, akkor A_1, A_2, \dots, A_n független. (A gyakorlatban ez a tipikus, fontos előfordulása ennek a függetlenségnek.)