Valószínűségszámítás előadás informatika BSC/A szakosoknak és matematikai elemző BSC-seknek

2017/2018 1. félév

2. előadás

Visszatevéses mintavétel

- N termék, melyből M selejtes
- n elemű minta visszatevéssel
- A: pontosan k selejtes van a mintában

(k=0,…,n)
$$P(A)=\binom{n}{k}\left(\frac{M}{N}\right)^k\left(1-\frac{M}{N}\right)^{n-k}$$
azaz a valószínűség kifejezhető a $p=M/N$ selejtarány segítségével:
$$P(A)=\binom{n}{k}p^k(1-p)^{n-k}$$
Mintavétel

Visszatevés nélküli mintavétel

- N termék, melyből M selejtes
- n elemű minta visszatevés nélkül
- A: pontosan k selejtes van a mintában

(k=0,…,n)
$$P(A)=\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$$
Mintavétel

További lehetőségek

- Elektronok energia-szintek közötti megoszlásai, Bose-Einstein statisztika:
- A lehetőségek száma (n részecske, k energia-szint)

$$w(n,k)=\binom{n+k-1}{n}$$

A valószínűség további tulajdonságai

A valószínűség végesen is additív:
ha A_1, A_2 , ..., A_n páronként kizáró események, akkor
$$P(A_1\cup A_2\cup\ldots\cup A_n)=P(A_1)+P(A_2)+\ldots+P(A_n)$$
Bizonyítás. $A_{n+1}=A_{n+2}=\ldots=\emptyset$ választással alkalmazzuk a σ-additivitást.
Tehát a korábban belátott tulajdonságok a Kolmogorov-féle valószínűségi mezőre is érvényesek.

Megszámlálható valószínűségi mező

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

Geometriai valószínűségi mező

- Ω=V (az n-dimenziós Euklideszi tér véges térfogatú tartománya)
- $\mathcal{A}$ a tengelyekkel párhuzamos oldalú téglákat:
  $$T = \mathop{\times}_{i=1}^{n} [a_i, b_i)$$
  tartalmazó legszűkebb σ-algebra (elnevezés: Borel-halmazok, tartalmazza pl. az összes nyílt halmazt).
- A∈ $\mathcal{A}$, A⊆ V esetén P(A):=λ(A)/ λ(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]X[0,π] tartománynak. Feltesszük, hogy a valószínűség a geometriai módon számolható, azaz egy adott tartományba a tartomány területével arányos valószínűséggel esik.

A valószínűség folytonossága

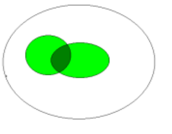
Állítás. *Ha $A_i \in \mathcal{A}$ (i=1,2,...) és $A_1 \supseteq A_2 \supseteq$...*

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

Bizonyítás. $A_1 = A \cup (A_1 \setminus A_2) \cup (A_2 \setminus A_3) \cup \ldots$
diszjunkt felbontás, tehát a
$$P(A_1 \setminus A_2) + P(A_2 \setminus A_3) + \ldots$$
sor konvergens. A fenti felbontást A_n-re alkalmazva:
$$P(A_n) = P(A) + P(A_n \setminus A_{n+1}) + P(A_{n+1} \setminus A_{n+2}) + \ldots$$

Események uniójának valószínűsége

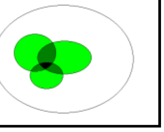
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$



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:
$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = \sum_{i=1}^{n} (-1)^{i+1} S_i^{(n)}$$
ahol
$$S_i^{(n)} = \sum_{1 \le j_1 < j_2 < \ldots < j_i \le n} P(A_{j_1} \cap A_{j_2} \cap \ldots \cap A_{j_i})$$

az *i* tényezős metszetek valószínűségeinek összege.

Alkalmazások

- Ha az egyes események és metszeteik is egyformán valószínűek, akkor
  $$P(A_1 \cup A_2 \cup \ldots \cup A_n) = \sum_{i=1}^{n} (-1)^{i+1} \binom{n}{i} P(A_1 \cap A_2 \cap \ldots \cap A_i)$$
- Átfogalmazás metszetekre:
  $$P(\bar{A}_1 \cap \bar{A}_2 \cap \ldots \cap \bar{A}_n) =$$
  $$= 1 - P(A_1 \cup A_2 \cup \ldots \cup A_n) = \sum_{i=0}^{n} (-1)^i S_i^{(n)}$$

(Megállapodás: S_0=1.)

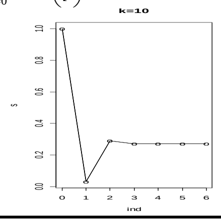
- Példa: Mi a valószínűsége, hogy adott (*k*) számú kockadobásból minden számot legalább egyszer megkaptunk?

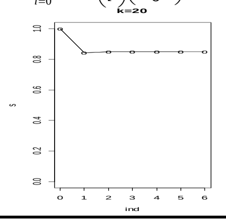
Megoldás

- A_i: az i számot nem dobtuk

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

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





Bonferroni egyenlőtlenség

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

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

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:

$$r_{A \cap B} / r_B$$

Feltételes valószínűség 2.

- Megfelelője a valószínűségekre:

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

az A esemény B-re vonatkozó feltételes valószínűsége (feltétel: $P(B)$>0).
- Példa: kockadobás. A={páros számot dobunk}
  B={3-nál nagyobbat dobtunk}
  $P(A|B)$=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 , ..., sorozata *teljes eseményrendszer*, ha egymást páronként kizárják és egyesítésük Ω.
- Tulajdonság: $P(A_1) + P(A_2) + \ldots = 1$
- Legtöbbször véges sok elemből álló teljes eseményrendszereket vizsgálunk.

Teljes valószínűség tétele.

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

$$P(A) = P(A \mid B_1)P(B_1) + P(A \mid B_2)P(B_2) + \ldots$$

- *Bizonyítás.* $A = (A \cap B_1) \cup (A \cap B_2) \cup \ldots$

diszjunkt tagokra bontás, tehát

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

és $P(A \cap B_i) = P(A \mid B_i)P(B_i)$ 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 találomra válaszott ember színvak?
- A teljes eseményrendszer: {férfi} {nő}.

p=0.05/2+0.0025/2=0.02625

Bayes tétele

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

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

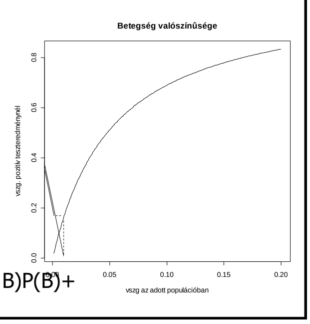
(Visszakövetkeztetés az első lépés eredményére.)

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=beteg, E=egészséges} a teljes eseményrendszer):



P(B|poz)=P(poz|B)P(B)/(P(poz|B)P(B)+ P(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

- Húzunk 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.