A sztochasztika alapjai
MBNXK262

2. előadás: Szitaformula, geometriai valószínűség, feltételes valószínűség

Kevei Péter

2023/24 tavasz

Ismétlés

A véletlen kísérlet lehetséges kimeneteleinek halmaza az **eseménytér**, jele Ω.
Az **esemény** olyan a kísérlettel kapcsolatban tett állítás, melynek igaz vagy hamis volta eldönthető a kísérlet lefolytatása után. Az **események halmaza** $\mathcal{A}$, az Ω részhalmazainak egy rendszere. Ha Ω véges, akkor $\mathcal{A} = 2^{\Omega}$.
Egy $\mathbf{P} : \mathcal{A} \to [0, 1]$ halmazfüggvény *valószínűségi mérték* az $(\Omega, \mathcal{A})$ mérhetőségi téren, ha

- $\mathbf{P}(\Omega) = 1$;
- ha az $A_1, A_2, \ldots \in \mathcal{A}$ halmazok (páronként) diszjunktak, akkor
$$\mathbf{P}(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbf{P}(A_i),$$
azaz a halmazfüggvény σ-additív.

A fenti tulajdonságokkal rendelkező $(\Omega, \mathcal{A}, \mathbf{P})$ hármast **valószínűségi mezőnek** nevezzük.

Ismétlés

- Legyenek $A_1, A_2, \ldots, A_n \in \mathcal{A}$. Ha $A_i \cap A_j = \emptyset$, minden $i \neq j$ párra, akkor

$$\mathbf{P}(A_1 \cup \ldots \cup A_n) = \mathbf{P}(A_1) + \ldots + \mathbf{P}(A_n).$$

- $\mathbf{P}(A^c) = 1 - \mathbf{P}(A)$.
- $A \subset B \Rightarrow \mathbf{P}(B \setminus A) = \mathbf{P}(B) - \mathbf{P}(A)$, és $\mathbf{P}(A) \leq \mathbf{P}(B)$.
- $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$.

Ismétlés

Szitaformula:

$$\mathbf{P}(A_1 \cup \ldots \cup A_n) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{1 \le i_1 < i_2 < \ldots < i_k \le n} \mathbf{P}(A_{i_1} \cap \ldots \cap A_{i_k}).$$

$n = 2$-re:

$$\mathbf{P}(A_1 \cup A_2) = \mathbf{P}(A_1) + \mathbf{P}(A_2) - \mathbf{P}(A_1 \cap A_2),$$

és $n = 3$-ra:

$$\begin{aligned}\mathbf{P}(A_1 \cup A_2 \cup A_3) &= \mathbf{P}(A_1) + \mathbf{P}(A_2) + \mathbf{P}(A_3) \\ &- (\mathbf{P}(A_1 \cap A_2) + \mathbf{P}(A_1 \cap A_3) + \mathbf{P}(A_2 \cap A_3)) \\ &+ \mathbf{P}(A_1 \cap A_2 \cap A_3).\end{aligned}$$

A valószínűség tulajdonságai

Állítás

$\mathbf{P}(A \cup B) \leq \mathbf{P}(A) + \mathbf{P}(B)$.

A valószínűség tulajdonságai

Állítás

$\mathbf{P}(A_1 \cup \ldots \cup A_n) \leq \mathbf{P}(A_1) + \ldots + \mathbf{P}(A_n).$

A párosítási probléma

Feladat

Veszünk n darab kártyát 1-től n-ig megszámozva. Összekeverjük, és véletlen sorrendben lerakjuk őket egy sorba. A k-adik helyen párosítás történik, ha a k-adik helyre a k sorszámú kártya kerül. (Tehát véletlen permutációk fixpontjait tekintjük.)

Mennyi a valószínűsége, hogy nem történik párosítás? Jelölje p_n ezt a valószínűséget.

A párosítási probléma: a valószínűségi mező

- Eseménytér:

  $$\Omega = \{\text{az } \{1, \ldots, n\} \text{ halmaz permutációi}\}$$
  $$= \{\omega = (\omega_1, \ldots, \omega_n) \, : \, \omega_1, \ldots, \omega_n \text{ az } 1, 2, \ldots, n \text{ egy felsorolása}\},$$

  azaz egy $\omega \in \Omega$ esetén $1 \mapsto \omega_1$, $2 \mapsto \omega_2$, …, $n \mapsto \omega_n$.
  $|\Omega| = n \times (n-1) \times \ldots \times 1 = n!$.
- Események: $\mathcal{A} = 2^\Omega$ a hatványhalmaz.
- Valószínűségi mérték: $\mathbf{P} : \mathcal{A} \to [0, 1]$:

$$\mathbf{P}(A) = \frac{|A|}{|\Omega|} = \frac{\text{kedvező esetek száma}}{\text{összes esetek száma}}.$$

A párosítási probléma

Jelölje A_k azt az eseményt, hogy a k-adik helyen párosítás történik, $k = 1, 2, \ldots, n$, és

$$A := \{\text{van párosítás}\}.$$

A párosítási probléma

A párosítási probléma

A párosítási probléma

A párosítási probléma

Mennyi a valószínűsége, hogy pontosan k darab párosítás történik? Legyen

$$p_{n,k} = \mathbf{P}(n \text{ kártya van, és pontosan } k \text{ párosítás történik}).$$

A párosítási probléma

Mennyi a valószínűsége, hogy pontosan k darab párosítás történik? Legyen

$$p_{n,k} = \mathbf{P}(n \text{ kártya van, és pontosan } k \text{ párosítás történik}).$$

Nyilván $p_n = p_{n,0}$. Jelölje $N_{n,k}$ azon kimenetelek számát, amikor pontosan k párosítás történik n kártyával. Ezekkel a jelölésekkel $p_m = N_{m,0}/m!$.

$$p_{n,k} = \frac{N_{n,k}}{n!} = \frac{\binom{n}{k} N_{n-k,0}}{n!} = \frac{\binom{n}{k}(n-k)!p_{n-k}}{n!}$$
$$= \frac{p_{n-k}}{k!} = \frac{1}{k!}\sum_{j=0}^{n-k}\frac{(-1)^j}{j!}.$$

A párosítási probléma

Az utóbbi alakból rögtön látjuk, hogy

$$\lim_{n\to\infty} p_{n,k} = \frac{1}{k!}\sum_{j=0}^{\infty}\frac{(-1)^k}{k!} = \frac{e^{-1}}{k!}.$$

Geometriai valószínűségi mező

Definíció

Geometriai valószínűségi mező esetén

- a lehetséges kimenetelek halmaza $\Omega = H \subset \mathbb{R}^n$ szép halmaz, aminek a mértéke (hossza, területe, térfogata) pozitív és véges.
- az események $\mathcal{A}$ halmaza az Ω szép részhalmazai.
- egy $A \subset H$ esemény valószínűsége arányos a halmaz mértékével, azaz
$$\mathbf{P}(A) = \frac{\lambda(A)}{\lambda(H)},$$
ahol λ az n-dimenziós térfogat (hossz, terület, térfogat).

Példák: $\Omega = [0,1] \times [0,1]$, $\Omega = [0,1] \times [0,2]$,
$\Omega = B(0,1) = \{(x,y) : x^2 + y^2 \leq 1\}$.

Példa

Egy négyzet belsejében egyenletes eloszlás szerint választunk egy pontot. Mennyi a valószínűsége, hogy a választott pont közelebb van valamelyik oldalhoz, mint 1/4?

Feltételes valószínűség

Definíció

$(\Omega, \mathcal{A}, \mathbf{P})$ egy valószínűségi mező, A, B események, és $\mathbf{P}(B) > 0$. Ekkor az A esemény B eseményre vonatkozó feltételes valószínűsége

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

Ha annyi információnk van a véletlen kísérletről, hogy a B esemény bekövetkezett, akkor az A esemény valószínűsége $\mathbf{P}(A|B)$.

Állítás

Rögzítsünk egy tetszőleges B eseményt, melyre $\mathbf{P}(B) > 0$. Ekkor $\mathbf{P}_B(A) = \mathbf{P}(A|B)$ valószínűségi mérték $\mathcal{A}$-n.

Állítás

Rögzítsünk egy tetszőleges B *eseményt, melyre* $\mathbf{P}(B) > 0$. *Ekkor* $\mathbf{P}_B(A) = \mathbf{P}(A|B)$ *valószínűségi mérték* $\mathcal{A}$*-n.*

Állítás (Szorzási szabály)

Legyenek $A_1, A_2, \ldots, A_n$ *tetszőleges olyan események, melyekre* $\mathbf{P}(A_1 \cap \ldots \cap A_{n-1}) > 0$*. Ekkor*

$$\begin{aligned}\mathbf{P}(A_1 \cap \ldots \cap A_n) = &\, \mathbf{P}(A_1)\mathbf{P}(A_2|A_1)\mathbf{P}(A_3|A_1 \cap A_2)\ldots \\ & \times \mathbf{P}(A_n|A_1 \cap \ldots \cap A_{n-1}).\end{aligned}$$

$n = 2$-re:

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

Példa

Egy dobozban 12 kék és 3 fehér golyó van. Visszatevés nélkül húzunk két golyót egymás után. Mi annak a valószínűsége, hogy mindkét golyó kék?
Jelölje K_i azt, hogy az i-edik kék. Ekkor

$$\mathbf{P}(K_1 \cap K_2) =$$

Teljes eseményrendszer

Definíció

A $B_1, B_2, \ldots$ események teljes eseményrendszert alkotnak, ha

- minden $i \neq j$ párra $B_i \cap B_j = \emptyset$;
- $\cup_{n=1}^{\infty} B_n = \Omega$.

Tétel (Teljes valószínűség tétele)

Legyen $B_1, B_2, \ldots$ teljes eseményrendszer, melyre $\mathbf{P}(B_n) > 0$ minden n-re. Ekkor tetszőleges $A \in \mathcal{A}$ esemény esetén

$$\mathbf{P}(A) = \sum_{n=1}^{\infty} \mathbf{P}(A|B_n)\mathbf{P}(B_n).$$

Szindbád

Szindbádnak jogában áll N háremhölgy közül egyet kiválasztania oly módon, hogy az előtte egyenként elvonuló hölgyek valamelyikére rámutat. Tegyük fel, hogy egyértelmű szigorúan monoton szépségi sorrendet tud felállítani, és a háremhölgyek bármely elvonulási sorrendje egyformán valószínű. Szindbád k hölgyet elenged, majd kiválasztja az elsőt, aki szebb az összes előtte elvonultnál. Mennyi a valószínűsége, hogy a legszebb hölgyet választja ki? Milyen k esetén lesz ez a valószínűség a legnagyobb, ha N elég nagy?

Szindbád

Jelölje A_i azt az eseményt, hogy a i-edik lány a legszebb, $i = 1, 2, \ldots, N$, és B azt, hogy Szindbád a legszebb lányt választja.

Szindbád

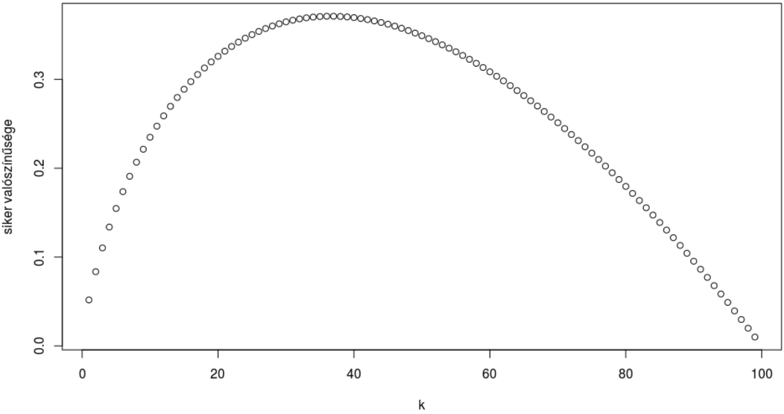
Ha $i > k$, akkor Szindbád pontosan akkor választja a legszebbet, ha az első $i - 1$ lány közül a legszebb az első k-ban volt.
Tehát

$$\mathbf{P}(B|A_i) = \frac{k}{i-1}.$$

Összegezve

$$\mathbf{P}(B) = \sum_{i=1}^{N} \mathbf{P}(A_i)\,\mathbf{P}(B|A_i) = \sum_{i=k+1}^{N} \frac{1}{N}\frac{k}{i-1}.$$

Szindbád



ábra: A siker valószínűsége k függvényében, $N = 100$

Szindbád

És ez mennyi?

$$\mathbf{P}(B) = \frac{k}{N} \sum_{i=k}^{N-1} \frac{1}{i}.$$

Integrálközelítő összeg becslése integrállal:

$$\log \frac{N}{k} = \int_k^N \frac{1}{x} \mathrm{d}x \leq \sum_{i=k}^{N-1} \frac{1}{i} \leq \int_{k-1}^{N-1} \frac{1}{x} \mathrm{d}x = \log \frac{N-1}{k-1}.$$

Tehát

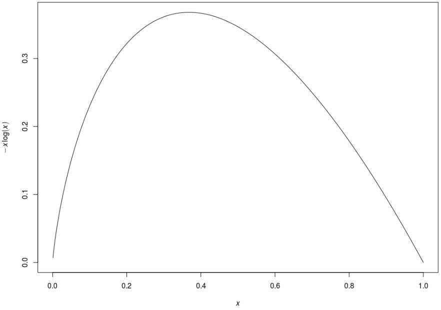
$$\mathbf{P}(B) \sim \frac{k}{N} \log \frac{N}{k}.$$

Szindbád

Hogy válassza Szindbád k-t?

$$\mathbf{P}(B) \sim \frac{k}{N} \log \frac{N}{k}.$$

Az $f(x) = x \log \frac{1}{x}$ függvény maximuma kell $[0, 1]$-en. HF
Maximumhely $x = 1/e$,
maximum értéke: $1/e$,
tehát $k = [N/e]$.



Szindbád

- Szindbád megtudja N-et.
- Elenged $k = [N/e]$ lányt.
- Majd kiválasztja az elsőt, aki szebb az összes eddiginél.
- Ekkor nagyjából $1/e \approx 0,368$ valószínűséggel a legszebbbbet választja.

Bayes-formula

Tétel

Legyenek A és B olyan események, hogy $\mathbf{P}(A) > 0$, $\mathbf{P}(B) > 0$. Ekkor

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

Bayes-tétel

Tétel

Legyen $B_1, B_2, \ldots$ *teljes eseményrendszer, melyre* $\mathbf{P}(B_n) > 0$ *minden* n*-re. Ekkor tetszőleges pozitív valószínűségű* $A \in \mathcal{A}$ *esemény esetén, tetszőleges* k*-ra*

$$\mathbf{P}(B_k|A) = \frac{\mathbf{P}(A|B_k)\mathbf{P}(B_k)}{\sum_{n=1}^{\infty} \mathbf{P}(A|B_n)\mathbf{P}(B_n)}.$$

Monty Hall probléma

Egy játékos 3 csukott ajtó közül 1-et választhat. Kettő mögött kecske van, egy mögött autó, de nem tudja, hogy melyik hol van. A játékos autót szeretne nyerni. Miután egy ajtót kiválaszt, a játékvezető kinyittat egy ajtót, amit nem választott a játékos, és ami mögött kecske van. Ezután felajánlja a játékosnak, hogy cserélhet ajtót. Megéri-e váltani?

Monty Hall probléma

Tegyük fel, hogy a játékos az 1-es ajtót választotta, a játékvezető a 2-est nyittatta ki. Jelölje A_i azt az eseményt, hogy az autó az i-edik ajtó mögött van, B pedig azt, hogy a játékvezető a 2-es ajtót nyittatta ki.