Kurs: PSK Informatyka 3
st. stacj. / niestacj.
tryb: on-line,
Wykład 2

Temat. Model Probabilistyczny Kołmogorowa (MPK) jako opis teorii ZL.

Z przykładów nad pojedynczymi rzutami monetą widzimy, że na opis ZL składają się dwa zbiory oraz miara.
Uogólnimy teraz te rozważania formułując dokładny opis ZL.

Pojęcie MPK

Przez MPK rozumiemy standardową postać
(Ω, Σ, P), gdzie:

(i) Ω jest niepustym zbiorem nazywanym przestrzenią probabilistyczną.
Elementy Ω oznaczamy przez ω, czyli $\omega \in \Omega$.
Wtedy ω nazywamy zdarzeniem elementarnym
albo punktem próbkowym)

1

(ii) Σ jest zbiorem złożonym z pewnych podzbiorów Ω. Zatem Σ jest <u>rodziną zbiorów</u>) o własności

(∘) $\emptyset \in \Sigma$

(∘∘) $A \in \Sigma \Rightarrow A^c \in \Sigma$

(∘∘∘) jeśli $\{A_n, n \in \mathbb{N}_0 \subset \mathbb{N}\} \subset \Sigma$, to $\bigcup_{n \in \mathbb{N}_0} A_n \in \Sigma$ (mówimy, że Σ jest <u>zamknięta</u> na branie <u>przeliczalnych sum</u>)

<u>Uwaga.</u>

1°. Najmniejszą rodziną tego typu Σ jest $\{\emptyset, \Omega\}$.

2°. Największą jest rodzina złożona ze wszystkich podzbiorów Ω nazywana <u>rodziną potęgową</u> i oznaczana $\mathcal{P}(\Omega)$ albo 2^{Ω}.

W szczególności jeśli $|\Omega| = n$, to $|\mathcal{P}(\Omega)| = 2^{|\Omega|}$

Elementy rodziny Σ nazywamy zdarzeniami, a Σ σ-ciałem.

Przykład 1.

Niech $\Omega = \mathbb{R}$, oraz weźmy rodzinę

$$\mathcal{A} = \{ [a,b),\ a<b,\ a,b \in \Omega \}.$$

$\mathcal{A}$ nie jest σ-ciałem! (dlaczego?)

Istnieje najmniejsze σ-ciało zawierające $\mathcal{A}$.

Oznaczmy je $\mathcal{B}_{\mathbb{R}}$ i nazywamy σ-ciałem borelowskim

Przykład 2.

Z aksjomatyki podanej wyżej można wyprowadzić inne własności Σ.

(•) $\Omega \in \Sigma$

(••) $A, B \in \Sigma$, to $A \cap B \in \Sigma$
$A \setminus B \in \Sigma$

ZAD1. Uzasadnić powyższe własności.

(iii) P jest <u>miarą</u> określoną na Σ, czyli zdarzeniach,

a więc

$$\Sigma \ni A \longrightarrow P(A) \in [0,1],$$

gdzie

$(\cdot)$ $P(\emptyset) = 0$

$(\cdot\cdot)$ $\forall_{A \in \Sigma}\ P(A^c) = 1 - P(A)$

$(\cdot\cdot\cdot)$ $\{A_n,\ n \in \mathbb{N}_0 \subset \mathbb{N}\} \subset \Sigma$

(przeliczalna podrodzina zdarzeń)

$\forall_{n \neq m}\ A_n \cap A_m = \emptyset$ (parami rozłączne)

$$P\Big(\bigcup_{n \in \mathbb{N}_0} A_n\Big) = \sum_{n \in \mathbb{N}_0} P(A_n) \ -$$

zasada <u>σ-addytywności</u> P.

P nazywamy <u>funkcją prawdopodobieństwa</u>,

a liczbę $P(A)$ – prawdopodobieństwem zdarzenia A.

Zad1 $\forall A \in \Sigma,\ 0 \leq P(A) \leq 1$.

Uwagi. 1° $\sum_{n \in \mathbb{N}_0} P(A_n)$, gdy $\mathbb{N}_0$ nie jest skończony

oznacza sumę szeregu liczbowego.

2° dla $A, B \in \Sigma$, $A \cap B = \emptyset$ mamy

$$P(A \cup B) = P(A) + P(B).$$

3° dla $A, B \in \Sigma$ dowolnych powyższy wzór

nie jest prawdziwy.

Wtedy

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

ZAD2.

Korzystając z powyższego wyprowadzić wzór

dla $P(A \cup B \cup C)$, $A, B, C \in \Sigma$

4°. P jest funkcją niemalejącą, gdy

$$\forall A, B \in \Sigma \quad A \subseteq B \Longrightarrow P(A) \leq P(B)$$

5°. $P(\Omega) = 1$

ZAD 3. Uzasadnić poprawny wzór.

Przegląd wybranych MPK (Ω, Σ, P)

Model dyskretny.

$\Omega = \{\omega_n, n \in \mathbb{N}_0 \subset \mathbb{N}\}$ – jest dyskretny przeliczalny, jeśli $\mathbb{N}_0$ skończony, to Ω jest skończony.

$\Sigma = \mathcal{P}(\Omega)$ – każdy $A \subset \Omega$ jest zdarzeniem, $A \in \Sigma$.

P określony jest następująco:

(o) $P(\emptyset) \stackrel{df}{=} 0$ (bo tak musi być!)

(o.) $P(\{\omega_n\}) \stackrel{df}{=} p_n \in (0,1)$ $n \in \mathbb{N}_0$

oraz $\sum_{n \in \mathbb{N}_0} p_n = 1$ (suma wag $= 1$).

Wtedy łatwo rozszerzyć tę def na dowolne A.

Np. $A = \{\omega_1, \omega_2, \omega_3\}$, to

$$P(A) = P(\{\omega_1\} \cup \{\omega_2\} \cup \{\omega_3\}) = P(\{\omega_1\}) + P(\{\omega_2\}) + P(\{\omega_3\}) = p_1 + p_2 + p_3.$$

~~A~~ przykłady MD

① 2-punkty

$|\Omega| = 2$, $|\mathcal{P}(\Omega)| = 2^2 = 4$

$\Omega = \{\omega_1, \omega_2\}$, $P(\{\omega_1\}) = p$

$P(\{\omega_2\}) = q = 1 - p$

Np. $\omega_1 = 0$
$\omega_2 = 1$

Uwaga. Prosty powiązać z przykładem rzut monetą.

② $B(n, p)$ – model Bernoulliego, $n \geq 2$
$p \in (0, 1)$

$\Omega = Bin_n$ – zbiór wszystkich ciągów binarnych długości n

$\Omega \ni \omega = (b_1, b_2 \ldots b_n)$, $b_j \in \{0, 1\}$

$\Sigma = \mathcal{P}(\Omega)$

$P(\{\omega\}) = p^k q^{n-k}$, gdzie

$k = \#\{j \; b_j = 1\}$

Weźmy pewne szczególne zdarzenia:

S_k, $k = 0, 1, 2, \ldots, n$, gdzie

$\omega \in S_k \equiv \omega = (b_1, b_2, \ldots, b_n)$ i dokładnie mamy k-jedynek.

Wtedy $|S_k| = \binom{n}{k}$ (dlaczego?)

Stąd $P(S_k) = \binom{n}{k} p^k (1-p)^{n-k}$.

Ale $\{S_0, S_1, \ldots, S_n\}$ tworzy partycję Ω, czyli

$\Omega = S_0 \dot{\cup} S_1 \dot{\cup} \ldots \dot{\cup} S_n$, więc

z addytywności P,

$$1 = P(\Omega) = P\left(\dot{\bigcup_{k=0}^{n}} S_k\right) = \sum_{k=0}^{n} P(S_k),$$ czyli

$$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = 1.$$

Z drugiej strony mamy wzór dwumianowy Newtona, który mówi, że

$$\forall_{a,b \in \mathbb{R}} \quad (a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

Niech $a = p$, $b = q = 1-p$.

Wtedy

$a+b = 1$, co daje oczekiwany wynik!

③ Model Poissona $\mathcal{P}(\lambda)$, $\lambda > 0$

$\Omega = \{0, 1, 2 \dots\}$

$\Sigma = \mathcal{P}(\Omega)$

$$\mathbb{P}(\{k\}) \overset{df}{=} e^{-\lambda} \frac{\lambda^k}{k!}, \quad k \in \Omega.$$

Wtedy mamy liczyć $\sum_{k \geq 0} e^{-\lambda} \frac{\lambda^k}{k!}$ zbieżny jako 1

(dlaczego?)

4) Model dyskretny skończony jednorodny.

$\Omega = \{\omega_1, \omega_2, \dots, \omega_n\} \quad n \geq 2$

$\Sigma = \mathcal{P}(\Omega)$

$P(\{\omega_1\}) = P(\{\omega_2\}) = \dots = P(\{\omega_n\}) = p$

Zatem z addytywności

$$np = 1 \implies p = \frac{1}{n}$$

Wtedy $\forall A \in \Sigma$

$$P(A) = \begin{cases} 0, & A = \emptyset \\ \frac{|A|}{|\Omega|} & A \neq 0 \end{cases}$$

Uwaga. Był to pierwszy w historii tej teorii model rozważany przez Pascala, Fermata i innych (początek XVII w).

Model warunkowy

Bardzo ważny model z punktu widzenia zastosowań.

Dla lepszego zrozumienia podamy jego konstrukcję

Krok 1. Bierzemy dowolny MPK
(Ω, Σ, P)

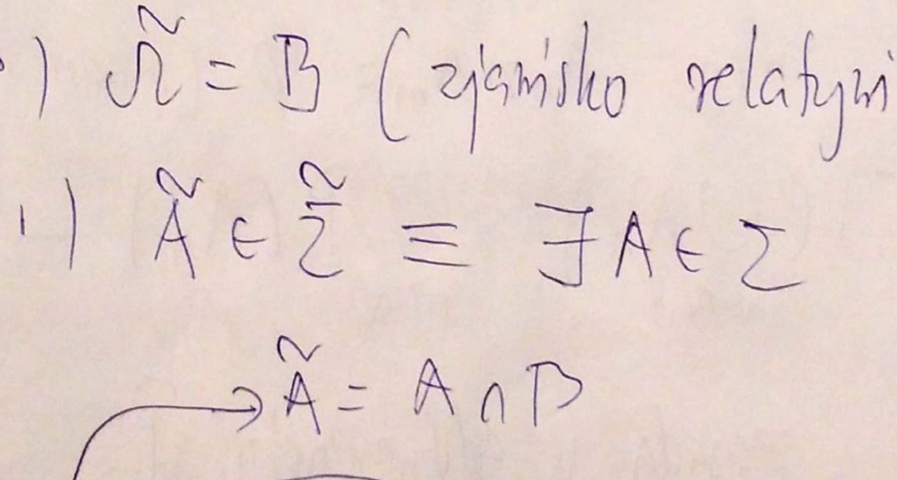
Krok 2 Ustalamy zdarzenie $B \in \Sigma$, $P(B) > 0$

Krok 3. Definiujemy nowy MPK $(\tilde{\Omega}, \tilde{\Sigma}, \tilde{P})$,
gdzie:

(·) $\tilde{\Omega} = B$ (zjawisko relatywizacji)

(··) $\tilde{A} \in \tilde{\Sigma} \equiv \exists A \in \Sigma$

$\tilde{A} = A \cap B$



$\tilde{A}$ nazywamy zdarzeniem „śladowym”

(...) $\tilde{P}$ definiujemy następująco

$$\tilde{\Sigma} \ni \tilde{A} \longrightarrow \tilde{P}(\tilde{A}) \stackrel{df}{=} \frac{P(A \cap B)}{P(B)}.$$

Wtedy $(\tilde{\Omega}, \tilde{\Sigma}, \tilde{P})$ nazywamy <u>modelem warunkowym</u>.

<u>Uwagi</u> 1°. Liczbę $\tilde{P}(\tilde{A})$ nazywamy prawd. warunkowym <u>zdarzenia A pod warunkiem B</u>

Bardzo często piszemy

$$\tilde{P}(\tilde{A}) = P(A|B)$$

2°. Mamy dwa prawdopodobieństwa: P, $\tilde{P}$, zatem dwie liczby dla $A \in \Sigma$,

$$P(A), \quad \tilde{P}(\tilde{A}).$$

~~Które~~ <u>Co oznacza, i nie odzmię się</u>

Sprawdźmy to.

$$P(A) = \tilde{P}(\tilde{A}) \iff$$

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

$$\boxed{P(A \cap B) = P(A)P(B)}$$

Ostatni związek jest fundamentalny w T.P. i zastosowaniach.

Nazywamy jest <u>stochastyczną niezależnością</u>

<u>Zad 3</u>

Uzasadnić, że jeśli A, B są st. niez. to A^c, B i A^c, B^c również.

Wróćmy do prawdop. warunkowego $P(A|B)$.

Zamieńmy rolami A z B. W tym celu należy założyć, że $P(A) > 0$.

<u>Wtedy</u>:

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

i ostatecznie

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

Jest to (słynny) wzór BAYESA.

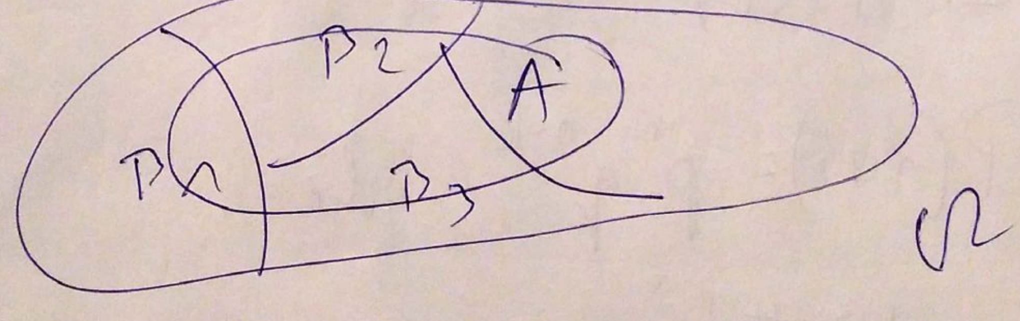
W szczególności jeśli warunki A, B są jednakowo prawdopodobne $(P(A) = P(B))$, to

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

Załóżmy teraz, że mamy partycję warunków, czyli

$$\Omega = B_1 \,\dot\cup\, B_2 \,\dot\cup \dots \dot\cup\, B_n$$

$$P(B_j) > 0, \quad j = 1 \dots n$$



Wtedy dla $A \in \Sigma$ mamy:

$$A = A \cap \Omega = A \cap \left(\bigcup_{j=1}^{n} B_j \right) = \bigcup_{j=1}^{n} (A \cap B_j).$$

Stąd

$$P(A) = \sum_{j=1}^{n} P(A \cap B_j) = \sum_{j=1}^{n} \frac{P(A \cap B_j)}{P(B_j)} P(B_j)$$

i ostatecznie

$$\boxed{P(A) = \sum_{j=1}^{n} P(A \mid B_j) P(B_j)}$$

Jest to wzór na tzw. <u>prawdopodobieństwo zupełne</u>.

Mówimy, że zdarzenie A można <u>mierzyć</u> na dwa sposoby:

jako $P(A)$ (bezpośrednio)

jako $P(A \mid B_j) P(B_j)$ (pośrednio),

o ile <u>dysponujemy partycją</u>.