Materiały do wykładu PMPiSt 29.10.2020
tryb ON-LINE

<u>Temat</u>. Rodzaje <u>rozkładu prawdopodobieństwa</u>. (RP)

MPK jako projekcja ZL wymaga <u>uzupełnienia</u>. Powodów jest kilka:

a) <u>niejednoznaczności</u>: danemu ZL odpowiada na ogół wiele MPK

$$ZL \to \begin{cases} MPK_1 \\ MPK_2 \\ \vdots \\ MPK_n \end{cases}$$

b) na ogół jest trudny do przetwarzania numerycznego, bowiem Ω - zbiór wyników doświadczenia <u>nie musi</u> być <u>zbiorem liczbowym</u>.

Dalej pokażemy to uzupełnienie. Zaczniemy od kilku definicji i przykładów.

Def 1 (Rozkłady prawdopodobieństwa dla przypadku dyskretnego)

Przez dyskretny rozkład prawdopodobieństwa (DRP) rozumiemy każdą funkcję d, gdzie

(i) $d: A \longrightarrow [0,1]$,

$A = \{a_n, n \in \mathbb{N}_0 \subset \mathbb{N}\} \subset \mathbb{R}$

(ii) $\forall_{n \in \mathbb{N}_0} \quad d(a_n) = p_n \in (0,1)$

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

Uwaga: ~~(...)~~ dziedzina d jest przeliczalnym podzbiorem $\mathbb{R}$. W szczególności A może być skończony.

Jeśli A nie jest skończony, to $\sum_{n \in \mathbb{N}_0} p_n$ oznacza szereg liczbowy i jego sumę.

ZAD 1. Przypomnieć sobie podstawowe fakty n/t szeregów liczbowych.

P1 (2-punktowy)

$A = \{v_1, v_2\}$, $v_1 < v_2$ rzeczywiste

$d : A \to [0,1]$, gdzie

$d(v_2) = p \in (0,1)$, $d(v_1) = q = 1-p$.

Zauważmy, że możemy to przedstawić tabelą

	v_1	v_2
d:	q	p

W szczególności, gdy $v_1 = 0$, $v_2 = 1$, to mówimy, że mamy standardowy rozkład 2-punktowy.

P2. Niech $A = \{a_1, a_2, \ldots, a_n\}$

$d(a_j) = p_j \in (0,1)$, $j = 1, \ldots, n$

$\sum_{j=1}^{n} p_j = 1$, co można przedstawić tabelą

	a_1	a_2	$\ldots$	a_n
d:	p_1	p_2	$\ldots$	p_n

P.3. $B(n,p)$ - rozkład Bernoulliego / dwumianowy Newtona.

$$d: A \to [0,1]$$

$$A = \{0,1,2,\dots,n\}, \quad n \geq 2$$

$$P_k = d(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad p \in (0,1)$$

ZAD 2. Dlaczego $\sum_{k=0}^{n} P_k = 1$?

P4. $P(\lambda)$ - rozkład Poissona $(\lambda > 0)$

$$d: A \to [0,1]$$

$$A = \{0,1,2,\dots\} \quad \text{(nie jest skończony!)}$$

$$P_k = d(k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

ZAD 3. Dlaczego $\sum_{k \geq 0} P_k = 1$?

Dalej potrzebujemy pojęcia ~~ścisłej~~ funkcji dystrybuanty rozkładu prawdopodob. (FDRP).

Def 2.

Każdą funkcję F taką, że:

$$F : \mathbb{R} \to [0,1] \text{ i posiadającą}$$

co najmniej następujące własności:

a) $\lim\limits_{t \to -\infty} F = 0$, $\lim\limits_{t \to +\infty} F = 1$

b) F jest niemalejąca: $\forall_{t_1, t_2} \; t_1 < t_2 \Rightarrow F(t_1) \le F(t_2)$

c) F jest (co najmniej) lewo-stronnie ciągła:

$$\forall_{t_0} \lim_{t \to t_0^-} F = F(t_0)$$

nazywamy FDRP.

ZAD 4. Przypomnieć sobie pojęcia: granicy funkcji, ciągłości funkcji

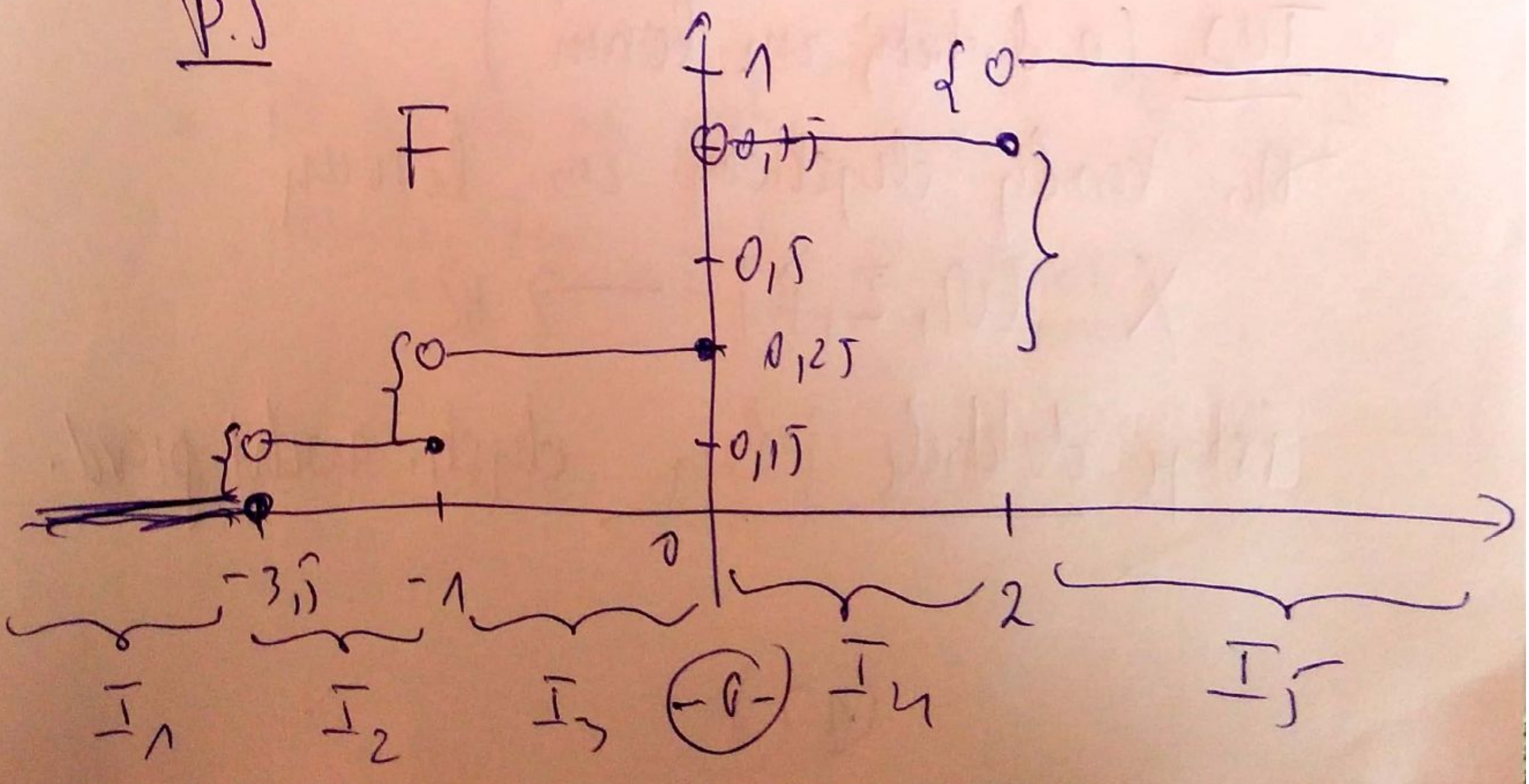
Myśli dodatkowo

d) F jest stała na przedziałach, które się parami rozłączne i sumują się do $\mathbb{R}$ (a więc tworzą partycję $\mathbb{R}$), to

F nazywamy dyskretną FDRP (DFDRP)

Dalej zakładamy, że F jest typu DFDRP oraz (dla uproszczenia), że partycja $\mathbb{R}$ jest skończona. Wtedy postać (wykres) F wygląda jak w poniższym przykładzie:

P.5



Komentarz :

- partycja $\Omega = \{I_1, I_2, I_3, I_4, I_5\}$
- F spełnia warunki (a)–(d) Def 2, a więc jest DFDRP.
- Niech $A = \{-3{,}5, -1, 0, 2\}$ – generuje partycję
- $\forall_{a \in A} \lim_{t \to a^+} F - F(a)$ – miara „skoku"

  np. $a = -1$, „skok" $= 0{,}1$ itd.
- suma wszystkich „skoków" $= 1$.
- Ze względu na postać wykresu DFDRP nazywamy też „funkcją schodkową".

Niech F będzie DFDRP, np. jak w P5.

Zauważmy, że wtedy mamy:

- zbiór $A = \{-0{,}5, -1, 0, 2\}$
- możemy określić

$$d : A \longrightarrow [0,1]$$

$$d(-0{,}5) = (\text{„skok"}) = 0{,}15$$

$$d(1) = 0{,}1$$

$$d(0) = 0{,}5$$

$$d(2) = 0{,}25$$

$$\Sigma = 1$$

- $\sum_{a \in A} d(a) = 1$.

Mamy więc

TW. 1 (O DRP)

Dla każdej DFDRP F istnieje dokładnie

jeden dyskretny r. prawd. $d : A \to [0,1]$, i

$$\forall_{a \in A} \quad d(a) = \lim_{t \to a^+} F - F(a)$$

Weźmy teraz DRP $d: A \to [0,1]$, np.

$A = \{0, -1, -2, -0{,}5\}$

$d(0) = 0{,}5$, $d(-1) = 0{,}1$, $d(2) = 0{,}25$

$d(-0{,}5) = 0{,}15$

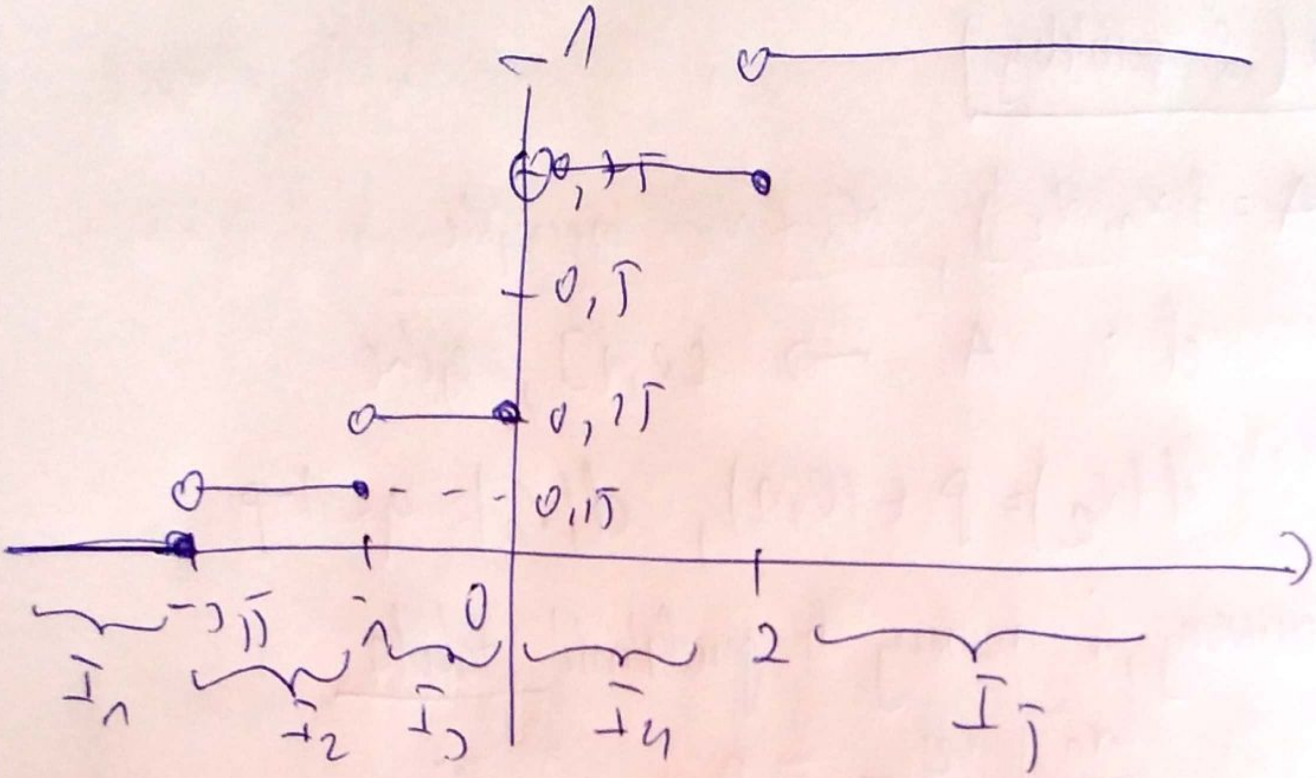
Pokażmy jak zakodować powyższą informację za pomocą F.

Krok 1. Porządkujemy A (rosnąco) i d zapisujemy za pomocą tabeli:

$-0{,}5$	-1	0	2
$0{,}15$	$0{,}1$	$0{,}5$	$0{,}25$

Krok 2. Na układ. współrz. rysujemy:

- partycję



• na każdym obszarze partycji (zaczynając od lewej →) zaznaczamy wykres funkcji stałej o wartościach

$I_1 \to 0$, $I_2 \to 0{,}15$,

$I_3 \to 0{,}15+0{,}1$, ... itd.

jak na rys. wyżej.

Mamy więc

<u>Tw2 (o DRP)</u>

Każdemu d odpowiada dokładnie jedna F typu DFDRP.

WNIOSEK 1.

$d \Leftrightarrow F$, czyli opisy te są sobie równoważne.

PROBLEM.

1) Dlaczego w def. $d(F)$ pojawia się termin "prawdopodobieństwo"?
2) Jaki jest związek pomiędzy ZL a $d(F)$?

Aby uchwycić odp. na 1) & 2) potrzebujemy kolejnego pojęcia – zmiennej losowej.

Def 3 (zmiennej losowej).

Mówimy, że X jest zmienną losową, jeśli

1. $\exists$ ZL, a więc MPK

$$(\Omega, \Sigma, P)$$

(2) $X: \Omega \longrightarrow \mathbb{R}$, taka, że

$$\underset{t \in \mathbb{R}}{\forall} \{\omega \in \Omega : X(\omega) < t\} \in \Sigma$$

Jeśli dodatkowo zbiór

$X(\Omega) = \{X(\omega), \omega \in \Omega\} \subset \mathbb{R}$ jest

przeliczalny, to powiemy, że X jest dyskretną zmienną losową.

Niech dalej X będzie dyskretna. Wtedy

$$X(\Omega) = \{a_n \in \mathbb{R} : n \in \mathbb{N}_0 \subset \mathbb{N}\}.$$

Mamy więc zdarzenia:

$$A_n = \{\omega \in \Omega : X(\omega) = a_n\}, \ n \in \mathbb{N}_0.$$

Zauważmy, że: (i) $A_n \cap A_m = \emptyset$, $n \neq m$

(ii) $\bigcup_{n \in \mathbb{N}_0} A_n = \Omega$

ZADJ

Uzasadnić, że $\{A_n, n \in \mathbb{N}_0\}$ nie spełnia (i') & (ii').

ZATEM $A = \{A_n, n \in \mathbb{N}_0\}$ jest partycją Ω.

Definiujmy: $p_n = P(A_n), n \in \mathbb{N}_0$.

Mamy zatem:

$$d : A \longrightarrow [0,1],$$

$$d(a_n) = p_n$$

$$\sum_{n \in \mathbb{N}_0} p_n = 1.$$

TWJ (o dyskretnej zm. losowej)

Dla każdej dyskretnej zm. losowej

$$X : (\Omega, \Sigma, P) \longrightarrow \mathbb{R}$$

istnieje dokładnie jeden dysk. rozkład prawd.

d_X (a mi DFDRP F_X), i

(i) $\forall_{n \in \mathbb{N}_0} \; p_n = d(a_n) = P(\{\omega \in \Omega : X(\omega) = a_n\})$

(ii) przyporządkowanie

$\mathbb{R} \ni t \longrightarrow P(\{\omega \in \Omega : X(\omega) < t\}) \in [0,1]$

ma wszystkie charakterystyczne f. schodkowej, której oznaczamy przez F_X, a zatem

$F_X(t) = P(\{\omega \in \Omega : X(\omega) < t\})$

(iii) $\forall_{a_n} \; \lim_{t \to a_n^+} F - F(a_n) = d(a_n) = p_n =$

$= P(\{\omega \in \Omega : X(\omega) = a_n\})$

Aby teoria zaczęła działać potrzebujemy tw. odwrotnego do Tw 3.
Zachodzi
Tw 4

Dla każdego d (a więc i F) istnieje co najmniej jedno ZL i jego MPK (Ω, Σ, P) oraz odwzorowanie

$$X : \Omega \longrightarrow \mathbb{R}, \text{ gdzie}$$

$$X(\Omega) = A \text{ – dziedzina f. } d$$

mająca własności dyskretnej zm. losowej i zachodzą własności (i) — (iii) Tw. 3.

Uwagi.

1) X pełni rolę translatora:

$$\begin{tikzcd} \Omega \ni \omega \arrow[r] \arrow[d] & X(\omega) \in \mathbb{R} \arrow[d] \\ \text{„stany"} & \text{„nowy"} \end{tikzcd}$$

wynik obs.

ZL

2) Dalej mówiąc o rozkładzie prawdopodobieństwa mamy na myśli X i jej opis za pomocą F_X lub d_X w rozumieniu Tw. 3 i Tw. 4.