Poprzedni wykład

Prawdopodobieństwo warunkowe

$$P(A\,|\,B) \equiv \frac{P(A \cap B)}{P(B)} \qquad P(B) > 0$$

Reguła całkowitego prawdopodobieństwa (B_i tworzą partycję)

$$P(A) = \sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} P(A\,|\,B_i)P(B_i)$$

Wzór Bayesa

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

Analiza bayesowska

Analiza bayesowska (jakościowy przykład wyników)

- Rzut monetą: do wyznaczenia P(O)
  a priori + eksperyment 90 orłów w 100 rzutach → a posteriori

1. P(O)=0.5
   → analiza doprowadzi do P(O)≈0.86
2. P(O) – rozkład wokół 0.5
   → analiza doprowadzi do rozkładu wokół P(O) ≈0.86
3. P(O) – rozkład wokół 0.3
   → analiza doprowadzi do rozkładu wokół P(O) ≈0.80
4. P(O) – rozkład jednorodny na [0,1]
   → analiza doprowadzi do rozkładu wokół P(O) ≈0.90

Im więcej danych tym bardziej różne P(O) a priori będą zbiegać się do tego samego P(O) a posteriori.

Wzór „łańcuchowy”

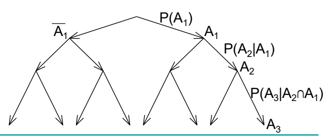
Zał. $P(A_1 \cap A_2 \cap \ldots \cap A_{n-1}) > 0$

Wtedy zachodzi wzór:

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1)P(A_2\,|\,A_1)P(A_3\,|\,A_1 \cap A_2)\times$$
$$\times \ldots \times P(A_n\,|\,A_1 \cap A_2 \cap \ldots \cap A_{n-1})$$

-dowód – iteracja definicji prawdopodobieństwa warunkowego

-uzasadnia metodę „drzewka”



Niezależność zdarzeń

- Dwa zdarzenia A i B (P(A)>0 i P(B)>0) są **niezależne** gdy:

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

- Warunek wystarczający i konieczny niezależności:

$$P(A \cap B) = P(A)P(B)$$

Niezależność zdarzeń

- Dowód ⇨

$$P(A\,|\,B) = \frac{P(A \cap B)}{P(B)} \quad \Big/ \cdot P(B)$$
$$P(A\,|\,B)P(B) = P(A \cap B)$$
$$P(A)P(B) = P(A \cap B)$$

- Dowód ⇦

$$P(A \cap B) = P(A)P(B)$$
$$P(A\,|\,B)P(B) = P(A)P(B) \quad /: P(B)$$
$$P(A\,|\,B) = P(A)$$

- Analogicznie dla drugiej równości *P(B|A)=P(B)*

Niezależność zdarzeń

- Zdarzenia niezależne nie muszą być rozłączne
  Tw. Jeżeli zdarzenia A i B są rozłączne i niezależne to P(A)=0 lub P(B)=0 lub P(A)=P(B)=0
  Dowód:

$$P(A \cap B) = P(\phi) = 0 \quad \text{(z rozłączności)}$$
$$P(A \cap B) = P(A)P(B) \quad \text{(z niezależności)}$$

czyli $P(A)P(B) = 0$

- Przykład: losowanie liczb całkowitych
- Nie należy mylić **niezależności** zdarzeń z **rozłącznością** zdarzeń!

Przykład: losowanie liczb całkowitych

Zdarzenie A – liczba jest podzielna przez 3
Zdarzenie B – liczba jest podzielna przez 7

- Losujemy liczbę całkowitą z przedziału [1,20]

A={3,6,9,12,15,18} B={7,14} A∩B=ϕ

$P(A)=6/20 \quad P(B)=2/20 \quad P(A\cap B)=0 \quad P(A)\cdot P(B)=6\cdot 2/20^2=12/20^2\neq P(A\cap B)$

→ Zdarzenia A i B są rozłączne i nie są niezależne

- Losujemy liczbę całkowitą z przedziału [1,21]

A={3,6,9,12,15,18,21} B={7,14,21} A∩B={21}

$P(A)=7/21 \quad P(B)=3/21 \quad P(A\cap B)=1/21 \quad P(A)\cdot P(B)=7\cdot 3/21^2=1/21=P(A\cap B)$

→ Zdarzenia A i B nie są rozłączne i są niezależne

- Losujemy liczbę całkowitą z przedziału [1,22]

A={3,6,9,12,15,18,21} B={7,14,21} A∩B={21}

$P(A)=7/22 \quad P(B)=3/22 \quad P(A\cap B)=1/22 \quad P(A)\cdot P(B)=7\cdot 3/22^2=21/22^2\neq P(A\cap B)$

→ Zdarzenia A i B nie są rozłączne i nie są niezależne

Jak dobrać przedział otrzymać czwartą możliwość ?

Przykład: losowanie liczb całkowitych

Zdarzenie A – liczba jest podzielna przez 3
Zdarzenie B – liczba jest podzielna przez 7

- Losujemy liczbę całkowitą z przedziału [1,6]

A={3,6} B=ϕ A∩B=ϕ

$P(A)=2/6 \quad P(B)=0 \quad P(A\cap B)=0 \quad P(A)\cdot P(B)=2\cdot 0/6=0=P(A\cap B)$

→ Zdarzenia A i B są rozłączne i są niezależne

→ Nie należy mylić **niezależności** zdarzeń z **rozłącznością** zdarzeń!

Niezależność zdarzeń

- Warunek wystarczający i konieczny **niezależności warunkowej**: $P(A\cap B\,|\,C)=P(A\,|\,C)P(B\,|\,C)$
- Nie należy mylić niezależności zdarzeń z niezależnością warunkową zdarzeń!

- **Zdarzenia $A_1, A_2, \ldots, A_n$ są parami niezależne** gdy:

$$\forall i\neq j: P(A_i\cap A_j)=P(A_i)P(A_j)$$

((n-1)n/2 warunków)

- **Zdarzenia $A_1, A_2, \ldots, A_n$ są globalnie niezależne** gdy:

$$\forall k\leq n: P(A_{k1}\cap A_{k2}\cap\ldots\cap A_{kk})=P(A_{k1})P(A_{k2})\ldots P(A_{kk})$$

gdzie A_{kj} to dowolne, różne zdarzenia spośród zdarzeń $A_1, \ldots, A_n$.

Zmienne losowe

Syt. Eksperyment losowy E i związana z nim przestrzeń zdarzeń S, zbudowana ze zdarzeń s.

- **Zmienną losową** nazywamy funkcję X, która każdemu zdarzeniu s∈S przyporządkowuje liczbę rzeczywistą X(s)=x

$$\boxed{\forall s\in S \xrightarrow{X(s)} x\in S_X}$$

- X – nazwa zmiennej losowej (funkcji)
- x – wartości zmiennej losowej X
- S_X – zbiór wartości zmiennej losowej X (dyskretny → dyskretna zmienna losowa; ciągły → ciągła zmienna losowa; mieszany → mieszana zmienna losowa)

Np. rzut kostką; X=liczba wyrzuconych oczek; S_X={1,2,3,4,5,6}
Np. rzut kostką; X=kwadrat liczby wyrzuconych oczek; S_X={1,4,9,16,25,36}
Np. rzut kostką; X=odległość od „4"; S_X={3,2,1,0,1,2}={0,1,2,3}
Np. rzut kostką; X=liczba rzutów aż wypadnie sześć oczek; S_X={1,2,3,…}

Zmienne losowe

$$\boxed{\forall s\in S \xrightarrow{X(s)} x\in S_X}$$

Zdarzeniem ze względu na S_X nazywamy podzbiór S_X.

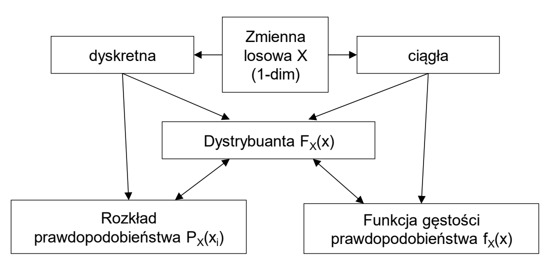
- Zdarzenie A∈S i zbiór wartości $A_X \in S_X$ są równoważne (ekwiwalentne) gdy

$$\boxed{\begin{array}{l}\forall s\in A: X(s)\in A_X \\ \forall s_X\in A_X: \exists s\in S: X(s)\in A_X\end{array}}$$

- Inaczej mówiąc zdarzenie A i zdarzenie (ze względu na S_X) A_X są równoważne, wówczas $P(A)=P(A_X)$
- Czyli **zamiast operować na zdarzeniach możemy operować na zbiorach wartości w przestrzeni S_X**

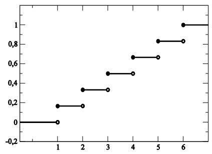
Przykład: Eksperyment: rzucamy trzy razy uczciwą monetą
S={OOO,OOR,ORO,ORR,ROO,RRO,ROR,RRR}, zdarzenia elementarne równoprawdopodobne
Zdarzenie A: wypadła parzysta liczba orłów
Zdarzenia elementarne sprzyjające zdarzeniu A: {OOR,ORO,ROO,RRR}
P(A)=4/8=1/2
Zmienna losowa X – liczba orłów w trzech rzutach
Zbiór wartości X: S_X={0,1,2,3}, P(X)=(1/8, 3/8, 3/8, 1/8)
Zdarzenie ze względu na S_X: A_X – liczba orłów w trzech rzutach jest parzysta: {0,2}
$P(A_X)$=1/8+3/8=1/2
Czyli zdarzenie „wypadła parzysta liczba orłów" jest równoważne w tym eksperymencie zdarzeniu „liczba orłów wynosiła 0 lub 2"

Zmienna losowa i funkcje ją opisujące



Dystrybuanta

- **Dystrybuantą** zmiennej losowej X nazywamy funkcję dającą prawdopodobieństwo otrzymania wartości zmiennej losowej mniejszej bądź równej od danej wartości x: $\boxed{F_X(x) = P(X \le x)}$
- Przykład: Dystrybuanta w rzucie kostką



Dystrybuanta - własności

- $0 \le F_X(x) \le 1$ (gdyż $F_X(x)$ jest sumą prawdopodobieństw)
- $\lim_{x\to\infty} F_X(x) = 1$ (gdyż zdarzenie x≤+∞ jest pewne)
- $\lim_{x\to-\infty} F_X(x) = 0$ (gdyż zdarzenie poniżej min{x} jest niemożliwe)
- $F_X(x)$ jest niemalejąca tzn. $x_1 < x_2 \Rightarrow F_x(x_1) \le F_x(x_2)$ gdyż zdarzenie $\{X \le x_1\} \subset \{X \le x_2\}$
- $F_X(x)$ jest ciągła prawostronnie $\lim_{\delta\to 0^+} (F_X(x+\delta) - F_X(x)) = 0$
- $F_X(x)$ jest ciągła dla ciągłych zmiennych losowych
- $F_X(x)$ jest bezwymiarowa

Związek dystrybuanty z prawdopodobieństwem

- Tw. $\boxed{P(a < X \le b) = F_X(b) - F_X(a)}$
- Dow:

Zdarzenie $\{X \le b\} \equiv \underbrace{\{X \le a\} \cup \{a < X \le b\}}_{\text{rozłączne}} \qquad /P(\,)$

$$P(\{X \le b\}) = P(\{X \le a\}) + P(\{a < X \le b\})$$
$$F_X(b) = F_X(a) + P(\{a < X \le b\})$$
$$P(\{a < X \le b\}) = F_X(b) - F_X(a)$$

Związek dystrybuanty z prawdopodobieństwem - wnioski

- Wniosek 1

Weźmy $a = x - \delta, \quad b = x$

$P(x - \delta < X \le x) = F_X(x) - F_X(x-\delta) \qquad /\lim_{\delta\to 0}$

$P(X = x) = F_X(x) - \lim_{\delta\to 0} F_X(x-\delta)$

ale dla ciągłej zmiennej losowej $\lim_{\delta\to 0} F_X(x-\delta) = F_X(x)$

zatem $\forall x: \quad P(X = x) = F_X(x) - F_X(x) = 0$

- Wniosek 2 (z Wniosku 1)

$P(a \le X \le b) = P(a < X \le b) = P(a \le X < b) = P(a < X < b)$

(związek z prawdopodobieństwem geometrycznym)

Dystrybuanta warunkowa

Syt:

X-zmienna losowa,

A – zdarzenie takie, że P(A)>0

- **Dystrybuantą warunkową** zmiennej losowej X pod warunkiem, że zaszło zdarzenie A nazywamy funkcję

$$F_X(x \mid A) = \frac{P(\{X \le x\} \cap A)}{P(A)}$$

Funkcja rozkładu prawdopodobieństwa

- Syt:

X – **dyskretna** zmienna losowa

S_X – odpowiedni zbiór wartości zmiennej X

- **Funkcją rozkładu prawdopodobieństwa** $P_X(x_k)$ nazywamy funkcję przyporządkowującą wartościom zmiennej losowej prawdopodobieństwo jej wystąpienia

$\forall k: \; P_X(x_k) = P(X = x_k)$

Funkcja rozkładu prawdopodobieństwa - własności

- $P_X(x_k) \geq 0$ (własność prawdopodobieństwa)
- $\sum_k P_X(x_k) = 1$ gdyż $P(S_X) = P(S) = 1$
- dla $x_1 < x_2 < \ldots$ oraz $k > 1$ mamy

  $P_X(x_k) = F_X(x_k) - F_X(x_{k-1})$
- $P_X(x_k)$ jest bezwymiarowe

Funkcja gęstości prawdopodobieństwa

- Syt:

  X – **ciągła** zmienna losowa o różniczkowalnej dystrybuancie

  S_X – odpowiedni zbiór wartości zmiennej X
- **Funkcją gęstości prawdopodobieństwa** $f_X(x)$ nazywamy funkcję $\boxed{f_X(x) = \frac{d}{dx} F_X(x)}$

$$P(x < X \leq x + \varepsilon) = F_X(x+\varepsilon) - F_X(x) =$$

$$= \frac{F_X(x+\varepsilon) - F_X(x)}{\varepsilon} \varepsilon \xrightarrow{\lim_{\varepsilon \to 0}} \frac{dF_X(x)}{dx} \varepsilon$$

Funkcja gęstości prawdopodobieństwa - własności

- $f_X(x) \geq 0$
- $P(x < X \leq x + dx) = P(x \leq X < x + dx) = \ldots = f_X(x)dx$
- $\int_{-\infty}^{x} f_X(t)dt = \int_{-\infty}^{x} \frac{d}{dt} F_X(t)dt = F_X(t)\Big|_{-\infty}^{x} = F_X(x) - 0 = F_X(x)$
- $\int_{-\infty}^{+\infty} f_X(t)dt = F_X(t)\Big|_{-\infty}^{+\infty} = 1 - 0 = 1$
- $P(a < X < b) = F_X(b) - F_X(a) = \int_a^b f_X(t)dt$
- $[\, f_X(t)\,] = \frac{1}{[x]}$ (wymiar)

Funkcja gęstości prawdopodobieństwa własności

- Uwaga 1:

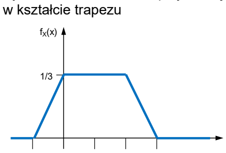
  Jeżeli zmienna losowa X przybiera wartości tylko z pewnego obszaru to poza tym obszarem definiujemy $f_X(x)=0$
- Uwaga 2:

  Zarówno funkcja rozkładu prawdopodobieństwa jak i funkcja gęstości prawdopodobieństwa muszą być znormalizowane, gdy tak nie jest należy przeprowadzić normalizację $f_X(x) \to N \cdot f_X(x)$
- **Warunkową funkcją gęstości prawdopodobieństwa** nazywamy funkcję

$$f_X(x \mid A) = \frac{d}{dx} F_X(x \mid A)$$

Przykład: FGP (trapez) i odwracanie dystrybuanty

Syt: Zmienna losowa X opisywana jest funkcją gęstości prawdopodobieństwa w kształcie trapezu



$$f_X(x) = \begin{cases} 0 & dla\ x \in (-\infty, -1] \\ \frac{1}{3}x + \frac{1}{3} & dla\ x \in (-1, 0] \\ \frac{1}{3} & dla\ x \in (0, 2] \\ \frac{-1}{3}x + 1 & dla\ x \in (2, 3] \\ 0 & dla\ x \in (3, +\infty] \end{cases}$$

- Wysokość trapezu obliczamy z warunku normalizacji

$$Pole\ trapezu = \frac{(a+b)\cdot h}{2} = \frac{(4+2)\cdot h}{2} = 3h \;\to\; 1 = 3h \;\to\; h = \frac{1}{3}$$

- Funkcję gęstości prawdopodobieństwa obliczamy wiedząc przez jakie punkty przechodzą poszczególne odcinki