Statystyczna analiza danych SAD

Wykład 2

Elementy rachunku prawdopodobieństwa

(MODELE PROBABILISTYCZNE)

Elementy rachunku prawdopodobieństwa

Rachunek prawdopodobieństwa bada zdarzenia, które są wynikami doświadczeń losowych. Tworzy modele matematyczne zjawisk losowych.

Przykładami doświadczeń losowych są gry hazardowe.

Ich analiza (połowa XVII wieku – prace Blaise Pascala i Pierre de Fermata) dała początek teorii prawdopodobieństwa.

Elementy rachunku prawdopodobieństwa

Podstawowe pojęcia:

- doświadczenie losowe
- zdarzenie elementarne
- zdarzenie losowe
- prawdopodobieństwo zdarzenia – aksjomaty
- prawdopodobieństwo warunkowe
- niezależność zdarzeń losowych

Własności prawdopodobieństwa

Wzór Bayes'a

Definicja 1

Doświadczenie losowe:

- może być powtarzane w tych samych warunkach,
- wynik nie jest znany przed wykonaniem doświadczenia,
- znany jest zbiór wszystkich możliwych wyników - opisany przed przeprowadzeniem doświadczenia.

Elementy rachunku prawdopodobieństwa

Przykłady:

Doświadczenie	**Wyniki**
podanie lekarstwa	{leczy, nie leczy}
czas życia elementu (np. procesora)	$[0, \infty)$
czas naprawy elementu	$[0, \infty)$
liczba błędów w aplikacji	{0,1,2,...}
liczba orłów w 100 rzutach monetą	{0,1,2,...,100}
liczba samochodów na autostradzie w godzinach szczytu	{0,1,2,...}

Elementy rachunku prawdopodobieństwa

Definicja 2

- **Przestrzenią zdarzeń elementarnych** (przestrzenią próbkową) nazywamy zbiór wszystkich możliwych wyników doświadczenia losowego i oznaczamy symbolem S.
- **Zdarzeniem elementarnym** nazywamy każdy element przestrzeni S ($s \in S$); s - niepodzielny (pojedynczy) wynik doświadczenia losowego.
- **Zdarzenie to** podzbiór przestrzeni S ($A \subset S$, $B \subset S$).
- **Zdarzenie** A **zaszło**, gdy wynik doświadczenia losowego (zdarzenie elementarne) jest elementem zbioru A.

Przykłady:

- Rzut monetą: $S = \{O, R\}$ - skończona,

  Podzbiory S (zdarzenia): $\varnothing$, $\{O\}$, $\{R\}$, S.

- Podwójny rzut kostką sześcienną:

  $S = \{(i, j): i, j = 1,2,\ldots,6\}$ - skończona.

  Np. $A = \{$suma oczek nieparzysta mniejszych od $6\}$

  $= \{ (1,2), (1,4), (3,2), (2,1), (4,1), (2,3) \}$.

Elementy rachunku prawdopodobieństwa

- Czas obsługi klienta w systemie masowej obsługi

  (np. sieć telefoniczna, komputerowa, czas mierzony w min.):

  $S = [0, \infty)$ - nieskończona, nieprzeliczalna.

  Zdarzenia – podprzedziały S, np.: $A = [0, 5)$, $B = [0, T)$,

  $C = [0, \infty)$, (A = {obsługa klienta trwa krócej niż 5 min.})

- Liczba awarii urządzenia w określonym czasie:

  $S = \{0,1,2,\ldots\}$ - nieskończona, przeliczalna.

Zdarzenia utożsamiamy ze zbiorami stąd:

rachunek zdarzeń $\Leftrightarrow$ **rachunek zbiorów.**

Definicja 3

- **Zdarzenie przeciwne do** A $\Leftrightarrow$ dopełnienie zbioru A

$$A' = S - A$$

- **Iloczyn zdarzeń** A i B $\Leftrightarrow$ iloczyn zbiorów A i B

$$A \cap B$$

- **Suma zdarzeń** A i B $\Leftrightarrow$ suma zbiorów A i B:

$$A \cup B$$

- **Różnica zdarzeń** A i B $\Leftrightarrow$ różnica zbiorów A i B:

$$A - B$$

Elementy rachunku prawdopodobieństwa

- Zdarzenia A i B **wzajemnie się wykluczają,** jeśli $A \cap B = \varnothing$, gdzie $\varnothing$ jest zbiorem pustym.

- **Zajście zdarzenia** A **pociąga za sobą zajście zdarzenia** B, jeśli $A \subset B$

.

- **Diagramy Venna -** ilustracja działań na zdarzeniach.

Przykłady:

- Dwukrotny rzut monetą: S= {OO,OR,RO,RR}. Niech

  A={orzeł w I rzucie} = {OO, OR} , B ={OR,RR}

  C= {orzeł w II rzucie}={OO, RO}, D={OO}

$$A \cup B = \{OO, OR, RR\}, A \cap B = \{OR\},$$

$$B \cup C = S, B \cap C = \emptyset, B - A = \{RR\},$$

$$D \subset A, D \subset C$$

Przykłady:

- Dwukrotny rzut monetą: S= {OO,OR,RO,RR}. Niech

  A={orzeł w I rzucie} = {OO, OR} , B ={OR,RR}

  C= {orzeł w II rzucie}={OO, RO}, D={OO}

$$A \cup B = \{\mathrm{OO}, \mathrm{OR}, \mathrm{RR}\}, A \cap B = \{\mathrm{OR}\},$$

$$B \cup C = S, B \cap C = \emptyset, B - A = \{\mathrm{RR}\},$$

$$D \subset A, D \subset C$$

Elementy rachunku prawdopodobieństwa

- Doświadczenie polega na rzutach monetą aż do momentu wypadnięcia orła po raz pierwszy. Wówczas

  S= {O,RO,RRO,RRRO,…}.

A={wykonano <6 rzutów} = {O, RO,RRO,RRRO,RRRRO} ,

B ={wyrzucono reszkę} = S – {O}.

Elementy rachunku prawdopodobieństwa

Definicja 4

Zdarzenia $A_1, A_2, A_3, \ldots$ **wzajemnie się wykluczają**, jeśli dowolne dwa zdarzenia A_i oraz A_j, $i \neq j$, wzajemnie się wykluczają: $A_i \cap A_j = \emptyset$, dla dowolnych $i, j, i \neq j$.

Uwaga

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

$$A - B = A \cap B' \qquad B - A = B \cap A'.$$

,

PRAWDOPODOBIEŃSTWO

Eksperyment Karla Pearsona (1857-1936):

Wykonał 24 000 rzutów monetą. Częstość wyrzucenia orła wyniosła

n/N = 0,5005

Szansa (**prawdopodobieństwo**) zajścia zdarzenia A = graniczna częstość wystąpienia A przy nieograniczenie rosnącej liczbie powtórzeń doświadczenia losowego = statystyczna definicja prawdopodobieństwa (Richard von Mises 1919)

Definicja 5 (aksjomatyczna prawdopodobieństwa, A. Kołmogorow, 1933)

(A1) Dla każdego $A, A \subset S,$

$$0 \leq P(A) \leq 1$$

(A2) $P(S) = 1, \qquad P(\emptyset) = 0$

(A3) Jeśli zdarzenia $A_1, A_2, A_3, \ldots$ wzajemnie się wykluczają, to

$$P(A_1 \cup A_2 \cup A_3 \ldots) = P(A_1) + P(A_2) + P(A_3) + \ldots$$

Elementy rachunku prawdopodobieństwa

<u>Uwaga</u> Niech

$$A_{n+1} = A_{n+2} = \cdots = \emptyset$$

Z **(A2)** i **(A3)** wynika

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + \ldots P(A_n)$$

<u>Stwierdzenie 1</u> Niech $S = \{s_1, s_2, \ldots, s_M\}$.

Jeśli $P(\{s_1\}) = P(\{s_2\}) = \ldots = P(\{s_M\})$,
to dla dowolnego m - elementowego podzbioru A zbioru S zachodzi

- $$P(A) = \frac{m}{M}$$
- $$P(A) = \sum_{s_i \in A} P(\{s_i\})$$

D.

- Z aksjomatów **(A2), (A3)** i założenia o jednakowym prawdopodobieństwie jednoelementowych zdarzeń:

$$1 = P(S) = P(\{s_1\}) + P(\{s_2\}) + \ldots P(\{s_M\}) = M \times P(\{s_i\})$$

gdzie *i* = 1,2,..., *M*. Stąd $P(\{s_i\}) = 1/M$.

- $A = \{s_{i_1}, s_{i_2}, \ldots, s_{i_m}\}, \; 1 \le i_1 < i_2 < \ldots < i_m \le M,$

$$P(A) = P(\{s_{i_1}\}) + P(\{s_{i_2}\}) + \ldots + P(\{s_{i_m}\}) = m \times \frac{1}{M} = \frac{m}{M}.$$

Przykład

Oblicz $P(A)$, gdzie A = {wypadnięcie orła po raz pierwszy za trzecim razem, w trzech rzutach monetą symetryczną}.

S = {OOO, OOR, ORO, ROO, ORR, ROR, RRO, RRR}, A={RRO}

Liczebność S = 2^3 = 8, liczebność A = 1.

Zatem $P(A) = \frac{1}{8}$.

Wzory obliczeniowe kombinatoryki

Reguła iloczynu: jeśli w dwuetapowym doświadczeniu na pierwszym etapie można uzyskać r różnych wyników a na drugim etapie niezależnie k różnych wyników to liczba wszystkich wyników doświadczenia wynosi

$$r \cdot k$$

Przykład

Liczba liczb dwucyfrowych =   9x10=90

Liczba liczb dwucyfrowych podzielnych przez 5 = 9x2=18

Wzory obliczeniowe kombinatoryki

Kombinacja (k – elementowa) to k - elementowy podzbiór zbioru n – elementowego.

Liczba k - elementowych kombinacji zbioru n - elementowego,

$$C(n,k) = \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

$$0! = 1.$$

Przykład Ile jest możliwych wyborów 2 asów z talii 52 kart?

$$C(4,2) = \binom{4}{2} = \frac{4!}{2! \times 2!} = 6.$$

$$V(n,k) = n(n-1)(n-2)\times\ldots\times(n-k+1) =$$

$C(n,k)\times k!$ = liczba k - elementowych wariacji bez powtórzeń n - elementowego zbioru, $k \le n$

<u>Przykład</u> Ile jest możliwych sposobów opuszczenia windy przez 5 osób w bloku 10 –cio piętrowym w taki sposób, że nie ma osób wysiadających na tych samych piętrach ?

$$V(10,5) = 10\times 9\times 8\times 7\times 6 = 30240.$$

Elementy rachunku prawdopodobieństwa

Zadanie. Losujemy 13 kart z talii 52 kart (np. w brydżu).

$$P(\text{zestaw 6 pików}) = \frac{\binom{13}{6}\binom{39}{7}}{\binom{52}{13}} = ?$$

$$P(\text{4 karty tej samej wysokości}) =$$

$$\frac{\binom{13}{1}\binom{4}{4}\binom{48}{9}}{\binom{52}{13}} = ?$$

Elementy rachunku prawdopodobieństwa

Twierdzenie 1

Niech $A \subset S$. Wówczas

$$\boxed{P(A') = 1 - P(A)}$$

Dowód.

$A \bigcup A' = S$, $A \bigcap A' = \emptyset$.

Z definicji 5: $P(A \bigcup A') = P(A) + P(A') = P(S) = 1$, $\square$

Przykład.

Prawdopodobieństwo wyrzucenia co najmniej jednego orła w trzech rzutach monetą symetryczną:

$$1 - \frac{1}{8} = \frac{7}{8}.$$

Twierdzenie 2

Niech $A \subset S$ i $B \subset S$ oraz $A \subset B$. Wówczas

$$\boxed{P(A) \leq P(B)}.$$

Dowód

$B = A \bigcup (B - A)$ oraz $A \bigcap (B - A) = \emptyset$.

Zatem z definicji 5:

$$P(B) = P(A) + P(B - A) \geq P(A), \text{ gdyż } P(B - A) \geq 0. \quad \square$$

Elementy rachunku prawdopodobieństwa

Twierdzenie 3 Niech $A \subset S$ i $B \subset S$. Wówczas

$$P(A \bigcup B) = P(A) + P(B) - P(A \bigcap B),$$

$$P(A \bigcup B) = P(A) + P(B), \text{ jeśli } A \bigcap B = \emptyset.$$

Dowód

$$A \bigcup B = (A \bigcap B) \bigcup (A - B) \bigcup (B - A)$$

$$P(A \bigcup B) = P(A \bigcap B) + P(A - B) + P(B - A)$$

Elementy rachunku prawdopodobieństwa

$$A = A \bigcap S = A \bigcap (B \bigcup B') = (A \bigcap B) \bigcup (A \bigcap B')$$

$$A - B = A \bigcap B'.$$

$$P(A) = P(A \bigcap B) + P(A - B)$$

$$P(A - B) = P(A) - P(A \bigcap B)$$ Podobnie

$$P(B - A) = P(B) - P(A \bigcap B)$$

Podstawiając prawe strony powyższych dwu równości do równości w ramce otrzymujemy tezę twierdzenia. □

Prawdopodobieństwo warunkowe

- $S = \{s_1, s_2, \ldots, s_M\}, \; P(\{s_i\}) = \frac{1}{M}, \; i = 1,2,..,M,$
- $A \subset S, \; B \subset S,$
- $n(A), n(B), n(A \cap B)$ - liczności zdarzeń sprzyjających
- $P(B|A) = \frac{n(A \cap B)}{n(A)} = \frac{n(A \cap B)/n(S)}{n(A)/n(S)} = \frac{P(A \cap B)}{P(A)}$
- $P(B|A) = \frac{P(A \cap B)}{P(A)}$ - prawdopodobieństwo

warunkowe zajścia zdarzenia *B* pod warunkiem zajścia *A*.

Definicja 6

Niech $A \subset S$ i $B \subset S$, $P(A) > 0$.

Prawdopodobieństwo warunkowe zdarzenia B

pod warunkiem zajścia zdarzenia A dane jest wzorem

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

Przykład

Obliczono, że 70% studentów zdało egzamin z matematyki w czasie sesji, natomiast 30% zdało egzamin z matematyki i angielskiego. Jakie jest prawdopodobieństwo, że student, który zdał egzamin z matematyki zdał również egzamin z angielskiego?

Rozwiązanie.

Niech M = "losowo wybrany student zdał matematykę",

A = "losowo wybrany student zdał angielski".

$$P(A|M) = \frac{P(A \cap M)}{P(M)} = \frac{0,3}{0,7} = \frac{3}{7}$$

Twierdzenie 4

Niech $A \subset S$ i $B \subset S$, oraz $P(A) > 0$ i $P(B) > 0$. Wówczas

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

Przykład

Urna zawiera 6 kul czarnych i 4 białe. Losujemy 2 kule bez zwracania. Niech

A = { 2 kule czarne },

B = { 1 kula czarna i 1 kula biała }.

C_i = { w i-tym ciągnieniu kula czarna },

B_i = { w i-tym ciągnieniu kula biała}, $i = 1, 2$.

$$P(A) = P(C_1 \cap C_2) = P(C_1)P(C_2 | C_1) = \frac{6}{10} \times \frac{5}{9} = \frac{1}{3},$$

$$P(B) = P[(B_1 \cap C_2) \cup (C_1 \cap B_2)] = P(B_1 \cap C_2) + P(C_1 \cap B_2) =$$

$$P(C_2 | B_1)P(B_1) + P(B_2 | C_1)P(C_1) = \frac{6}{9} \times \frac{4}{10} + \frac{4}{9} \times \frac{6}{10} = \frac{8}{15}.$$

$$P(A) = \frac{\binom{6}{2}}{\binom{10}{2}} = \frac{1}{3}, \qquad P(B) = \frac{\binom{6}{1}\binom{4}{1}}{\binom{10}{2}} = \frac{8}{15}.$$

Reguła wielokrotnego warunkowania

$$P(A \cap B \cap C) = P[C \cap (A \cap B)] =$$

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

$$P(C | A \cap B) P(B | A) P(A)$$

Dla k zdarzeń losowych:

$$P(A_1 \cap A_2 \cap \ldots \cap A_k) = P(A_k | A_{k-1} \cap A_{k-2} \cap \ldots \cap A_1) \times$$

$$\times P(A_{k-1} | A_{k-2} \cap \ldots \cap A_1) \quad \times \quad \ldots \quad \times \quad P(A_2 | A_1) \times P(A_1).$$

Przykład (drzewa prawdopodobieństw warunkowych)

Pewna uczelnia bada wyniki nauczania.

A_1 = {po I sem. ocena > 3 z przedm. ogólnych}

A_2 = {po II sem. ocena > 4 z przedm. specjalist.}

B = { „dobre” ukończenie studiów }

C = { „słabsze” ukończenie studiów }

D = { nieukończenie studiów }

$P(A_1)$ = 0,76, $P(A_2' | A_1)$= 0,80, $P(C | A_1 \cap A_2')$ = 0,69.

$P(A_1 \cap A_2' \cap C)$ = 0,76 x 0,80 x 0,69 = 0,42.

Uwaga:

Niech $A \subset S, P(A) > 0$. Prawdopodobieństwo warunkowe pod warunkiem zdarzenia A spełnia aksjomaty prawdopodobieństwa z definicji 5, jeli zamienimy S na A:

- $0 \leq P(B|A) \leq 1$, dla każdego zdarzenia B,
- $P(A|A) = 1, \qquad P(\emptyset|A) = 0$,
- Jeśli zdarzenia $B_1, B_2, B_3, \ldots$ wykluczają się wzajemnie , to

$$\boxed{P(B_1 \cup B_2 \cup \ldots|A) = P(B_1|A) + P(B_2|A) + \cdots}$$

Zdarzenia niezależne

Prawdopodobieństwo warunkowe pozwala określić niezależność zdarzeń. Zdarzenia A i B, o dodatnich prawdopodobieństwach, nazwiemy **niezależnymi**, jeśli **informacja o zajściu jednego z nich nie wpływa na prawdopodobieństwo zajścia drugiego**:

$$P(B|A) = P(B) \qquad \text{oraz} \qquad P(A|B) = P(A)$$

Zdarzenia niezależne

Definicja 7

Zdarzenia A i B nazywamy **niezależnymi,** jeśli

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

Zdarzenia $\boxed{A_1, A_2, \ldots, A_k}$, $k \geq 2$, nazywamy **niezależnymi**,

jeśli dla każdego m, $2 \leq m \leq k$, dla dowolnych różnych

zdarzeń $A_{i_1}, A_{i_2}, \ldots, A_{i_m}$ z rodziny $\{A_1, A_2, \ldots, A_k\}$:

$$\boxed{P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_m}) = P(A_{i_1}) \times \ldots \times P(A_{i_m})}$$

Zdarzenia niezależne parami

Określenie

Zdarzenia $A_1, A_2, \ldots, A_k$, $k \geq 2$, nazywamy niezależnymi parami, jeśli każde dwa zdarzenia spośród nich są niezależne.

Uwaga: Z niezależności parami nie wynika niezależność rodziny zdarzeń $A_1, A_2, \ldots, A_k$.

Natomiast **niezależność rodziny** zdarzeń implikuje **niezależność parami.**

Zdarzenia niezależne parami

Przykład Urna zawiera 4 kule – zieloną, niebieską, czerwoną i kulę zielono-niebiesko-czerwoną. Niech

A_1 = { w losowo wybranej kuli jest kolor zielony },

A_2 = { w losowo wybranej kuli jest kolor niebieski },

A_3 = { w losowo wybranej kuli jest kolor czerwony },

B = { losowo wybrana kula jest trójkolorowa }.

Pokaż, że zdarzenia A_1, A_2, A_3 **nie są niezależne**, ale są **parami niezależne**.

Wsk. $A_1 \cap A_2 = A_2 \cap A_3 = A_1 \cap A_3 = B$.

Przykład Niech $S = \{s_1, s_2, s_3, s_4\}$ oraz zdarzenia elementarne są jednakowo prawdopodobne.

- $A_1 = \{s_1, s_2\}, \ A_2 = \{s_1, s_3\}, \ A_3 = \{s_1, s_4\}$

- A_1, A_2, A_3 są niezależne parami, ale

$P(A_1 \cap A_2 \cap A_3)$ = 0,25 $\neq$ $P(A_1)P(A_2)P(A_3)$ = 0,125.

- $B_1 = \{s_1\}, \ B_2 = \{s_1, s_2\}, \ B_3 = \emptyset.$

- B_1, B_2, B_3 nie są parami niezależne, ale

$P(B_1 \cap B_2 \cap B_3)$ = $P(B_1)P(B_2)P(B_3)$.

Niezależność układu zdarzeń

<u>Przykład</u> Układ czterech przekaźników połączony jest w taki sposób, że: dwa pierwsze – szeregowo, połączone są szeregowo z dwoma pozostałymi połączonymi równolegle. Przekaźniki pracują niezależnie, prawdopodobieństwo awarii każdego z nich wynosi 0,1. Oblicz niezawodność układu przekaźników, tzn. prawdopodobieństwo poprawnej pracy.

- A_i = { przekaźnik i pracuje poprawnie }, i = 1,2,3,4.
- $D =$ { układ pracuje poprawnie } =

$$A_1 \cap A_2 \cap (A_3 \cup A_4)$$

$$D = (A_1 \cap A_2 \cap A_3) \cup (A_1 \cap A_2 \cap A_4).$$

Niezależność układu zdarzeń

$$D = (A_1 \cap A_2 \cap A_3) \cup (A_1 \cap A_2 \cap A_4).$$

Z twierdzenia 3 (prawdopodobieństwo sumy zdarzeń) oraz definicji 7 (niezależność zdarzeń):

Niezawodność =

$$P(D) = P(A_1 \cap A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_4) +$$

$$- P(A_1 \cap A_2 \cap A_3 \cap A_4) = 0{,}9^3 + 0{,}9^3 - 0{,}9^4 =$$

0,8019.

Prawdopodobieństwo sumy zdarzeń

Twierdzenie 5 Niech $A \subset S$ i $B \subset S$. Wówczas

$$P(A \bigcup B) = P(A) + P(B) - P(A \bigcap B),$$

$$P(A \bigcup B) = P(A) + P(B), \text{ jeśli } A \bigcap B = \emptyset.$$

Dowód

$$A \bigcup B = (A \bigcap B) \bigcup (A - B) \bigcup (B - A)$$

Prawdopodobieństwo sumy niezależnych zdarzeń

Wniosek 1 Niech A, B będą zdarzeniami niezależnymi. Wówczas

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

Wniosek 2 Niech A, B, C będą zdarzeniami. Wówczas

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

Zupełny układ zdarzeń

Definicja 8

Zdarzenia $B_1, B_2, ..., B_k$ tworzą **podział** przestrzeni zdarzeń elementarnych S (układ **zupełny** zdarzeń), jeśli $B_i \cap B_j = \emptyset$ dla $i \neq j$, oraz

$$B_1 \cup B_2 \cup ... \cup B_k = S.$$

Prawdopodobieństwo całkowite

<u>Twierdzenie 6</u> (o prawdopodobieństwie całkowitym).

Jeśli $\boxed{B_1, B_2, \ldots, B_k}$ tworzą **układ zupełny** zdarzeń oraz $P(B_i) \neq 0$, $i = 1,2,\ldots k$, to dla każdego zdarzenia A:

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

<u>D.</u> $P(A) = P(A \cap S) = P(A \cap (B_1 \cup B_2 \cup \ldots \cup B_k)) =$

$$P((A \cap B_1) \cup (A \cap B_2) \cup \ldots \cup (A \cap B_k)) =$$

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

Twierdzenie Bayesa

Twierdzenie 7 (**reguła Bayesa'a**). Jeśli zdarzenia $B_1, B_2, \ldots, B_k$ tworzą podział przestrzeni S oraz $P(B_i) > 0$, $i = 1,2,\ldots k$, to dla $A \in S$, takiego że $P(A) > 0$,

$$\boxed{P(B_m|A) = \frac{P(A|B_m)P(B_m)}{P(A)} = \frac{P(A|B_m)P(B_m)}{\sum_{j=1}^{k} P(A|B_j)P(B_j)}},$$

gdzie B_m jest dowolnym ustalonym zdarzeniem spośród zdarzeń $B_1, B_2, \ldots, B_k$, $1 \le m \le k$.

Reguła Bayesa

Przykład W konferencji naukowej bierze udział 30 % matematyków i 70 % informatyków. Wśród matematyków jest 50 % kobiet a wśród informatyków zaledwie 10 % stanowią kobiety. Wybrana losowo osoba jest (a) kobietą , (b) mężczyzną. Jakie jest prawdopodobieństwo, że wybrana osoba jest matematykiem ?

Określamy zdarzenia:

$A =$ {wybrana losowo osoba jest kobietą},

$B_1 =$ { wybrana losowo osoba jest matematykiem},

$B_2 =${ wybrana losowo osoba jest informatykiem},

$$P(B_1) = 0{,}3, \quad P(B_2) = 0{,}7, \quad P(A|B_1) = 0{,}5, \quad P(A|B_2) = 0{,}1,$$

$$P(A'|B_1) = 1 - P(A|B_1) = 1 - 0{,}5 = 0{,}5,$$

$$P(A'|B_2) = 1 - P(A|B_2) = 1 - 0{,}1 = 0{,}9.$$

Reguła Bayesa

$$P(B_1)=0{,}3,\quad P(B_2)=0{,}7,\; P(A|B_1)=0{,}5,\; P(A|B_2)=0{,}1,$$
$$P(A'|B_1)=1-P(A|B_1)=1-0{,}5=0{,}5,$$
$$P(A'|B_2)=1-P(A|B_2)=1-0{,}1=0{,}9.$$

(a)

$$P(B_1|A)=\frac{P(A|B_1)P(B_1)}{P(A|B_1)P(B_1)+P(A|B_2)P(B_2)}=$$

$$=\frac{0{,}5\times 0{,}3}{0{,}5\times 0{,}3+0{,}1\times 0{,}7}\approx 0{,}68.$$

Reguła Baeysa

Interpretacja:

Wśród matematyków jest dużo kobiet, zatem prawdop., że wybrana osoba jest matematykiem zwiększyło się, jeśli wiemy, że ta osoba jest kobietą.

$$P(B_1) = 0{,}3 \qquad P(B_1 | A) = 0{,}68$$

$$P(B_1) = 0{,}3 \qquad P(B_1 | A') = 0{,}19$$

Reguła Bayesa

$$P(B_1)=0{,}3,\quad P(B_2)=0{,}7,\qquad P(A|B_1)=0{,}5$$
$$P(A|B_2)=0{,}1,$$

$$P(A'|B_1)=1-P(A|B_1)=1-0{,}5=0{,}5,$$
$$P(A'|B_2)=1-P(A|B_2)=1-0{,}1=0{,}9.$$

(b)

$$P(B_1|A')=\frac{P(A'|B_1)P(B_1)}{P(A'|B_1)P(B_1)+P(A'|B_2)P(B_2)}=$$

$$=\frac{0{,}5\times 0{,}3}{0{,}5\times 0{,}3+0{,}9\times 0{,}7}\approx 0{,}19.$$

Dziękuję za uwagę