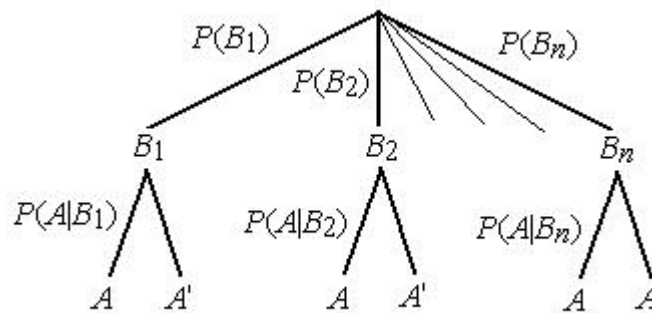


## PRAWDOPODOBIEŃSTWO CAŁKOWITE

JEŻELI ZDARZENIA  $B_1, B_2, \dots, B_n$  WYKLUCZAJĄ SIĘ PARAMI I MAJĄ PRAWDOPODOBIEŃSTWA DODATNIE, TO DLA KAŻDEGO ZDARZENIA  $A$  ZAWARTEGO W SUMIE ZDARZEŃ  $B_1 \cup B_2 \cup \dots \cup B_n$ :

$$P(A) = \sum_{i=1}^n P(A|B_i)P(B_i) = P(A|B_1)P(B_1) + \dots + P(A|B_n)P(B_n)$$

POWYŻSZY WZÓR NAZYWAMY WZOREM NA PRAWDOPODOBIEŃSTWO CAŁKOWITE I POZWALA NAM NA OBLICZANIE PRAWDOPODOBIEŃSTW WIELU ZDARZEŃ NIE TYLKO W DOŚWIADCZENIACH DWUETAPOWYCH. W DOŚWIADCZENIACH O WIĘKSZEJ LICZBIE ETAPÓW STOSUJEMY TEN WZÓR WIELOKROTNIE. ZDARZENIA  $B_i$  NAZYWAMY CZĘSTO HIPOTEZAMI. TWIERDZENIE O PRAWDOPODOBIEŃSTWIE CAŁKOWITYM MOŻNA ZILUSTROWAĆ ZA POMOCĄ TZW. DRZEWA STOCHASTYCZNEGO.



WZÓR NA PRAWDOPODOBIEŃSTWO CAŁKOWITE TO SUMA ILOCZYNÓW PO WSZYSTKICH DROGACH, KTÓRE KOŃCZĄ SIĘ W  $A$ . DRZEWO STOCHASTYCZNE ZACZYNA SIĘ POCZĄTKIEM, W WĘZŁACH DRZEWA UMIESZCZAMY WYNIKI KOLEJNYCH ETAPÓW DOŚWIADCZENIA. WĘZŁY ŁĄCZYMY KRAWĘDZIAMI. OBOK KAŻDEJ KRAWĘDZI DOPISUJEMY PRAWDOPODOBIEŃSTWO OTRZYMANIA WYNIKU DANEGO ETAPU. SUMA PRAWDOPODOBIEŃSTW PRZYPORZĄDKOWANYCH KRAWĘDZIOM WYCHODZĄCYM Z JEDNEGO WĘZŁA JEST RÓWNA 1. DOWOLNY CIĄG KRAWĘDZI ŁĄCZĄCY POCZĄTEK DRZEWA Z JEDNYM Z KOŃCOWYCH WĘZŁÓW NAZYWAMY GAŁĘZIĄ DRZEWA. KAŻDEJ GAŁĘZI ODPOWIADA JEDEN WYNIK DOŚWIADCZENIA WIELOETAPOWEGO. PRAWDOPODOBIEŃSTWO WYNIKU ODPOWIADAJĄCEGO DANEJ GAŁĘZI DRZEWA RÓWNA SIĘ ILOCZYNOWI PRAWDOPODOBIEŃSTW PRZYPISANYCH KRAWĘDZIOM, Z KTÓRYCH JEST ZŁOŻONA GAŁĘŻ (JEST TO TZW. REGUŁA MNOŻENIA DLA DRZEWA).

## TWIERDZENIE BAYESA

**TWIERDZENIE BAYESA** (OD NAZWISKA THOMASA BAYESA) TO TWIERDZENIE TEORII PRAWDOPODOBIENSTWA, WIAZĄCE PRAWDOPODOBIENSTWA WARUNKOWE ZDARZEŃ  $A|B$  ORAZ  $B|A$ . NA PRZYKŁAD, JEŚLI  $A$  JEST ZDARZENIEM "U PACJENTA WYSTĘPUJE WYSOKA GORĄCZKA", A  $B$  JEST ZDARZENIEM "PACJENT MA GRYPĘ", TWIERDZENIE BAYESA POZWALA PRZELICZYĆ ZNANY ODSETEK GORĄCZKUJĄCYCH WŚRÓD CHORYCH NA GRYPĘ  $P(A|B)$  I ZNANE ODSETKI GORĄCZKUJĄCYCH  $P(A)$  I CHORYCH NA GRYPĘ  $P(B)$  W CAŁEJ POPULACJI, NA PRAWDOPODOBIENSTWO, ŻE KTOŚ JEST CHORY NA GRYPĘ, GDY WIEMY, ŻE MA WYSOKĄ GORĄCZKĘ  $P(B|A)$ . TWIERDZENIE STANOWI PODSTAWĘ TEORETYCZNĄ SIECI BAYESOWSKICH, STOSOWANYCH W DATA MINING.

## WZÓR BAYESA

JEŻELI ZDARZENIA  $T_1, T_2, \dots, T_n$  WYKLUCZAJĄ SIĘ PARAMI I MAJĄ PRAWDOPODOBIENSTWA DODATNIE, TO DLA KAŻDEGO ZDARZENIA  $A$  ZAWARTEGO W SUMIE ZDARZEŃ  $T_1 \cup T_2 \cup \dots \cup T_n$ :

$$P(T_i|X) = \frac{P(T_i) \cdot P(X|T_i)}{P(X)}$$

$$P(T_i|X) = \frac{P(T_i) \cdot P(X|T_i)}{P(T_1) \cdot P(X|T_1) + P(T_2) \cdot P(X|T_2) + \dots + P(T_n) \cdot P(X|T_n)}$$

POWYŻSZY WZÓR NAZYWAMY WZOREM BAYESA. TWIERDZENIE BAYESA STOSUJEMY GŁÓWNIIE WTEDY, GDY ZNAMY WYNIK DOŚWIADCZENIA I PYTAMY O JEGO PRZEBIEG.

## DOWÓD

SKORO

$$P(X|T) = \frac{P(X \cap T)}{P(T)}$$

$$P(X \cap T) = P(T) \cdot P(X|T)$$

A

$$P(X \cap T) = P(T \cap X)$$

TO

$$P(X \cap T) = P(T) \cdot P(X|T) = P(X) \cdot P(T|X)$$

WIĘC

$$P(T) \cdot P(X|T) = P(X) \cdot P(T|X)$$

$$P(T|X) = \frac{P(T) \cdot P(X|T)}{P(X)}$$

## ZADANIA

1. ŻARÓWKI PEWNEJ MARKI SĄ PRODUKOWANE W DWU FABRYKACH  $X$  I  $Y$ . ŻARÓWKI Z FABRYKI  $X$  DZIAŁAJĄ DŁUŻEJ NIŻ 5000 GODZIN W 99% PROCENTACH PRZYPADKÓW, ŻARÓWKI Z FABRYKI  $Y$  TYLKO W 95% PRZYPADKÓW. FABRYKA  $X$  DOSTARCZA NA RYNEK 60% ŻARÓWEK TEJ MARKI. JAKIE JEST PRAWDOPODOBIENSTWO, ŻE ZAKUPIONA LOSOWO ŻARÓWKA BĘDZIE SPRAWNA DŁUŻEJ NIŻ 5000 GODZIN?
2. PRZEZ PEWNĄ MIEJSCOWOŚĆ PRZEJEŹDZA 4 RAZY WIĘCEJ SAMOCHODÓW OSOBOWYCH NIŻ CIĘŻAROWYCH. WIADOMO, ŻE JEDEN NA 75 SAMOCHODÓW OSOBOWYCH I JEDEN NA 25 SAMOCHODÓW CIĘŻAROWYCH TANKUJE PALIWO W TEJ MIEJSCOWOŚCI. OBLICZ PRAWDOPODOBIENSTWO ZE PRZEJEŹDZAJĄCY SAMOCHÓD ZATANKUJE PALIWO.
3. W PIERWSZEJ URNIE SĄ 3 KULE BIAŁE I 2 CZARNE, A W DRUGIEJ SĄ 4 CZARNE I 1 BIAŁA. RZUCAMY KOSTKĄ. JEŻELI WYPADNĄ MNIJ NIŻ 3 OCZKA, TO LOSUJEMY KULĘ Z PIERWSZEJ URNY, W PRZECIWNYM WYPADKU LOSUJEMY KULĘ Z DRUGIEJ URNY. JAKIE JEST PRAWDOPODOBIENSTWO WYLOSOWANIA KULI BIAŁEJ?
4. STACJA METEOROLOGICZNA ZBIERA INFORMACJE DOTYCZĄCE SZANSY WYSTĄPIENIA DNIA DESZCZOWEGO. WEDLE UZYSKANYCH DANYCH, JEŚLI DANEGO DNIA PADA, TO SZANSA, ŻE NASTĘPNY DZIEŃ TAKŻE BĘDZIE DESZCZOWY WYNOŚI 50%, JEŚLI ZAŚ DANY DZIEŃ JEST SŁONECZNY, TO KOLEJNY DZIEŃ JEST DESZCZOWY W 30% PRZYPADKÓW. SYNOPTYCY PRZEPOWIADAJĄ, NA 80%, NAJBLIŻSZY PONIEDZIAŁEK JAKO DZIEŃ DESZCZOWY. JAKIE JEST PRAWDOPODOBIENSTWO, ŻE ŚRODA BĘDZIE DESZCZOWA?
5. PEWNA FIRMA WIEDZĄC, ŻE 0,5% Z JEJ PRACOWNIKÓW TO NARKOMANI, POSTANOWIŁA PRZEPROWADZIĆ TEST ANTY-NARKOTYKOWY. JAKIE JEST PRAWDOPODOBIENSTWO PRAWIDŁOWEGO WYKRYCIA NARKOMANA, JEŻELI WIADOMO, ŻE TEST MA 99% SKUTECZNOŚCI (W 1% PRZYPADKÓW TEST WYPADA POZYTYWNIEM U OSOBY NIE ZAŻYWAJĄCEJ, A NEGATYWNIEM U NARKOMANA)?
6. W URNACH O NUMERACH 1-9 ZNAJDUJĄ SIĘ 4 KULE BIAŁE I 6 CZARNYCH, NATOMIAST W URNIE NR 10 JEST 10 KUL BIAŁYCH I 10 CZARNYCH. Z LOSOWO WYBRANEJ URNY WYLOSOWANO 2 KULE I OKAZAŁY SIĘ BIAŁE. JAKIE JEST PRAWDOPODOBIENSTWO, ŻE ZOSTAŁY WYLOSOWANE Z 10 URNY.
7. W ZBIORZE 100 MONET JEDNA MA PO OBU STRONACH ORŁY, A POZOSTAŁE SĄ PRAWIDŁOWE. RZUCONO 10-KROTNIEM JEDNĄ LOSOWO WYBRANĄ MONETĄ I OTRZYMANO 10 ORŁÓW. OBLICZYĆ PRAWDOPODOBIENSTWO TEGO, ŻE BYŁA TO MONETA Z DWOMA ORŁAMI.
8. TEST NA RZADKĄ CHOROBE, KTÓRĄ DOTKNIĘTA JEST ŚREDNIO JEDNA OSOBA NA 1000, DAJE TAK ZWANĄ FAŁSZYWĄ POZYTYWĄ ODPOWIEŹ U 5% ZDROWYCH (U CHOREGO ZAWSZE DAJE ODPOWIEŹ POZYTYWĄ). JAKA JEST SZANSA, ŻE OSOBA, U KTÓREJ TEST DAŁ ODPOWIEŹ POZYTYWĄ, JEST FAKTYCZNIE CHORA?
9. 5 MĘŻCZYŹN NA 100 ORAZ 25 NA 1000 KOBIEŹ JEST DALTONISTAMI. LOSOWO WYBRANO OSOBĘ I OKAZAŁO SIĘ, ŻE NIE ODRÓŻNIA ONA KOLORÓW. JAKIE JEST PRAWDOPODOBIENSTWO, ŻE BYŁ TO MĘŻCZYŻNA? (ILOŚĆ KOBIEŹ I MĘŻCZYŻN JEST RÓWNA)
10. PEWNE URZĄDZENIA SĄ SPROWADZANE OD 3 DOSTAWCÓW: A, B, C W ILOŚCIACH ODPOWIEDNIO: 50%, 20%, 30%. WADLIWOŚĆ URZĄDZEŹ: A – 1%, B – 2%, C – 3%. WYBRANE URZĄDZENIE JEST WADLIWE. JAKIE JEST PRAWDOPODOBIENSTWO, ŻE POCHODZI ONO OD DOSTAWCY A?

11. PRAWDOPODOBIENSTWO WYBUCHU GAZU W KOPALNI WYNOŚI 0,025. W KOPALNI JEST ZAINSTALOWANY SYSTEM ALARMOWY, KTÓRY W SYTUACJI ZAGROŻENIA ZAWODZI W 12% PRZYPADKÓW. FAŁSZYWY ALARM ZDARZA SIĘ W 15% PRZYPADKÓW. ZAŁÓŻMY, ŻE WŁĄCZA SIĘ ALARM. JAKIE JEST PRAWDOPODOBIENSTWO, ŻE JEST ON FAŁSZYWY?
12. W OKRĘGU WYBORCZYM BADANIE PREFERENCJI POLITYCZNYCH WYBRANYCH WYBORCÓW UJAWNIŁO, ŻE 70% POPIERA KANDYDATA A, ZAŚ 30% KANDYDATA B. JEDNAKŻE BADANIE TO, JEDNOCZEŚNIE WYKAZAŁO, ŻE WŚRÓD TYCH, KTÓRZY POPIERAJĄ KANDYDATA A, 32% PÓJDZIE GŁOSOWAĆ, A WŚRÓD TYCH POPIERAJĄCYCH KANDYDATA B, 90%. JAKIE JEST PRAWDOPODOBIENSTWO, ŻE WYBRANA OSOBA ODDAJĄCA GŁOS, ODDA GO NA KANDYDATA A?
13. WIEDZĄC, ŻE UŁAMEK  $p_k = 0,49$  POPULACJI TO KOBIETY, WŚRÓD KTÓRYCH UŁAMEK  $f = 0.005$  STANOWIĄ DALTONISTKI ORAZ WIEDZĄC, ŻE UŁAMEK  $p_d = 0.05$  CAŁEJ POPULACJI CIERPI NA DALTONIZM, OBLICZ JAKIE JEST PRAWDOPODOBIENSTWO, ŻE SPOŚRÓD LOSOWEGO DALTONISTY, WYLOSUJEMY KOBIETĘ.
14. WŚRÓD WSZYSTKICH CIĄŻ, JEDNA NA 300 PROWADZI DO BLIŹNIĄT MONOZYGOTYCZNYCH (JEDNOJAJOWE), KIEDY TO DZIECI SĄ TEJ SAMEJ PŁCI I SĄ PODOBNE DO SIEBIE, JAK DWIE KROPLE WODY, NATOMIAST Z JEDNEJ CIĄŻY NA 125 RODZĄ SIĘ BLIŹNIĘTA DIZYGOTYCZNE (DWUJAJOWE), KTÓRE MOGĄ BYĆ RÓŻNEJ PŁCI I NICZYM SIĘ NIE RÓŻNIĄ OD DZIECI URODZONYCH W DWÓCH KOLEJNYCH CIĄŻACH. ELVIS PRESLEY MIAŁ BRATA BLIŹNIAKA, KTÓRY ZMARŁ PODCZAS PORODU. WYZNACZ PRAWDOPODOBIENSTWO, ŻE PRESLEY BYŁ BLIŹNIAKIEM MONOZYGOTYCZNYM
15. (PARADOKS MONTY HALL'A) GRACZ POPULARNEGO TELETURNIJU STOI PRZED WYBOREM JEDNEJ Z TRZECH BRAMEK. ZA JEDNĄ Z BRAMEK STOI NAGRODA – SAMOCHÓD, A ZA POZOSTAŁYMI DWOMA NIE MA NIC. GRACZ WYBIERA JEDNĄ Z BRAMEK. NASTĘPNIE PROWADZĄCY (ZYGMENT HAJZER ☺) ODSŁANIA JEDNĄ Z PUSTYCH BRAMEK (INNĄ NIŻ TA WYBRANA PRZEZ GRACZA) I DAJE GRACZOWI MOŻLIWOŚĆ ZMIANY SWOJEGO WYBORU. CZY GRACZ POWINIEN ZMIEŃCIĆ SWÓJ WYBÓR CZY NIE?
16. (PROBLEM TRZECH WIĘŹNIÓW) W CELI ZAMKNIĘTYCH JEST 3 WIĘŹNIÓW: A, B, C. JEDEN Z NICH MA BYĆ STRACONY – DECYZJA JUŻ ZAPADŁA, ALE ŻADEN Z WIĘŹNIÓW NIE WIE, KTO MA ZOSTAĆ STRACONY. W PEWNYM MOMENCIE PRZYCHODZI STRAŻNIK I WYPUSZCZA WIĘŹNIA B. CZY PRAWDOPODOBIENSTWO, ŻE STRACĄ WIĘŹNIA A JEST TERAZ WIĘKSZE, MNIEJSZE, CZY TAKIE SAMO JAK NA POCZĄTKU?
17. (PARADOKS BERTRAND'A) NA OKRĘGU O PROMIENIU 1SKONSTRUOWANO LOSOWO CIĘCIWĘ. JAKIE JEST PRAWDOPODOBIENSTWO, ŻE CIĘCIWA BĘDZIE DŁUŻSZA, NIŻ BOK TRÓJKĄTA RÓWNOBOCZNEGO WPISANEGO W TEN OKRĄG?
18. (PARADOKS PUDEŁKA BERTRAND'A) W TRZECH PUDEŁKACH ZNAJDUJĄ SIĘ PO DWIE MONETY. W PIERWSZYM PUDEŁKU DWIE ŻŁOTE, W DRUGIM DWIE SREBRNE, A W TRZECIM JEDNA ŻŁOTA, DRUGA SREBRNA. WYBIERAMY LOSOWO PUDEŁKO I WYCIĄGAMY Z NIEGO JEDNĄ MONETĘ, KTÓRA OKAZUJE SIĘ ŻŁOTA. JAKIE JEST PRAWDOPODOBIENSTWO, ŻE DRUGA MONETA TEŻ JEST ŻŁOTA?