AGH
AKADEMIA GÓRNICZO-HUTNICZA
IM. STANISŁAWA STASZICA W KRAKOWIE

Rachunek prawdopodobieństwa

Statystyka

Dr inż. Janusz Majewski
Katedra Informatyki

Literatura

- Prezentacja wykorzystuje fragmenty książki: Amir D. Aczel „Statystyka w zarządzaniu”, PWN, 2007

Przestrzeń probabilistyczna, zdarzenia

Eksperymentem nazywamy proces, który prowadzi do jednego z możliwych *wyników*. Nazywamy je wynikami obserwacji lub wynikami pomiaru.

Przestrzeń prób jest zbiorem wszystkich możliwych wyników eksperymentu. Jest ona zbiorem uniwersalnym X związanym z danym eksperymentem.

Zdarzeniami są podzbiory w przestrzeni próby. Są to zbiory pewnych zdarzeń elementarnych. Mówimy, że zaszło dane zdarzenie, jeżeli w wyniku eksperymentu zaszło zdarzenie elementarne będące elementem podzbioru odpowiadającego temu zdarzeniu.

Przestrzeń probabilistyczna, zdarzenia

Zdarzenie *A*: wyciągnięcie asa

Kier	Karo	Trefl	Pik
H	*D*	*C*	*S*
A	*A*	*A*	*A*
K	*K*	*K*	*K*
D	*D*	*D*	*D*
W	*W*	*W*	*W*
10	10	10	10
9	9	9	9
8	8	8	8
7	7	7	7
6	6	6	6
5	5	5	5
4	4	4	4
3	3	3	3
2	2	2	2

Wyciągnięcie asa pik oznacza, że zaszło zdarzenie *A*

Rysunek 2.5. Przestrzeń prób w przypadku ciągnięcia kart z talii

Prawdopodobieństwo

Przy założeniu równej możliwości zdarzeń elementarnych, **prawdopodobieństwo zdarzenia** A jest względną miarą A w stosunku do miary przestrzeni prób X.

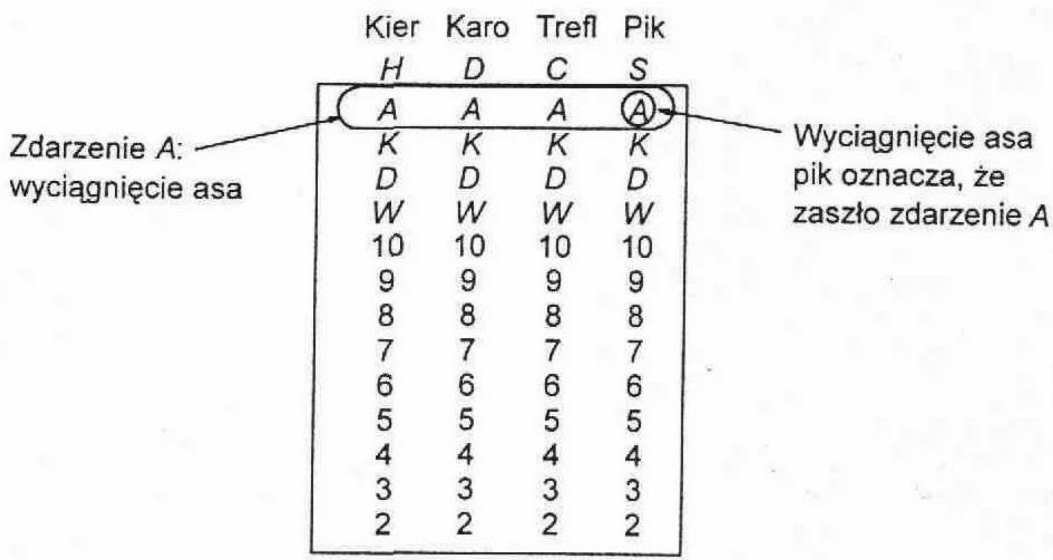
Ponieważ definicja ta zakłada uogólnioną zasadę równej możliwości, może ona nie być właściwa, gdy nie ma podstaw do przyjęcia tego założenia.

Oznaczając miarę zbioru A przez $|A|$, a miarę przestrzeni próby przez $|X|$, definicję prawdopodobieństwa możemy zapisać za pomocą wzoru[5].

Prawdopodobieństwo zdarzenia A:

$$P(A) = \frac{|A|}{|X|}. \tag{2.1}$$

Prawdopodobieństwo



Rysunek 2.5. Przestrzeń prób w przypadku ciągnięcia kart z talii

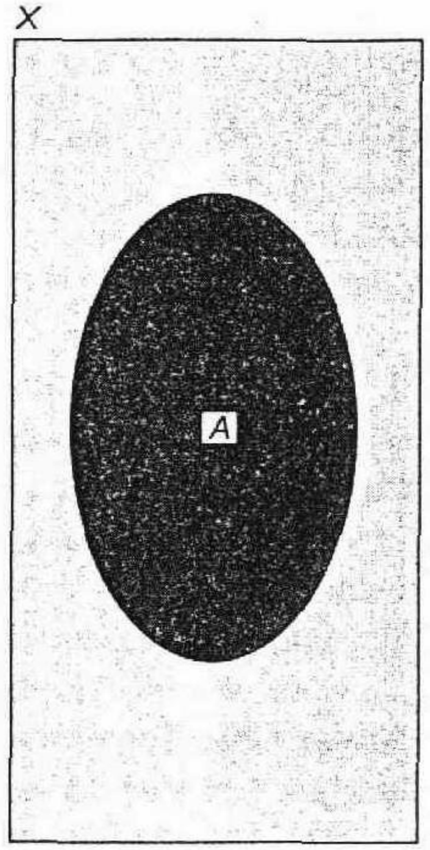
Dla skończonych przestrzeni prób równanie (2.1) może być uproszczone. Oznaczając przez $n(A)$ liczbę elementów zbioru A, a przez $n(X)$ liczbę elementów przestrzeni prób X, możemy napisać:

$$P(A) = \frac{n(A)}{n(X)}. \tag{2.2}$$

Prawdopodobieństwo wyciągnięcia asa: $P(A) = n(A)/n(X) = 4/52$.

Prawdopodobieństwo

Rysunek 2.6 ilustruje ideę prawdopodobieństwa jako stosunku miary zbioru odpowiadającego zdarzeniu do miary całej przestrzeni prób.

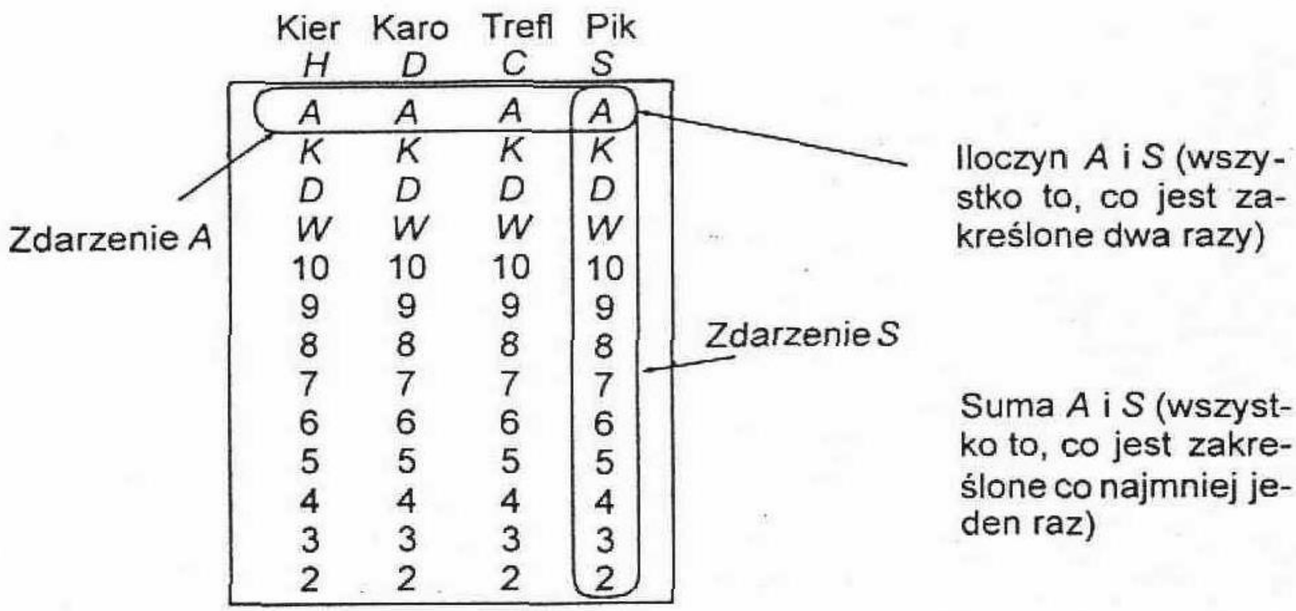


$$P(A) = \frac{|A|}{|X|} = \frac{\blacksquare}{\square + \blacksquare}$$

Rysunek 2.6. Prawdopodobieństwo jako stosunek dwóch miar

Przykład

Rozpatrzmy dwa zdarzenia, które mogą zajść, gdy pojedyncza karta zostanie wyciągnięta z dobrze potasowanej talii 52 kart. Niech A będzie zdarzeniem: wyciągnięcie asa, a S zdarzeniem: wyciągnięcie pika. Na rysunku 2.7 pokazano przestrzeń próby dla tego eksperymentu, zdarzenia A, S, ich przekrój i połączenie. Z równania (2.2) przez proste wyliczenie otrzymujemy: $P(A) = 4/52$, $P(S) = 13/52$, $P(A \cup S) = 16/52$, $P(A \cap S) = 1/52$.



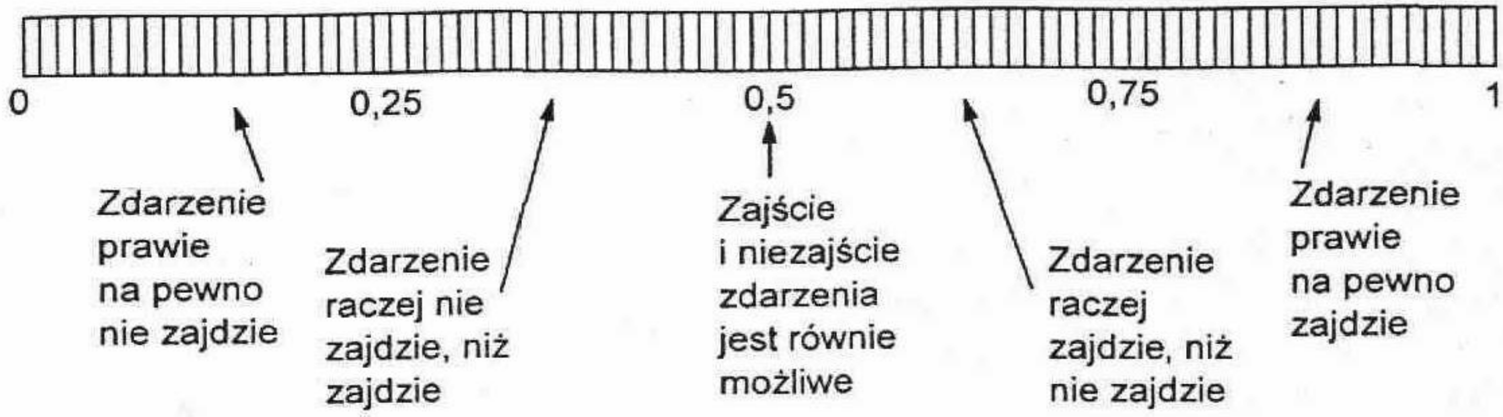
Rysunek 2.7. Zdarzenia A i S, ich połączenie (suma) i przekrój (iloczyn)

Reguły obliczania prawdopodobieństw

Prawdopodobieństwo jest miarą niepewności. Prawdopodobieństwo zdarzenia A jest liczbową miarą naszego przekonania, że zdarzenie to zajdzie.

Dla dowolnego zdarzenia *A*:

$$0 \leq P(A) \leq 1. \qquad (2.3)$$



Rysunek 2.9. Interpretacja prawdopodobieństwa

Reguły obliczania prawdopodobieństw

$$P(\overline{A}) = 1 - P(A). \tag{2.4}$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B). \tag{2.5}$$

Dla wykluczających się wzajemnie zdarzeń A i B:

$$P(A \cap B) = 0. \tag{2.6}$$

Z tego faktu wynika szczególny przypadek reguły sumowania wzajemnie wykluczających się zdarzeń. Ponieważ w tym przypadku $P(A \cap B) = 0$, więc:

$$P(A \cup B) = P(A) + P(B). \tag{2.7}$$

Reguły de Morgana

$$(\overline{A \cup B}) = \overline{A} \cap \overline{B}. \tag{2.8}$$

Ponieważ prawdopodobieństwo dopełnienia zdarzenia jest równe 1 minus prawdopodobieństwo tego zdarzenia, reguła de Morgana (równanie (2.8)) w zastosowaniu do prawdopodobieństwa zdarzeń przyjmuje postać:

$$P(A \cup B) = 1 - P(\overline{A} \cap \overline{B}). \tag{2.9}$$

Prawdopodobieństwo warunkowe

Warunkowe prawdopodobieństwo zdarzenia A przy zajściu zdarzenia B:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \tag{2.11}$$

o ile $P(B) \neq 0$.

$$P(A \cap B) = P(A|B)P(B) \tag{2.12}$$

oraz

$$P(A \cap B) = P(B|A)P(A). \tag{2.13}$$

Niezależność zdarzeń

Warunki niezależności zdarzeń A i B:

$$P(A|B) = P(A), \quad (2.14)$$

$$P(B|A) = P(B) \quad (2.15)$$

i najwygodniejszy w zastosowaniach:

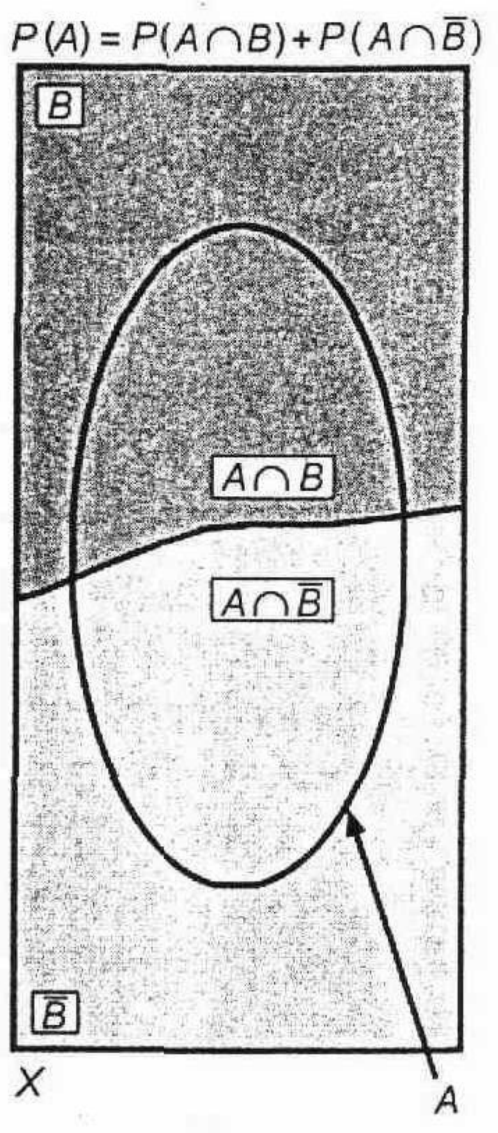
$$P(A \cap B) = P(A) \cdot P(B). \quad (2.16)$$

Reguła de Morgana w przypadku zdarzeń niezależnych:

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = 1 - P(\overline{A}_1 \cap \overline{A}_2 \ldots \cap \overline{A}_n) = 1 - P(\overline{A}_1)P(\overline{A}_2) \ldots P(\overline{A}_n). \quad (2.17)$$

Losowe pobieranie próby z wielkiej populacji implikuje niezależność wyników losowań.

Prawdopodobieństwo całkowite



Twierdzenie o prawdopodobieństwie całkowitym

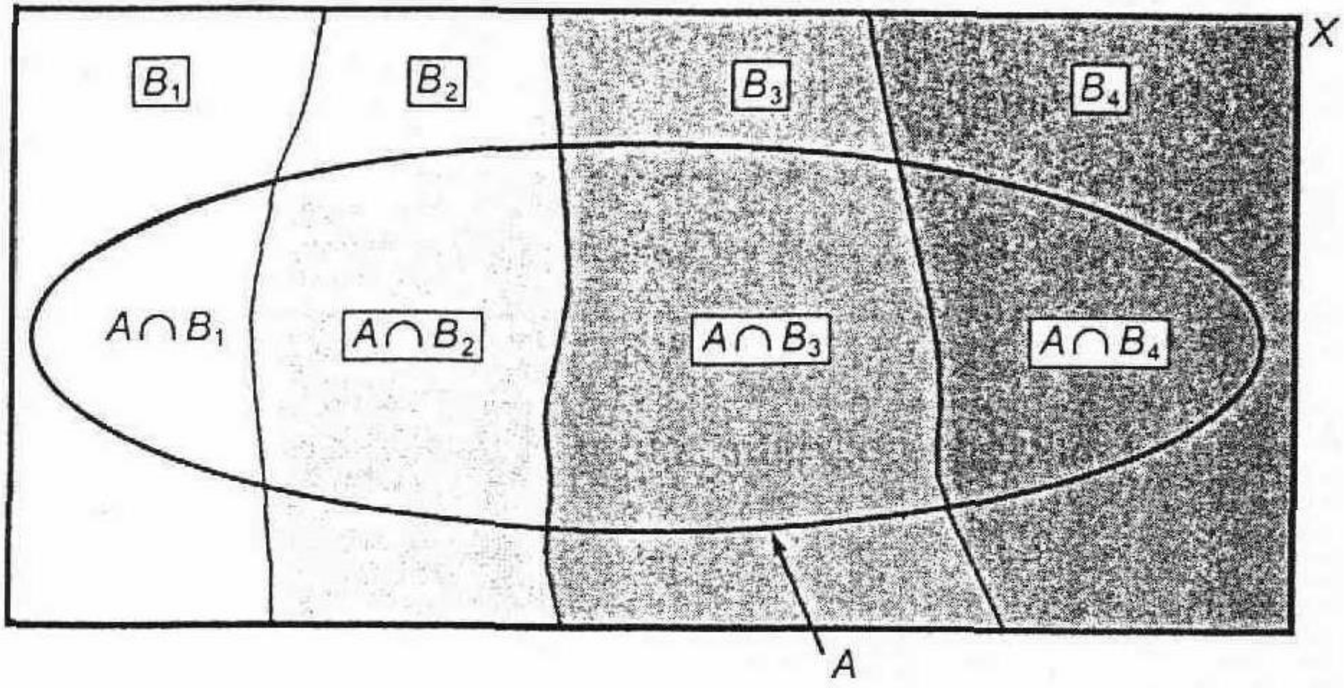
Twierdzenie o prawdopodobieństwie całkowitym:

$$P(A) = P(A \cap B) + P(A \cap \overline{B}). \tag{2.20}$$

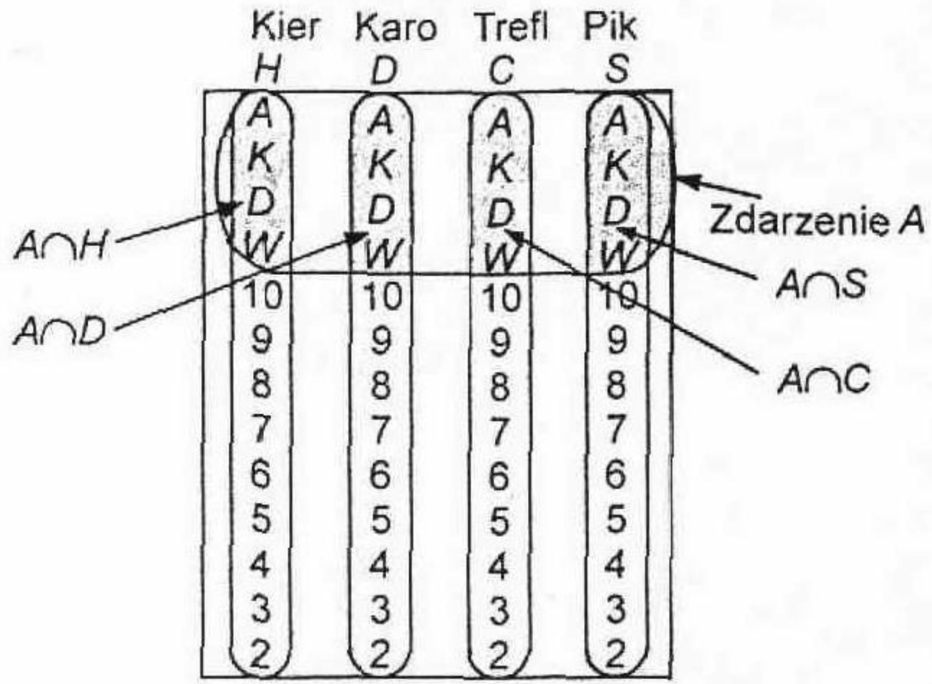
Prawdopodobieństwo całkowite

Twierdzenie o prawdopodobieństwie całkowitym można uogólnić na bardziej złożone sytuacje, w których przestrzeń prób jest podzielona na więcej niż dwa podzbiory (człony podziału). Jeżeli podzielimy ją na n podzbiorów $B_1, \ldots, B_n$, to twierdzenie o prawdopodobieństwie całkowitym można zapisać równaniem:

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



Prawdopodobieństwo całkowite



Rysunek 2.15. Całkowite prawdopodobieństwo wyciągnięcia figury z talii kart jako suma prawdopodobieństw wyciągnięcia karty z przekrojów zbioru figur i zbioru kolorów

Prawdopodobieństwo całkowite

Użyteczne wzory związane z prawdopodobieństwem całkowitym:

Przypadek dwuczłonowego podziału zbioru B:

$$P(A) = P(A \mid B)P(B) + P(A \mid \overline{B})P(\overline{B}). \tag{2.22}$$

Przypadek n-członowego podziału zbioru B na zbiory $B_1, \ldots, B_n$:

$$P(A) = \sum_{i=1}^{n} P(A \mid B_i)P(B_i). \tag{2.23}$$

Twierdzenie Bayesa

Twierdzenie Bayesa:

$$P(B|A)=\frac{P(A|B)P(B)}{P(A|B)P(B)+P(A|\overline{B})P(\overline{B})}. \qquad (2.27)$$

Uogólnione twierdzenie Bayesa:

$$P(B_1|A)=\frac{P(A|B_1)P(B_1)}{\sum_{i=1}^{n}P(A|B_i)P(B_i)}. \qquad (2.28)$$

Twierdzenie Bayesa – przykład zastosowania

Rozważane są 4 zestawy objawów (1, 2, 3 i 4) oraz 3 choroby (A, B, C). Warto zwrócić uwagę na wartość prawdopodobieństwa a priori choroby C w porównaniu z prawdopodobieństwem a priori choroby B oraz na prawdopodobieństwa a posteriori dla wszystkich chorób przy zestawie objawów 4.

objawy	gorączka	bóle mięśniowo-stawowe	bóle głowy	zmiany skórne	kaszel	katar	zaburzenia kardiolog.	zaburzenia neurolog.
1	+	+	+	+	+	-	-	-
2	+	-	+	-	-	+	-	-
3	-	+	-	-	+	+	-	-
4	+	+	+	+	-	-	+	+

Choroby	
A	grypa
B	przeziębienie
C	borelioza

Twierdzenie Bayesa – przykład zastosowania

a priori	choroba	1	2	3	4	
0,29	A	0,6	0,2	0,1	0,1	1
0,625	B	0,2	0,25	0,45	0,1	1
0,085	C	0,1	0,075	0,075	0,75	1
	A	0,56585366	0,26288952	0,091591	0,1867955	
	B	0,40650407	0,7082153	0,8882748	0,4025765	
	C	0,02764228	0,02889518	0,0201342	0,410628	
1		1	1	1	1	

Twierdzenie Bayesa – przykład zastosowania

		Objawy				
a priori	choroba	1	2	3	4	
0,29	A	0,6	0,2	0,1	0,1	1
0,625	B	0,2	0,25	0,45	0,1	1
0,085	C	0,1	0,075	0,075	0,75	1
	A	0,56585366	0,26288952	0,091591	0,1867955	
	B	0,40650407	0,7082153	0,8882748	0,4025765	
	C	0,02764228	0,02889518	0,0201342	0,410628	
1		1	1	1	1	

		Objawy			
a priori	choroba	1	2	3	4
P(A)	A	P(1|A)	P(2|A)	P(3|A)	P(4|A)
P(B)	B	P(1|B)	P(2|B)	P(3|B)	P(4|B)
P(C)	C	P(1|C)	P(2|C)	P(3|C)	P(4|C)
	A	P(A|1)	P(A|2)	P(A|3)	P(A|4)
	B	P(B|1)	P(B|2)	P(B|3)	P(B|4)
	C	P(C|1)	P(C|2)	P(C|3)	P(C|4)