Wybrane treści z rachunku prawdopodobieństwa w kontekście medycznym

M.Zalewska

Podstawowe pojęcia

- **Doświadczenie losowe** obserwacja zjawiska, którego przebiegu nie umiemy w pełni przewidzieć. Możemy oceniać z jakim prawdopodobieństwem wystąpią różne wyniki doświadczenia.
- **Przestrzeń próbkowa** zbiór wszystkich możliwych wyników doświadczenia losowego (Ω– omega), jej elementy to punkty próbkowe.
- **Zdarzenia losowe** to wypowiedzi dotyczące wyniku doświadczenia. Zdarzenia to pewne podzbiory przestrzeni próbkowej.

Przestrzeń próbkowa to
zbiór wszystkich możliwych
wyników doświadczenia

$$\Omega$$

Elementy zbioru to
pojedyncze wyniki
doświadczenia lub
zdarzenia elementarne

$$\omega \in \Omega$$

(Należy, jest elementem)

Zdarzenie to podzbiór
składający się z niektórych
wyników

$$A \subset \Omega$$

(Zawarte, jest podzbiorem)

Przykład: rzut kostką do gry

- Przestrzeń próbkowa $\Omega=\{1,2,3,4,5,6\}$

- Zdarzenie losowe – „wypadła parzysta liczba oczek” utożsamiamy ze zbiorem {2,4,6}

  Ω jest zdarzeniem pewnym

  – „wyrzucenie liczby oczek <7”

  Ø zbiór pusty jest zdarzeniem niemożliwym

  – „wyrzucenie liczby oczek >6”

Przykład: pomiar temperatury pacjenta

- Przestrzeń próbkowa $\Omega = (0, +\infty)$ przedział liczbowy

- Przykładowe wyniki doświadczenia:
  36.6
  37.7
  39.4………..

- Przykładowe zdarzenie losowe:
  „Pacjent ma temperaturę w normie" – utożsamiamy z przedziałem (36.6, 36.8)
  „ Pacjent ma temperaturę powyżej 38 "– utożsamiamy z przedziałem (38, + ∞)

Ω jest zdarzeniem pewnym – „temperatura ciała jest >0"

Ø zbiór pusty jest zdarzeniem niemożliwym – „ temperatura ciała jest >50"

Działania na zdarzeniach:

$A' = \Omega - A$ (dopełnienie) „nie zaszło A”

$A \cup B$ (połączenie) „zaszło A **lub** B”

$A \cap B$ (przecięcie) „zaszło A **i** B”

Ω zdarzenie pewne

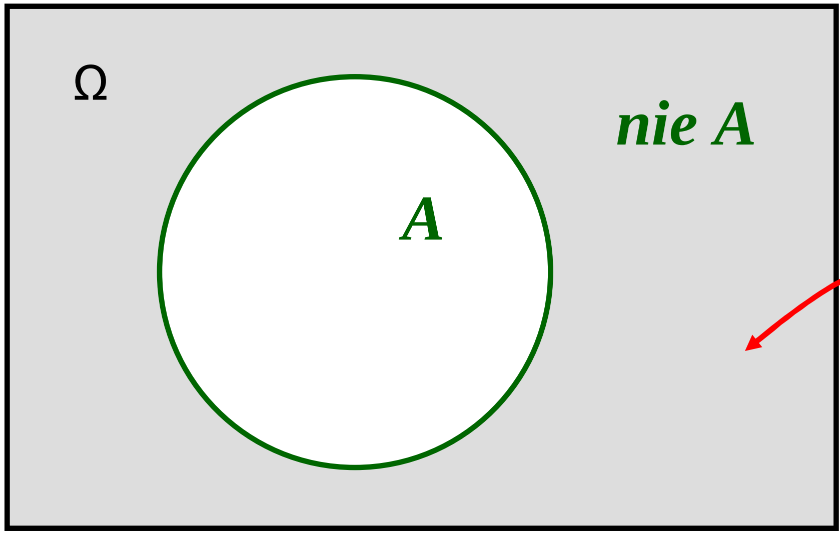
$\emptyset$ zdarzenie niemożliwe

$A \cap B = \emptyset$ zdarzenia A i B wykluczają się

$A \subset B$ jeśli zajdzie A, to musi zajść B

Dopełnienie zdarzenia

- Zdarzenie "*nie A*" występuje, gdy nie występuje *A*
- Diagram Venna : *A* (w kółku), "*nie A*" (zacienione)

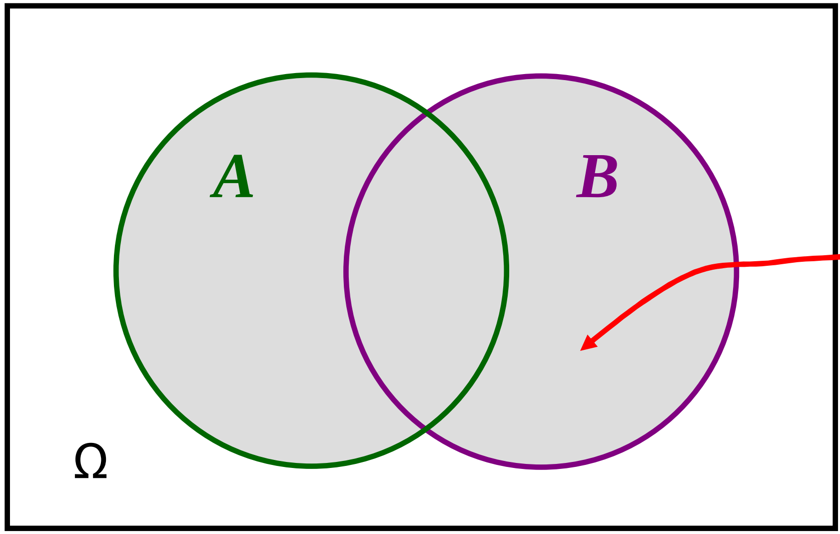


P(*nie A*) = 1 – P(*A*)

Jeżeli P(Sukcesu) = 0.7, to P(Porażki) = 1–0.7 = 0.3

Połączenie dwóch zdarzeń

- Zdarzenie $\omega \in A \cup B$ (należy do A lub do B)

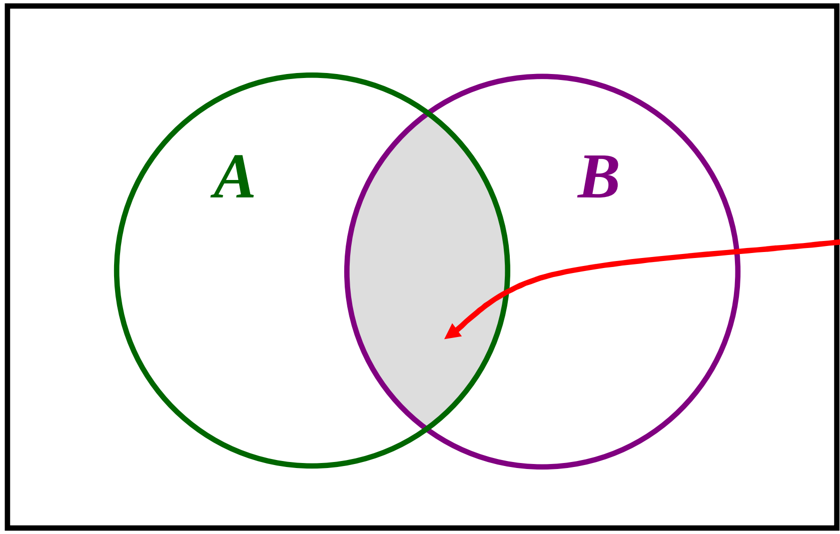


„otrzymanie liczby oczek <3" A={1} , B={2}
$A \cup B = \{1,2\}$ POŁĄCZENIE

Przecięcie dwóch zdarzeń

- Występuje, gdy oba zdarzenia wystąpią

$A \cap B$



Zdarzeniu „A i B" odpowiada zbiór $A \cap B = \{4,6\}$

$A=\{2,4,6\}$ Wypadła parzysta liczba oczek,

$B=\{4,5,6\}$ Wypadła liczba oczek>3

Przestrzeń probabilistyczna

- Przestrzeń probabilistyczna jest to para (Ω, P) czyli przestrzeń próbkowa Ω wyposażona w rozkład prawdopodobieństwa *P*
- Rozkład prawdopodobieństwa to przyporządkowanie wszystkim zdarzeniom liczb
- Rozkład prawdopodobieństwa musi spełniać szereg warunków (własności prawdopodobieństwa)

Podstawowe własności prawdopodobieństwa:

Jeśli $A \cap B = \varnothing$ **to** $P(A \cup B) = P(A) + P(B).$

$$P(\Omega) = 1 \qquad P(\varnothing) = 0$$

Dla dowolnych zdarzeń A i B

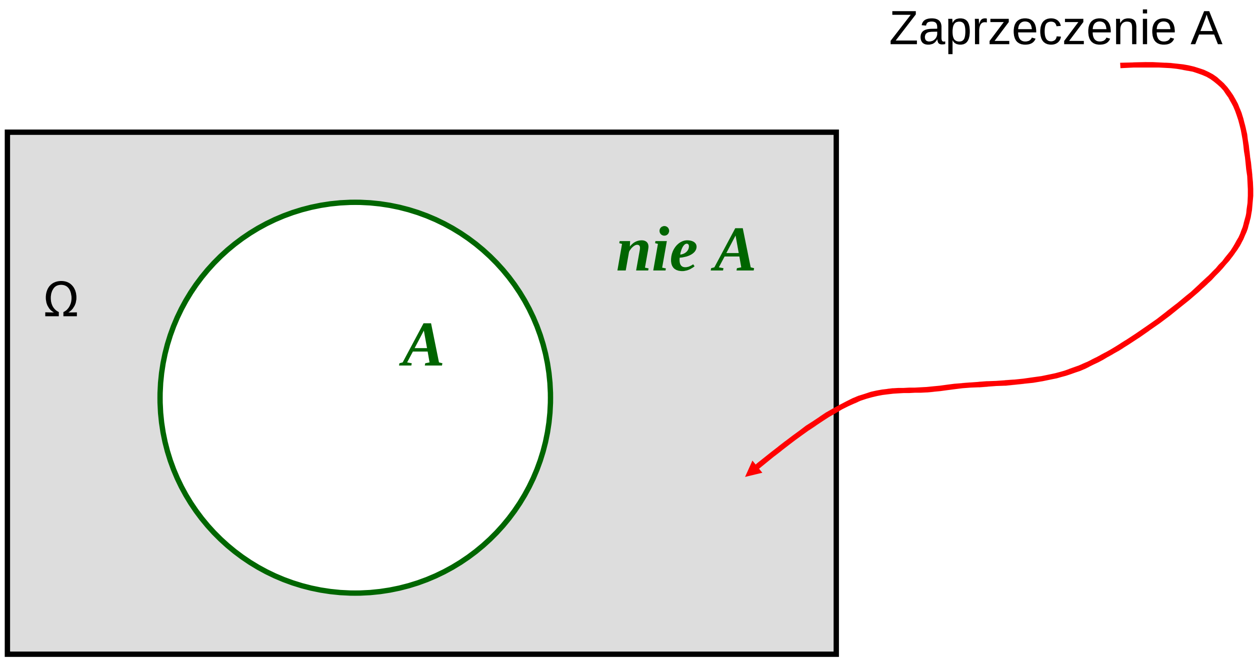
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Jeżeli $A \subset B$ **to** $P(A) \leq P(B)$

Dla każdego $A \subset \Omega$ **jest** $P(A) \leq 1$

$$P(A') = 1 - P(A)$$

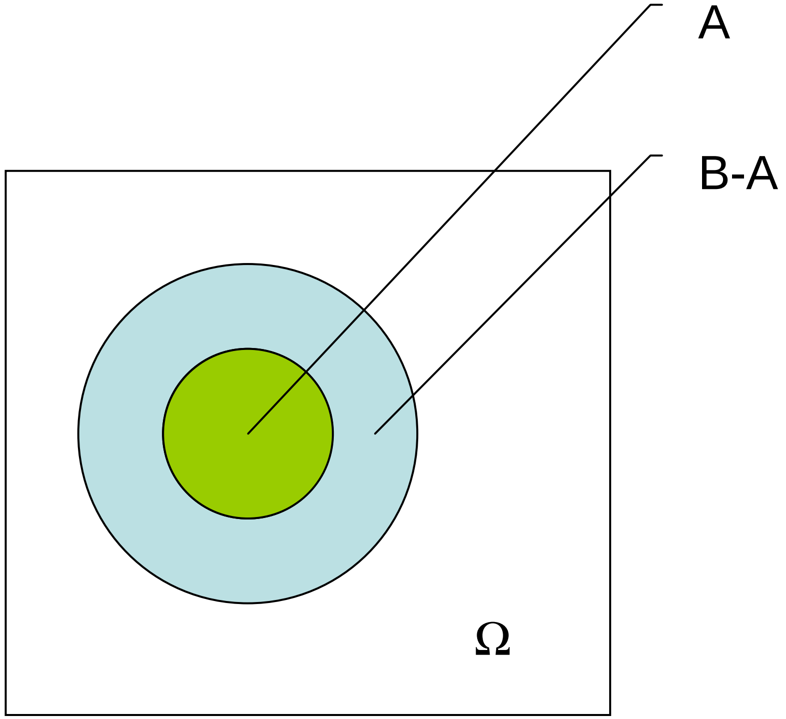
Prawdopodobieństwo zaprzeczenia (zdarzenia dopełniającego)



P(*nie A*) = P(A`) = 1 – P(*A*)

Jeżeli P(Sukcesu) = 0.7, to P(Porażki) = 1–0.7 = 0.3

Prawdopodobieństwo różnicy zdarzeń



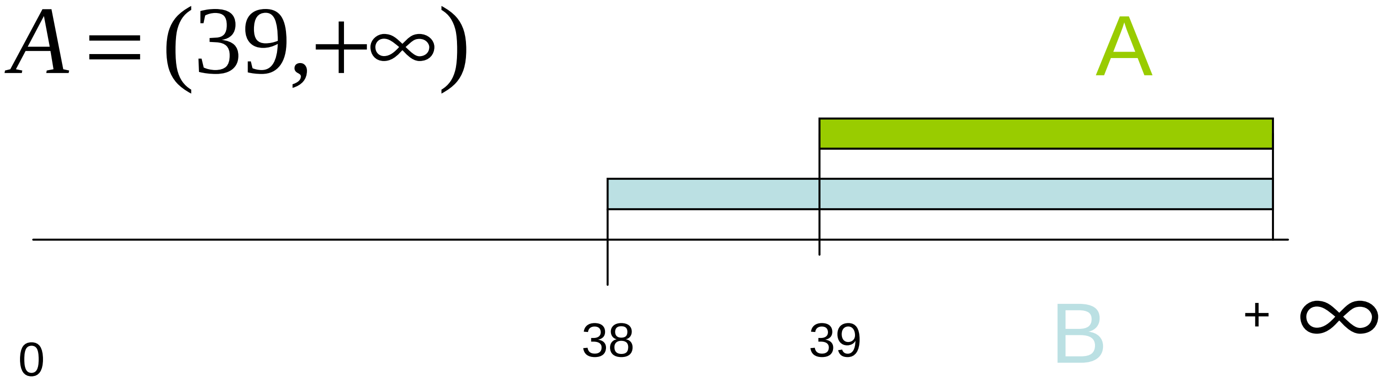
Jeżeli $A \subset B$ to $P(B\text{-}A)=P(B)-P(A)$

P(temperatura jest pomiędzy 38 a 39)=
P(temperatura>=38) –P(temperatura>39)

$$\Omega = (0, +\infty)$$

$$B = [38, +\infty)$$

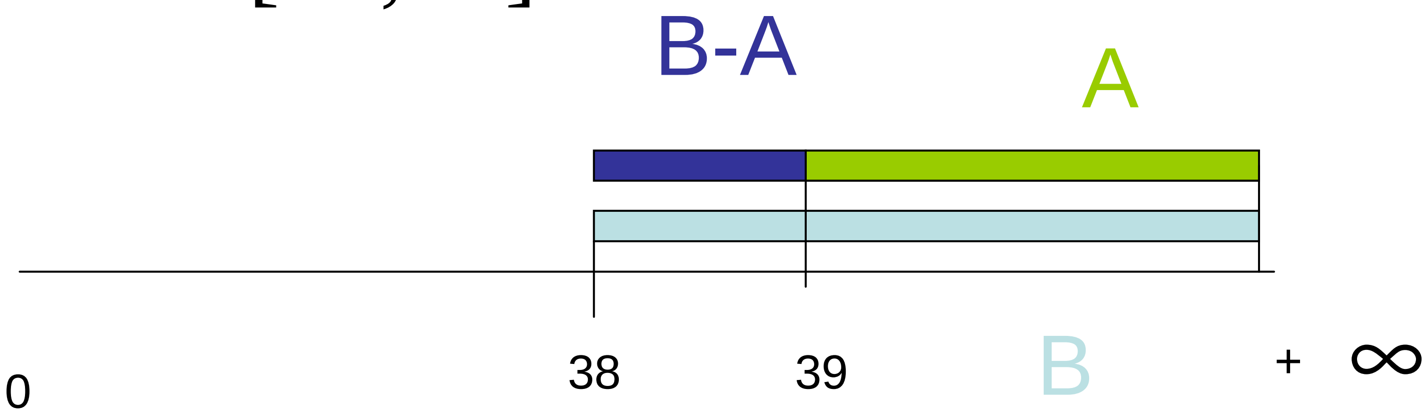
$$A = (39, +\infty)$$



$$A \subset B \qquad P(A) \leq P(B)$$

P(B-A)=P(B)-P(A)

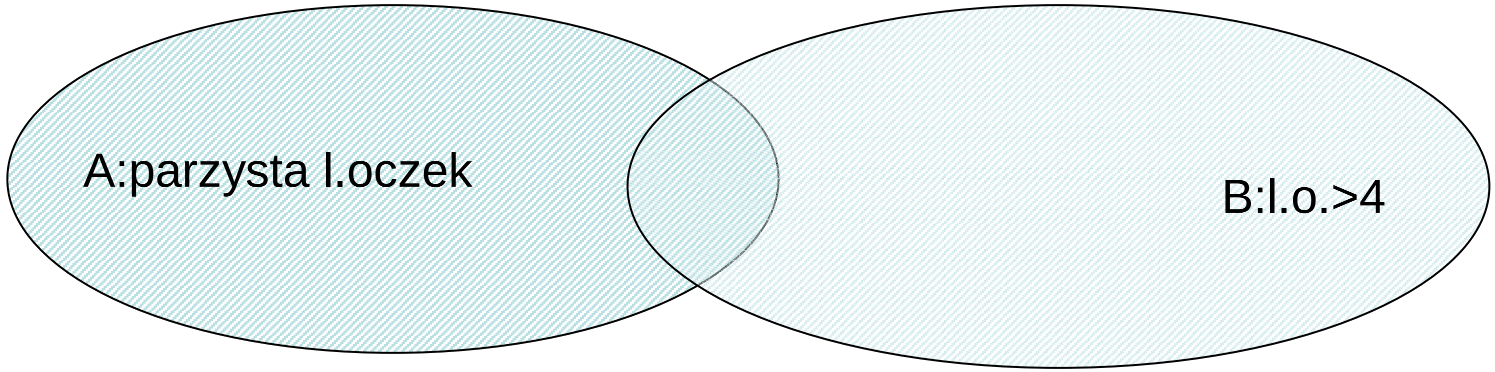
$$B - A = [38,39]$$



P(B-A)=P(B)-P(A)
P(temperatura jest pomiędzy 38 a 39)= P(temperatura>=38) – P(temperatura>39)

Jeżeli nie jest prawdą, że A $\subset$ B to trzeba korzystać ze wzoru:

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



Przykład: rzut kostką $\Omega=\{1,2,3,4,5,6\}$ A = {2,4,6} B = {5,6}

B-A = {6}

P(A) = 1/6 +1/6+ 1/6= 3/6 =1/2

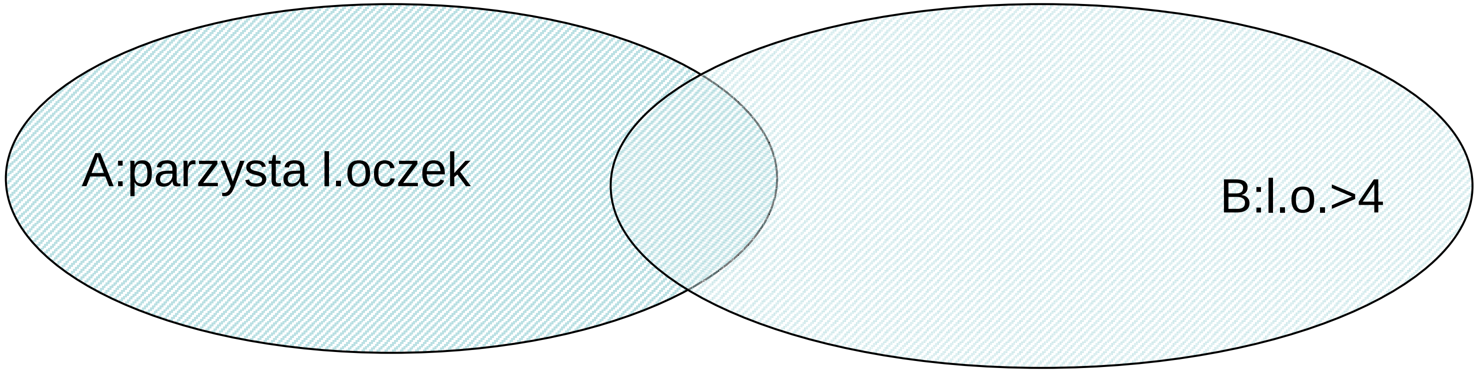
P(B) =1/6+1/6 = 2/6 =1/3

P(A i B)=1/6

P(B-A)= P(B) - P(B i A)=1/3 -1/6 = 1/6

Wzór „włączeń i wyłączeń” dla dowolnych zdarzeń A i B

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



Przykład: rzut kostką Ω={1,2,3,4,5,6} A = {2,4,6} B = {5,6}

$A \cup B$ = {2,4,5,6} $A \cap B$= {6}

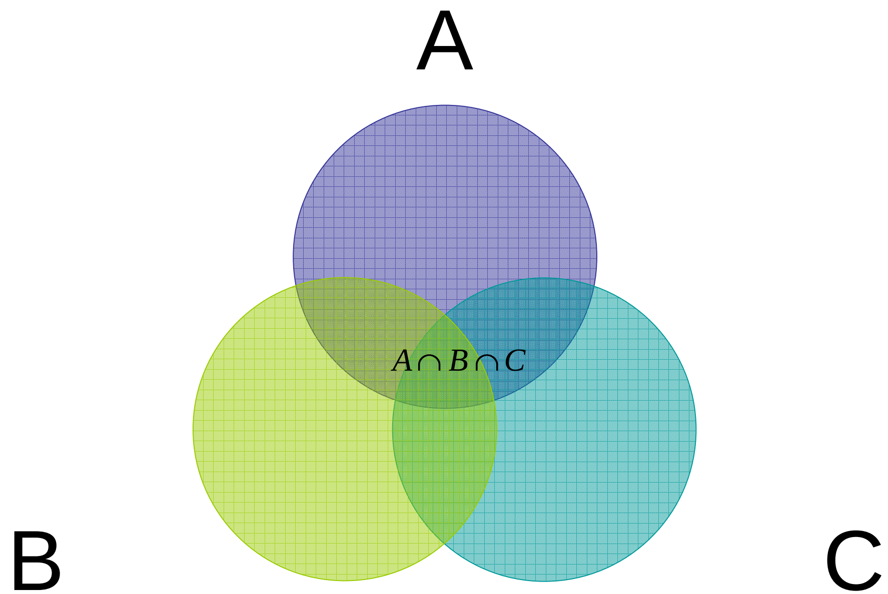
P(A) = 1/6 +1/6+ 1/6= 3/6 =1/2

P(B) =1/6+1/6 = 2/6 =1/3

P(A i B)=1/6

P(A lub B)= 1/2 +1/3 -1/6 = 4/6 =2/3

Wzór „włączeń i wyłączeń”:



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

Rozkład prawdopodobieństwa

Rozkład prawdopodobieństwa to przyporządkowanie wszystkim zdarzeniom liczb. Jest opisany za pomocą tabelki:

wynik	ω_1	ω_2	$\cdots$	ω_n
prawdopodobieństwo	p_1	p_2	...	p_n

gdzie

$$p_i \geq 0 \qquad \sum_{i=1}^{n} p_i = 1$$

Definicja. Zdarzenie losowe jest to podzbiór przestrzeni probabilistycznej

$$A \subset \Omega$$

Prawdopodobieństwo zdarzenia A jest to liczba

$$P(A) = \sum_{\omega_i \in A} p_i$$

Klasyczna definicja prawdopodobieństwa

- Klasyczna definicja prawdopodobieństwa zakłada, że wszystkie możliwe wyniki są jednakowo prawdopodobne

wynik	ω_1	ω_2	$\cdots$	ω_n
prawdopodobieństwo	1/n	1/n	...	1/n

Wtedy P(A) = #A / #Ω, gdzie # A oznacza liczbę elementów zbioru A, # Ω=n

Uwaga. Nie zawsze tak jest. Później przykład.

Rozkład prawdopodobieństwa c.d.

Przykład. 2 rzuty monetą
Rozkład prawdopodobieństwa:

wynik	OO	OR	RO	RR
prawdopodobieństwo	1/4	1/4	1/4	1/4

Przykład. Rzut dwiema jednakowymi monetami, nie odróżniamy OR od RO

wynik	2O	OR	2R
prawdopodobieństwo	1/4	1/2	1/4

Dodatkowy przykład

Przykład. Rzucamy monetą do pierwszego orła

$\Omega = \{O, RO, RRO, RRRO, \ldots,\}$

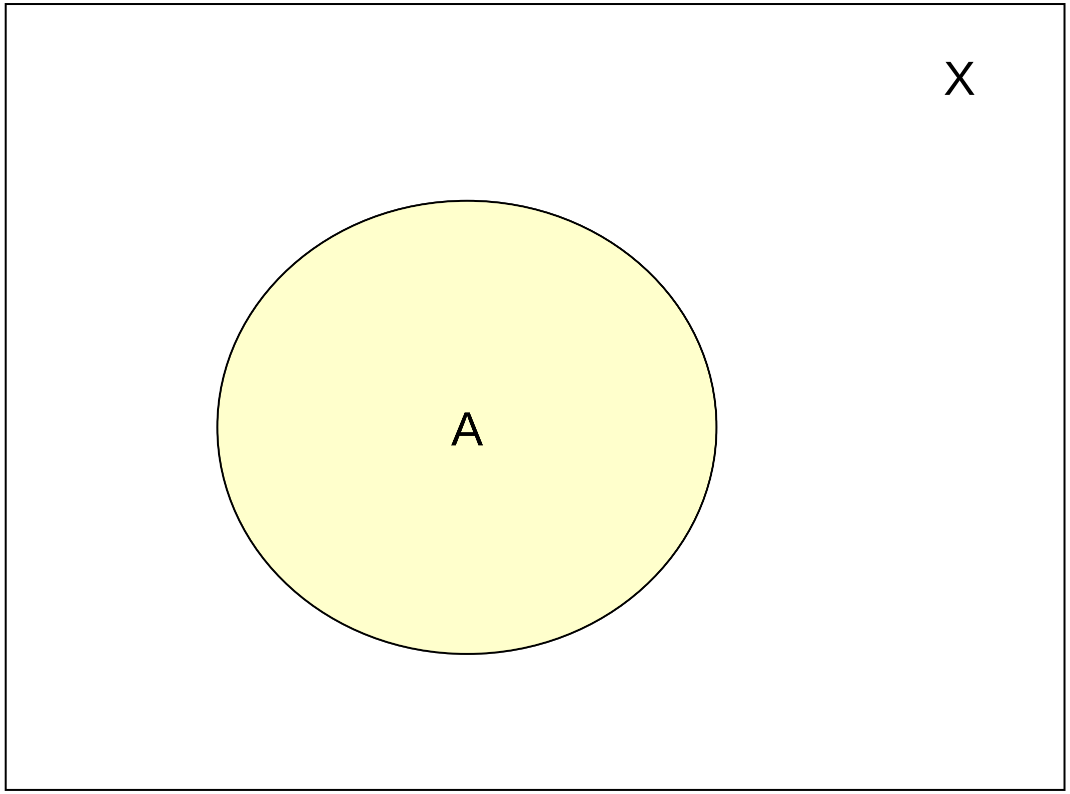
Mamy nieskończenie wiele wyników doświadczenia o różnym prawdopodobieństwie

wynik	O	RO	RRO	….	RRR…O	….
prawdopodobieństwo	1/2	1/4	1/8	…	$1/2^{k+1}$	…

Nie przyzwyczajajmy się, ze wyniki mają jednakowe prawdopodobieństwo

Możliwości – wariacje z powtórzeniami 2^k

Prawdopodobieństwo geometryczne



$$P(A) = |A| / |X| = \frac{\boxed{}}{\boxed{} + \boxed{}}$$

Kombinatoryka

- **Definicja.** Permutacja zbioru {1,2,3…,n} jest to uporządkowanie wszystkich elementów tego zbioru
- Przykład. Permutacje zbioru {1,2,3}

123 213 312

132 231 321

Twierdzenie. Liczba permutacji zbioru n-elementowego jest równa

$$n! = 1 \times 2 \times 3 \ldots \times n$$

Kombinatoryka c.d.

Przykład: permutacje

Ania, Bartek, Celinka : możliwości 3!=1*2*3=6

P(F)? A koło B

ABC +
ACB –
BAC +
BCA –
CAB +
CBA +

P(F)=4/6=2/3

Kombinatoryka c.d.

- **Definicja**. Wariacja z powtórzeniami k-wyrazowa ze zbioru $\{a_1, a_2, \ldots, a_n\}$ jest to ciąg (uporządkowany) k elementów tego zbioru (niekoniecznie różnych).
- **Twierdzenie**. Liczba k-wyrazowych wariacji z powtórzeniami ze zbioru n-elementowego jest równa $\mathbf{n^k}$ **.**

Przykład. Wariacje 2-wyrazowe z powtórzeniami ze zbioru {1,2,3}:

11	12	13
21	22	23
31	32	33

$3^2=9$

Kombinatorka c.d.

- **Definicja**. Wariacja bez powtórzeń

  k-wyrazowa ze zbioru $\{a_1, a_2, \ldots, a_n\}$ jest to ciąg (uporządkowany) k różnych elementów tego zbioru.

Twierdzenie. Liczba k-wyrazowych wariacji bez powtórzeń ze zbioru n-elementowego jest równa

$$n*(n-1)*(n-2) \qquad *(n-k+1) = n! / (n-k)!$$

Przykład. Wariacje 2-wyrazowe ze zbioru {1,2,3}:

12 13

21 23

31 32 3*2*1=6 (1*2*3)/(3-2)=6/1=6

Przykład: wariacje

Jakie jest prawdopodobieństwo, że wśród 5 losowo wybranych osób każda ma inny znak Zodiaku?

Zbiorem zdarzeń elementarnych będą ciągi k= 5 wyrazowe ze zbioru n=12 elementowego.

Takich ciągów jest $\#\Omega=12^5$ (ze zwracaniem)

R- różne znaki (w piątce żaden znak nie powtarza się)

#R=12*(12-1)*(12-2)*(12-3)*(12-4)=
12*11*10*9*8

$$P(R)=\frac{12\cdot 11\cdot 10\cdot 9\cdot 8}{12^5}=0.382$$

Kombinacja (bez powtorzeń)

- Definicja. Kombinacja k-wyrazowa ze zbioru $\{a_1, a_2, \ldots, a_n\}$ jest to podzbiór (nieuporządkowany) złożony z k (różnych) elementów tego zbioru
- Twierdzenie. Liczba kombinacji k-elementowych ze zbioru n-elementowego jest równa

$$\binom{n}{k} = n! / k!(n-k)! = \frac{n(n-1)(n-2)...(n-k+1)}{k!}$$

Przykład. Kombinacje 2-wyrazowe ze zbioru {1,2,3}:

{1,2} {1,3} {2,3} 3!/2!(3-2)!=3

Schemat losowania bez zwracania

Losujemy n kul z urny zawierającej b kul białych i c kul czarnych. Za każdym razem po wylosowaniu nie wrzucamy kuli do urny.

Przykłady

• Losowanie „Multilotek”
• Pobieranie próbki losowej z populacji wyborców (losowanie próbek reprezentatywnych).

Losowanie próbki bez zwracania

Wybieramy losowo n różnych elementów z populacji liczącej r elementów (losujemy n kul z urny zawierającej r kul, nie zwracając uprzednio wylosowanych kul do urny). Przyjmujemy, że wynikiem jest nieuporządkowany układ n kul, czyli kombinacja n z r.
Przestrzeń probabilistyczna Ω zawiera

$$\binom{r}{n}$$

jednakowo prawdopodobnych wyników. Wyobraźmy sobie, że wyróżniamy pewne m elementów populacji (mamy w urnie m „czarnych” i r-m „białych”

Twierdzenie. Prawdopodobieństwo wylosowania dokładnie *k* razy białej kuli spośród *b* białych podczas losowania *n* kul z wszystkich *b*+*c* bez zwracania jest równe:

$$P(k) = \frac{\binom{b}{k}\binom{c}{n-k}}{\binom{b+c}{n}}$$

Przykład. Multilotek. Gra polega na wytypowaniu 10 numerów spośród 80. Organizator gry losuje potem 20 numerów spośród 80. Prawdopodobieństwo tego, że wśród wylosowanych 20 numerów znajdą się wszystkie 10 wytypowane przez nas jest równe:

$$\frac{\binom{10}{10}\binom{70}{10}}{\binom{80}{20}} \approx \frac{1}{8911711.2}$$

(r=80, m=10, n=20, k=10). Trafienie „dziesiątki” zdarza się raz na 8 mln gier.

Przykład. Karty. Jakie jest prawdopodobieństwo że wśród wylosowanych 4 kart znajdą się 4 asy?

$$\frac{\binom{4}{4}\binom{48}{0}}{\binom{52}{4}}$$

Losowanie próbki ze zwracaniem

Wybieramy losowo n-krotnie po 1 elemencie z populacji liczącej r elementów,
Możemy wielokrotnie wylosować ten sam element (np. losujemy n kul
z urny zawierającej r kul zwracając za każdym razem wylosowaną kulę
wrzucając do urny).
Przyjmujemy, że wynikiem jest uporządkowany układ n kul, czyli wariacja z
powtórzeniami n z r. Przestrzeń probabilistyczna Ω zawiera

$$r^{n}$$

jednakowo prawdopodobnych wyników.

Schemat Bernoulliego

- Powtarzamy wielokrotnie (*n* razy) niezależnie (wynik następnego doświadczenia nie zależy od wyników poprzednich) doświadczenie losowe, w którym możliwe są dwa wyniki umownie nazwane „sukces” i „porażka”. Prawdopodobieństwo sukcesu w pojedynczym doświadczeniu oznaczymy przez p, porażki q=1-p

Przykłady

- Rzuty monetą „orzeł”=sukces,”reszka”=porażka
- Rzuty kostą „szóstka”= sukces, „inny wynik”=porażka
- Płeć noworodków „dziewczynka”=sukces, „chłopiec”=porażka
- Losowanie z urny ze zwracaniem (w urnie *b*-kul białych i *c* kul czarnych) „kula biała”=sukces, „kula czarna” =porażka
- Losowanie bez zwracania nie jest schematem Bernouliego bo wynik następnego losowania zależy od wyniku poprzedniego

Twierdzenie. W schemacie Bernoulliego, prawdopodobieństwo otrzymania dokładnie *k* sukcesów (i *n-k* porażek) jest równe

$$P(k) = \binom{n}{k} p^{k} (1 - p)^{n-k}$$

$$P(k) = \frac{n!}{k!(n-k)!} p^{k} (1 - p)^{n-k}$$

Rozkład prawdopodobieństwa opisany tym wzorem nazywa się dwumianowy

Rozkład dwumianowy - schemat Bernoulliego

Przykład. Jakie jest prawdopodobieństwo P wylosowania z populacji (w której proporcja kobiet wynosi p=1/2) dokładnie k=0,1,2,3,…n kobiet?

$$P(k) = \binom{n}{k}(1/2)^k (1-1/2)^{n-k} = \binom{n}{k} 1/2^n$$

Szansa, że wsród n=10 noworodków będzie dokładnie k=5 dziewczynek jest w przybliżeniu 25% bo:

$$P(5) = \binom{10}{5} 1/2^{10} = 252/1024 = 0.2461$$

0!=1

Przykład. (Wielokrotne rzuty kostką) n=30 razy rzucamy kostką. Prawdopodobieństwo Wyrzucenia szóstki w jednym rzucie jest równe p=1/6. Chcemy obliczyć prawdopodobieństwo zdarzenia polegającego na tym, że wśród 30 rzutów pojawi się dokładnie k =5 razy „szóstka”.

$$\binom{30}{5}\left(\frac{1}{6}\right)^{5}\left(\frac{5}{6}\right)^{25} \approx 0.1921$$

(n=30, k=5, p=1/6)

Prawdopodobieństwo warunkowe

- Oceniamy szanse zajścia zdarzenia losowego A wiedząc, że zaszło zdarzenie losowe B. Mówimy wtedy o prawdopodobieństwie warunkowym i używamy oznaczenia P(A|B).
- Prawdopodobieństwo zajścia zdarzenia A pod warunkiem, że zdarzenie B zaszło.

Przykłady:

- P(„Chory” jeżeli wiadomo, że (jest dane) „Wynik testu pozytywny”)
- P („Dostać pracę” jeżeli „Słaba prezentacja”)
- P(A I B)

Prawdopodobieństwo warunkowe cd

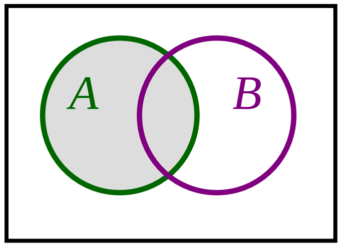
- **Definicja.** Prawdopodobieństwo warunkowe zdarzenia A pod warunkiem zajścia zdarzenia B jest określone wzorem:

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

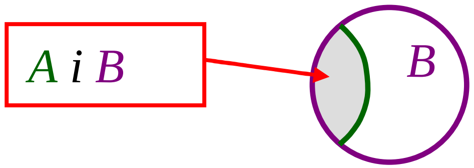
Zakładamy przy tym, że $P(B)>0$.

Prawdopodobieństwo warunkowe c.d.

- Bezwarunkowe
  - Prawdopodobieństwo *A*



- Warunkowe
  - wiemy, że *B* zaszło



Twierdzenie. (Wzór łańcuchowy) Z definicji wynika, że prawdopodobieństwo koniunkcji (łącznego zajścia):

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

Dla trzech zdarzeń:

$$\boxed{\begin{aligned} &P(A_1 \cap A_2 \cap A_3) \\ &= P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1 \cap A_2) \end{aligned}}$$

Przykład. Posługiwanie się wzorem łańcuchowym w życiu codziennym:

- prawdopodobieństwo, że osoba zachoruje na grypę w tym roku oceniam na 0.5

- jeśli zachoruje, to z prawdopodobieństwem 0.01 mogą wystąpić powikłania

- jeśli wystąpią powikłania to z prawdopodobieństwem 0.8 hospitalizacja z powodu powikłań po grypie

Zatem prawdopodobieństwo, że wystąpi hospitalizacja po powikłaniach po grypie

$$=0.5*0.01*0.8=0.004$$

Przykład. Prawdopodobieństwo warunkowe.

Dla osób chorych (A) test daje wynik pozytywny (B) z prawdopodobieństwem p=0.90.

Dla osób zdrowych (nieA, A\`), test daje wynik negatywny (nieB, B\`) z prawdopodobieństwem p=0.95.

Jakie jest prawdopodobieństwo, że pacjent jest chory jeśli test dał wynik pozytywny?
P(„Chory”|”Wynik testu pozytywny”)= P(A|B) ?

P(B|A)=0.90 jeśli wiadomo, że osoba jest chora

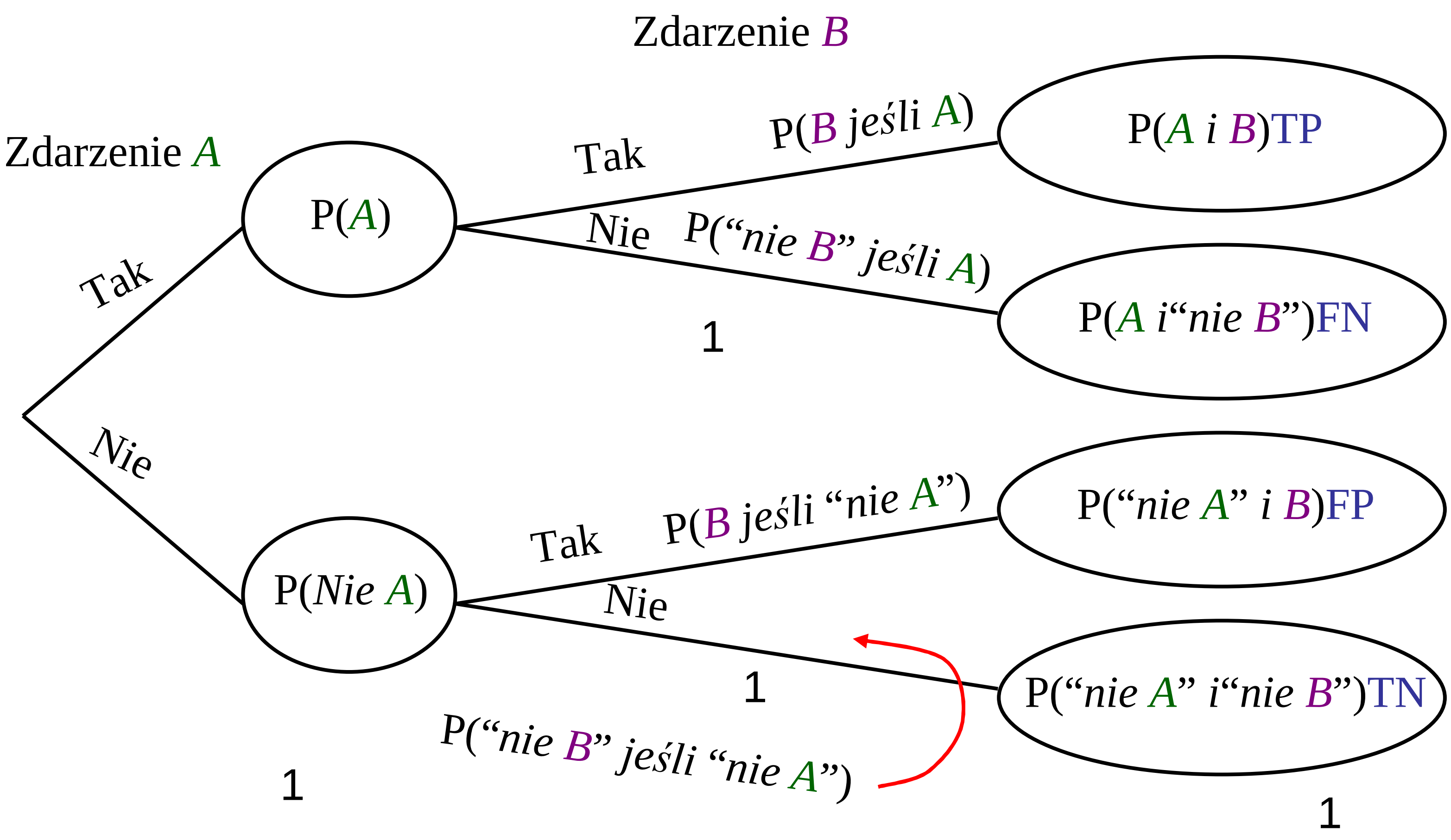
P(B\`|A\`)=0.95 jeśli wiadomo, że osoba nie jest chora

Prawdopodobieństwo występowania choroby P(A)=0.08

- Drzewka prawdopodobieństwa - metoda graficzna do rozwiązywania problemów związanych z ustaleniem prawdopodobieństwa wielu zdarzeń
- Tabele prawdopodobieństwa

Drzewka prawdopodobieństwa c.d.

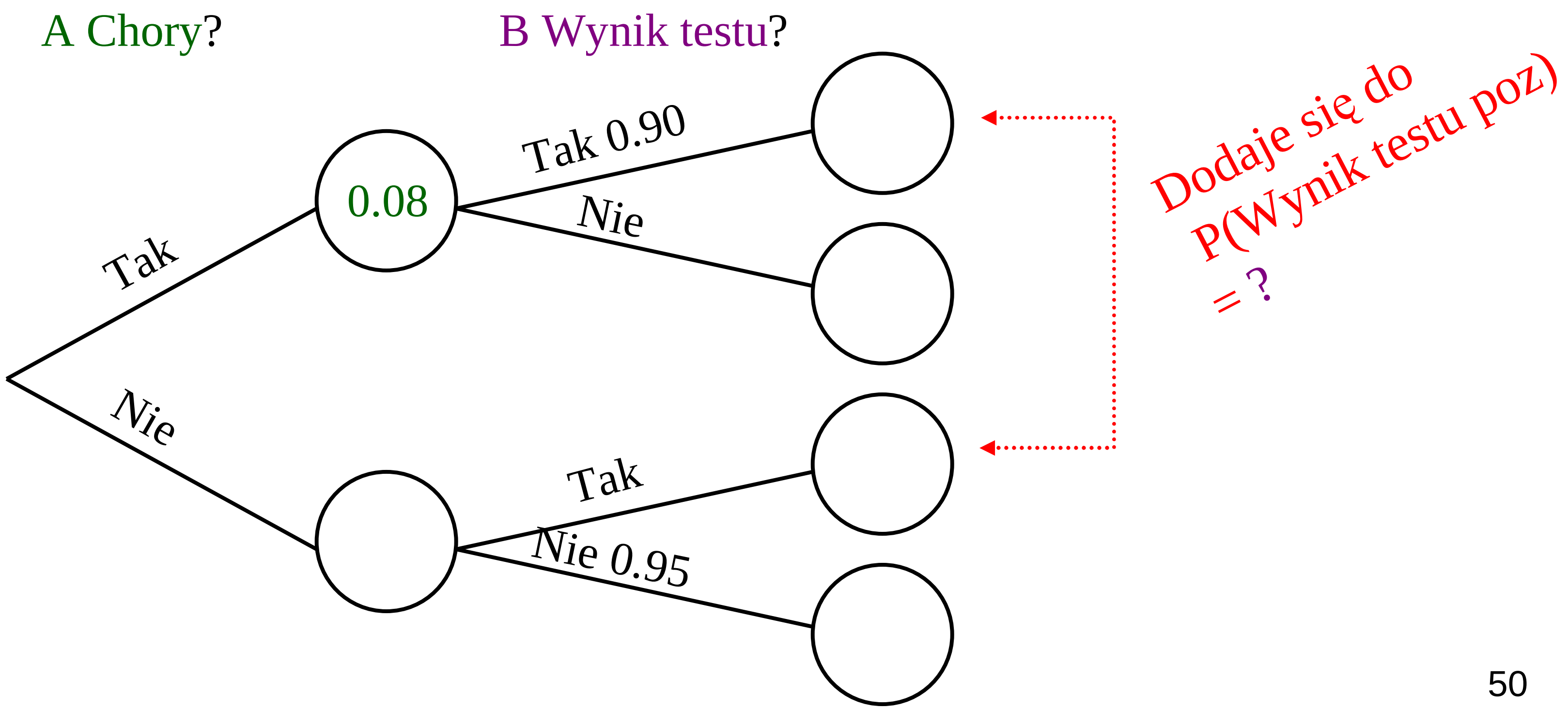
A - chory B - wynik testu pozytywny



◆Zanotować podstawowe informacje

- P(choryA) = 0.08,   P(zdrowyA`) = 0.92 – uzupełnione do 1
- P(B|A) = 0.90,   P(wynik testu negatywnyB`|A`)=0.95

Jakie jest prawdopodobieństwo, że pacjent jest chory, jeśli test dał wynik pozytywny ?

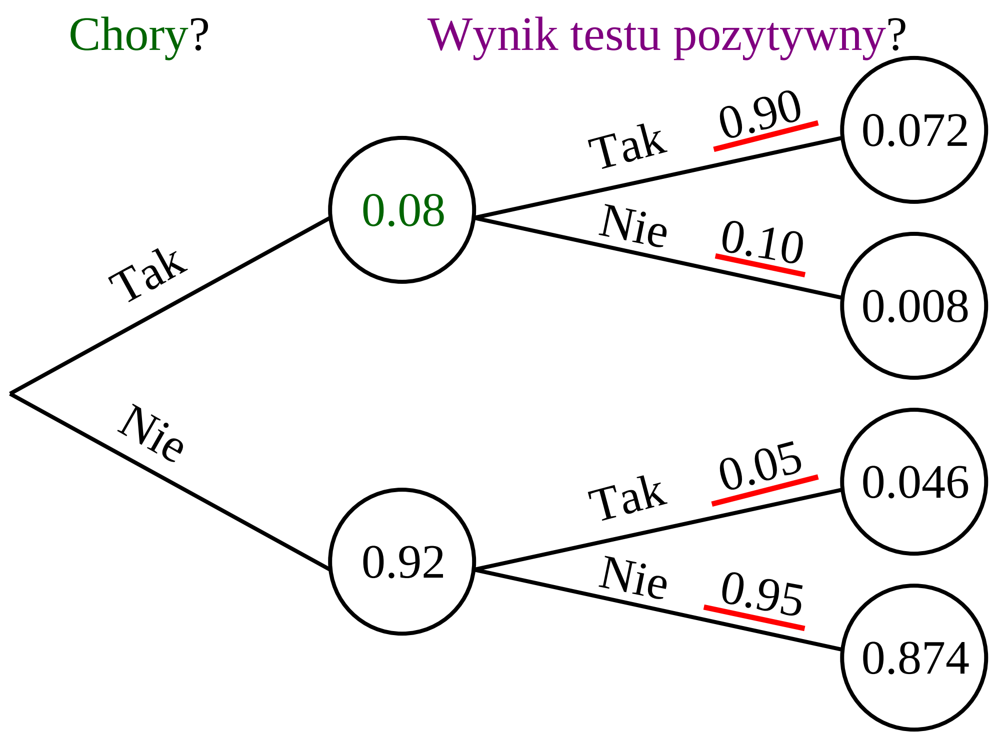


Przykład (pełne drzewko)

◆Zauważmy, że prawdopodobieństwo warunkowe jest wynikiem dzielenia

0.072/0.08 = 0.90, 0.008/0.08 = 0.10

0.046/0.92 = 0.05, 0.874/0.92 = 0.95



Przykład: Tabela połączonych prawdopodobieństw

Pokazuje prawdopodobieństwo dla każdego zdarzenia, dopełnienia i kombinacji używając „i”

Zauważmy: $$P(A \cap B) = P(B \mid A) \cdot P(A)$$

P(A *i* B)

P(“*nie* B” *i* A)

P(B|A)=0.90
P(A)=0.08

		B Wynik testu		
		Tak	Nie	
A	Tak	0.072	0.008	0.08
	Nie	0.046	0.874	0.92
		0.118	0.882	1

P(*nie* A)

P(B *i* “*nie* A”)

P(B)

P(“*nie* A” *i* “*nie* B”)

Przykład c.d.

$$P(A|B) = ?$$

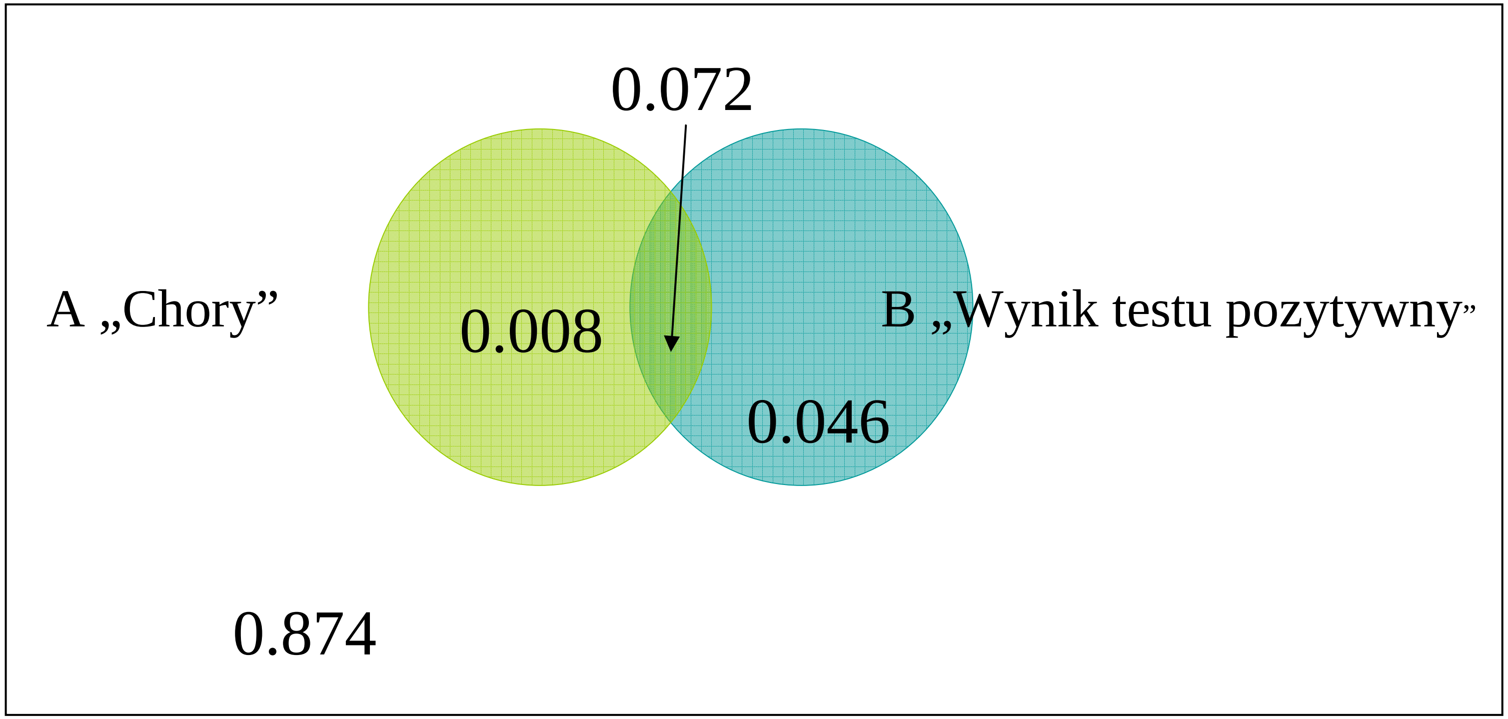
P(„Chory”|”Wynik testu pozytywny”)=?

$$\frac{P(A \cap B)}{P(B)} = \frac{0.072}{0.072 + 0.046} = \frac{0.072}{0.118} = 0.610$$

Pośród wszystkich pacjentów dla których test dał wynik pozytywny tylko 61% będzie chorych. Pozostałe 39% pacjentów będzie zdrowych. Niechętnie stawiana diagnoza dla rzadkich chorób.

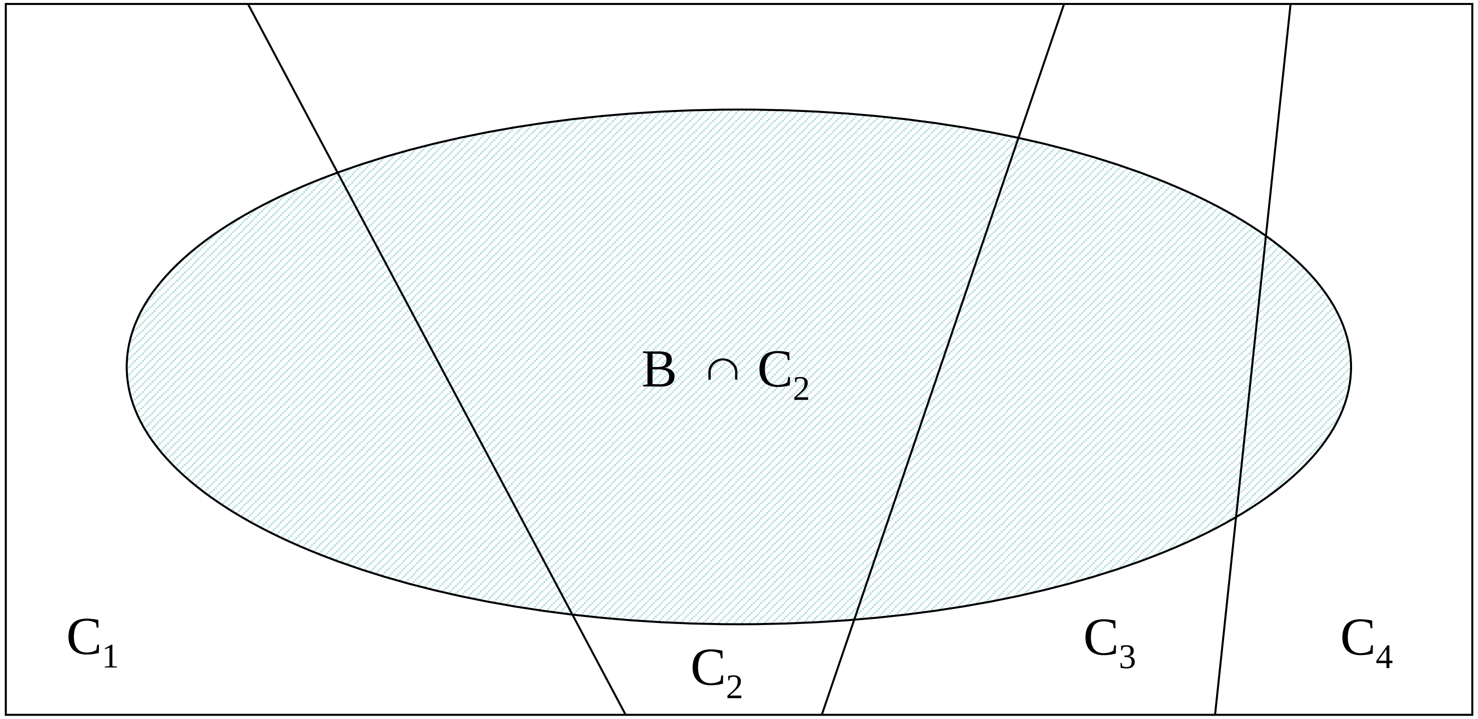
<u>Przykład c.d.</u> Prawdopodobieństwo warunkowe

Dodatkowa informacja którą mamy zawęża przestrzeń probabilistyczną do zbioru B



P(B)=P(A i B)+P(nieA i B) =0.072+0.046=0.118   0.072/0.118

Prawdopodobieństwo całkowite



$$P(B) = P((B \cap C_1) \cup (B \cap C_2) \cup ...(B \cap C_k)) =$$

$$\sum_i P(B \cap C_i) = \sum_i P(B|C_i)P(C_i)$$

Wzór Bayesa

Przy założeniach wzoru na prawdopodobieństwo całkowite, jeśli ponadto
P(B)>0 to

$$P(C_i \mid B) = \frac{P(C_i \cap B)}{P(B)} = \frac{P(B \mid C_i)P(C_i)}{\Sigma_k P(B \mid C_k)P(C_k)}$$

Wracając do przykładu:

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B \mid A)P(A) + P(B \mid A`)P(A`)}$$

$$= \frac{0.072}{0.072 + 0.046} = \frac{0.072}{0.118} = 0.610$$

Zastosowanie wzoru Bayesa

Przypomnijmy przykład:

Wiemy jakie jest prawdopodobieństwo, że test wykaże wynik

pozytywny jeśli pacjent jest chory: P(B|A) – dane $P(B|A)=0.90$

Zadajemy pytanie odwrotnie:

Jakie jest prawdopodobieństwo, że pacjent jest chory jeśli test dał wynik pozytywny? $P(A|B)=?$

P(„Chory”|”Wynik testu pozytywny”)= P(A|B) ?

Niezależność zdarzeń losowych

Dwa zdarzenia są **niezależne** jeżeli informacja o jednym nie ma wpływu na wystąpienie drugiego.

$$P(A|B) = P(A)$$

$$P(B|A) = P(B)$$

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

Przykład

- Przykład dotyczący testu c.d.
- $P(A\ i\ B) = 0.072$

  $P(A)\ \ P(B) = 0.072 \times 0.118 = 0.0084$

Zdarzenia „chory” i „wynik testu pozytywny”

nie są niezależne.

Przykład. Niezależność zdarzeń

PŁEĆ	GRYPA		Sumy
	G	Nie G	
M	0.20	0.30	0.50
K	0.20	0.30	0.50
Sumy	0.40	0.60	1

$$P(K \mid G) = \frac{P(K \cap G)}{P(G)} = \frac{0.20}{0.40} = 0.5 = P(K)$$

P(G)=0.40 , P(K)=0.50

P(G)P(K)=0.40*0.50=0.20

K i G są niezależne

Taki sam procent chorych wśród kobiet jak chorych w całej populacji.

Informacja o płci nie wpływa na ocenę szans „grypa” „nie grypa”.

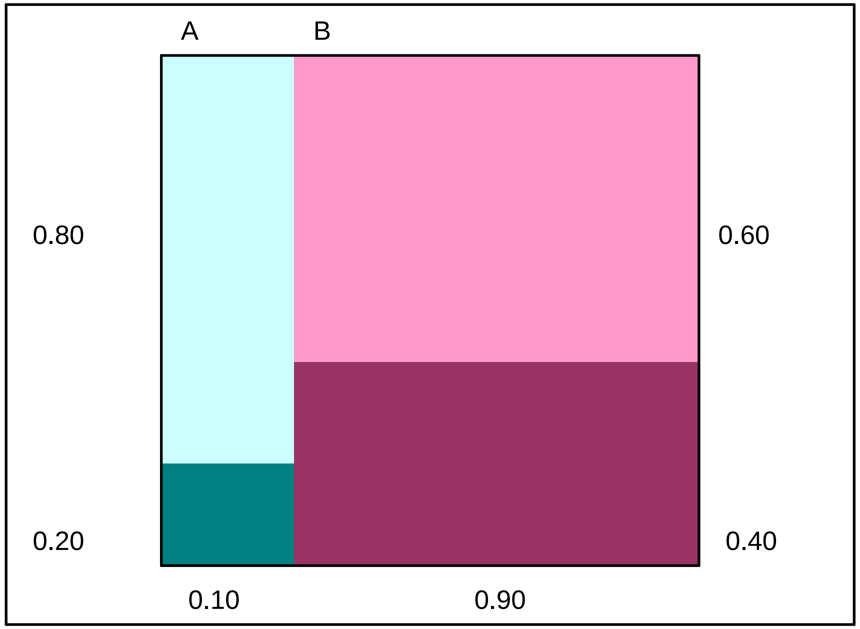
Przykład zależność

- Pewna chorobę leczymy dwoma lekami A i B.
- Lek A powoduje istotną poprawę w 80% przypadków.
- Lek B powoduje poprawę w 60% przypadków.
- Lek A jest drogi i stosujemy go w 10% przypadków.
- Lek B jest tańszy i stosujemy go w 90% przypadków.
- U ilu % pacjentów nastąpi istotna poprawa przy leczeniu lekami A i B?

<u>Przykład.</u> Prawdopodobieństwo całkowite

- Istotna poprawa spodziewana jest u 62% pacjentów

- P(popr)= 0.10*0.80+ 0.90*0.60=0.08+0.54 =0.62



Czułość i specyficzność

Przykład. (Koronacki, Ćwik) Wyobraźmy sobie, że poddano ocenie pewien test medyczny. Przebadano n=300 osób, w tym n_1=100 osób chorych i n_2=200 osób zdrowych.

TP- liczba przypadków, gdy test dał poprawnie wynik dodatni
TN- liczba poprawnych wyników ujemnych
FP- liczba fałszywych wyników dodatnich, tzn. test dał wynik dodatni, choć pacjent zdrowy
FN- liczba fałszywych wyników ujemnych

	Osoba klasyfikowana jako zdrowa (wynik testu negatywny)	Osoba klasyfikowana jako chora (pozytywny wynik testu)
Osoba zdrowa	TN	FP
Osoba chora	FN	TP

Uzyskano następujące wyniki:
1. Test rozpoznał poprawnie 97 przypadków choroby (TP)
2. Oraz u 176 osób zdrowych dał wynik negatywny (TN)

	Osoba klasyfikowana jako zdrowa (wynik testu negatywny)	Osoba klasyfikowana jako chora (pozytywny wynik testu)
Osoba zdrowa	176 (TN)	24 (FP)
Osoba chora	3 (FN)	97 (TP)

Suma liczb w pierwszym wierszu daje prawdziwą liczbę osób zdrowych – 200
Pierwsza kolumna daje liczbę zaklasyfikowań pacjentów do kategorii osób zdrowych -179

	Osoba klasyfikowana jako zdrowa (wynik testu negatywny)	Osoba klasyfikowana jako chora (pozytywny wynik testu)
Osoba zdrowa	TN	FP
Osoba chora	FN	TP

$$\frac{FP+FN}{TN+FP+FN+TP}=\frac{27}{300}=0.09$$

Liczba 0.09 będąca oszacowaniem prawdopodobieństwa błędnej klasyfikacji odnosi się do sumy błędów a nie oddzielnie każdego z dwóch typów błędów. Dlatego badacze oceniający testy medyczne wprowadzili pojęcia czułości i specyficzności testu.

	Osoba klasyfikowana jako zdrowa (wynik testu negatywny)	Osoba klasyfikowana jako chora (pozytywny wynik testu)
Osoba zdrowa	TN	FP
Osoba chora	FN	TP

Czułość (dla chorych) = $$\frac{TP}{TP + FN}$$ = 97/100 = 0.97

Specyficzność (dla zdrowych) = $$\frac{TN}{TN+FP} = 1 - \frac{FP}{TN+FP}$$

= 176/200 = 0.88 = 1 - 24/200

Czułość testu daje oszacowanie prawdopodobieństwa przewidzenia przez test choroby pod warunkiem, że pacjent jest chory na badaną chorobę.

Specyficzność testu daje oszacowanie prawdopodobieństwa przewidzenia przez test, że pacjent jest zdrowy pod warunkiem, że pacjent rzeczywiście nie cierpi na badaną chorobę.

Czułość i specyficzność podaje się w % .

Projektant testu chciałby, aby test był czuły, czyli dawał wynik dodatni, gdy pacjent jest chory, ale jednocześnie aby był specyficzny, czyli nie dawał wyniku dodatniego gdy pacjent jest zdrowy.

Test, który zawsze daje wynik dodatni ma
czułość maksymalną = 1,
ale również ma jednocześnie specyficzność = 0

- Wróćmy do przykładu drzewko prawdopodobieństwa s.51

	Osoba klasyfikowana jako zdrowa (wynik testu negatywny)	Osoba klasyfikowana jako chora (pozytywny wynik testu)
Osoba zdrowa	TN 0.874	FP 0.046
Osoba chora	FN 0.008	TP 0.072

P(A *i* B)

P(“*nie* B” *i* A)

P(A)

P(*nie* A)

		B: Tak	B: Nie	
A	Tak	0.072	0.008	0.08
	Nie	0.046	0.874	0.92
		0.118	0.882	1

P(B *i* “*nie* A”)

P(B)

P(“*nie* A” *i* “*nie* B”)

	Osoba klasyfikowana jako zdrowa (wynik testu negatywny)	Osoba klasyfikowana jako chora (pozytywny wynik testu)
Osoba zdrowa	TN 0.874	FP 0.046
Osoba chora	FN 0.008	TP 0.072

Obliczmy czułość i specyficzność testu

$$\frac{TP}{TP + FN}$$ = 0.072/(0.072+0.008) = 0.9

Dla osób chorych (A) test daje wynik pozytywny (B) z prawdopodobieństwem p=0.90.

$$\frac{TN}{TN+FP} = 1 - \frac{FP}{TN+FP}$$ =1-0.046/(0.874+0.046)=1-0.05=0.95

Dla osób zdrowych (nieA, A\`), test daje wynik negatywny (nie B, B\`) z prawdopodobieństwem p=0.95.

TP- frakcja pacjentów dla których zachodzą zdarzenia A i B
Chory i wynik testu pozytywny
TN- frakcja pacjentów dla których zachodzą zdarzenia nieA i nieB
Zdrowy i wynik testu negatywny
FP- frakcja pacjentów dla których zachodzą zdarzenia nieA i B
Zdrowy i wynik testu pozytywny
FN- frakcja pacjentów dla których zachodzą zdarzenia A i nieB
Chory i wynik testu negatywny.

$P(B|A)=0.90$ jeśli wiadomo, że osoba jest chora
$P(B`|A`)=0.95$ jeśli wiadomo, że osoba nie jest chora

Czułość i specyficzność to prawdopodobieństwa warunkowe.
Czułość to prawdopodobieństwo poprawnego działania testu dla osób chorych.
Specyficzność to prawdopodobieństwo poprawnego działania testu dla osób zdrowych (nie chorych na badaną chorobę).