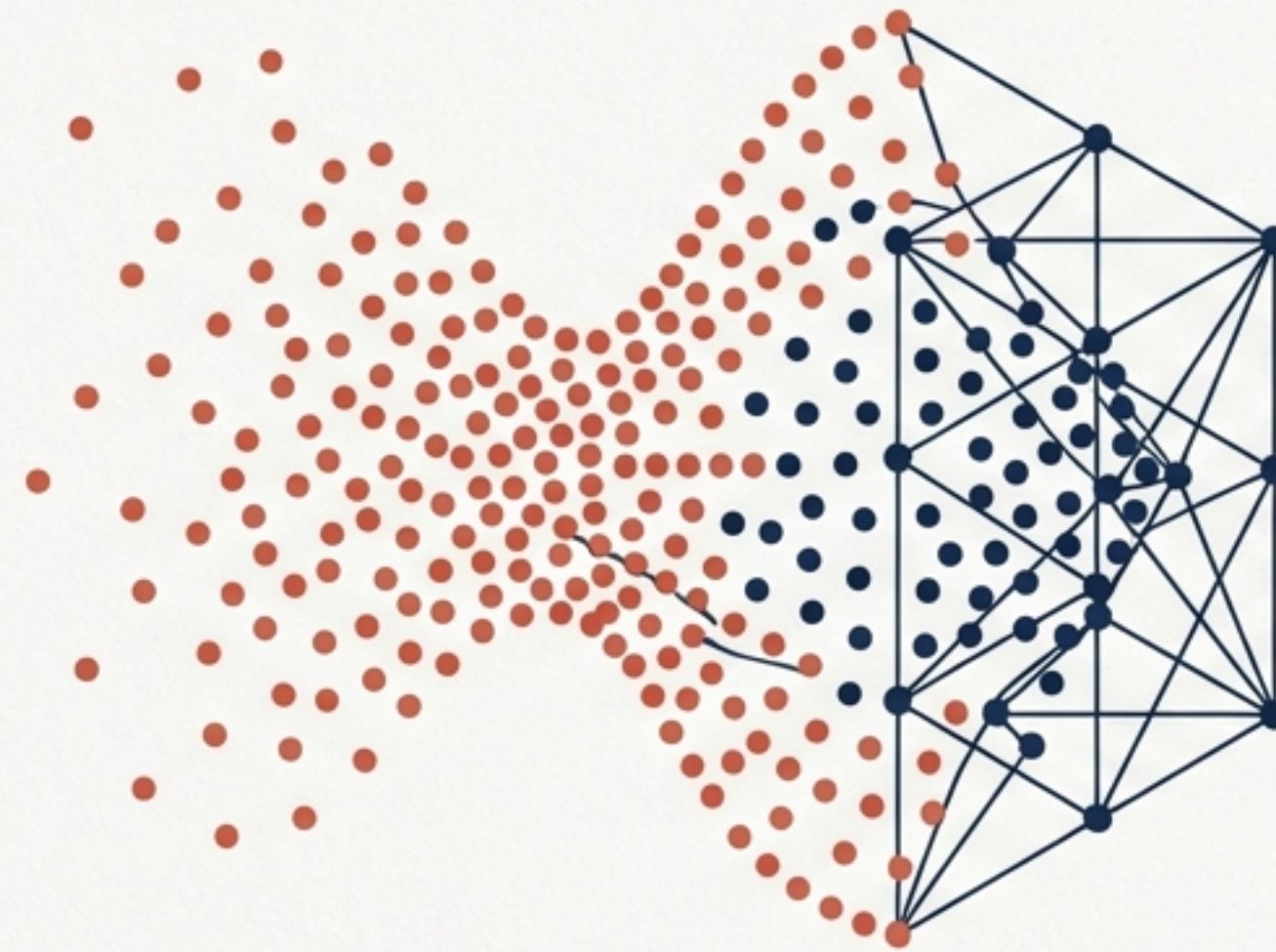


# Sztuczna Inteligencja: metody predyktywne i Analiza Bayesa

Od surowych danych do decyzji w warunkach niepewności



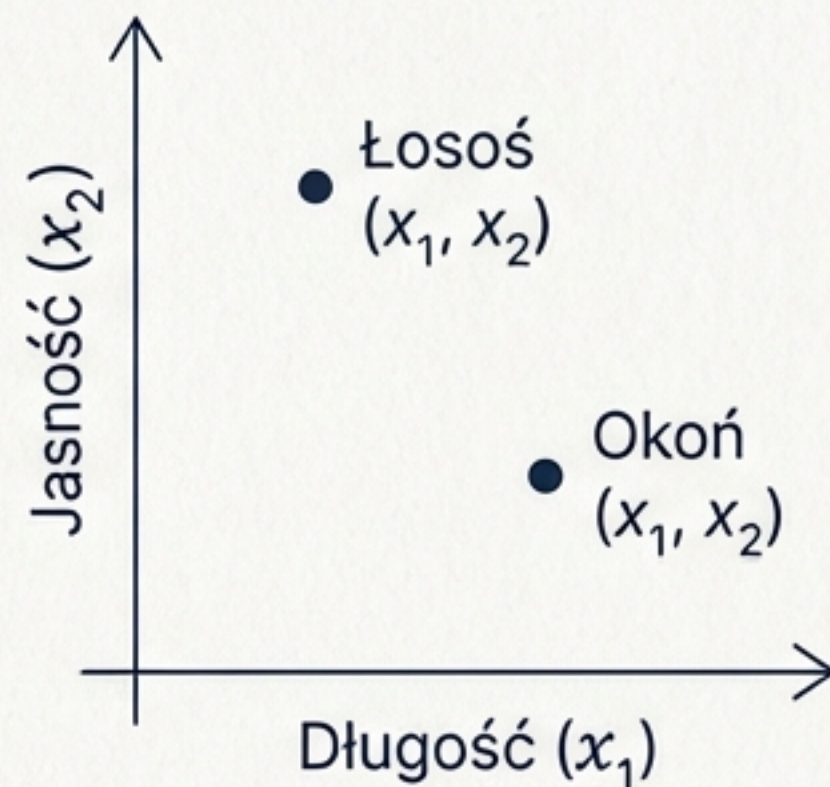
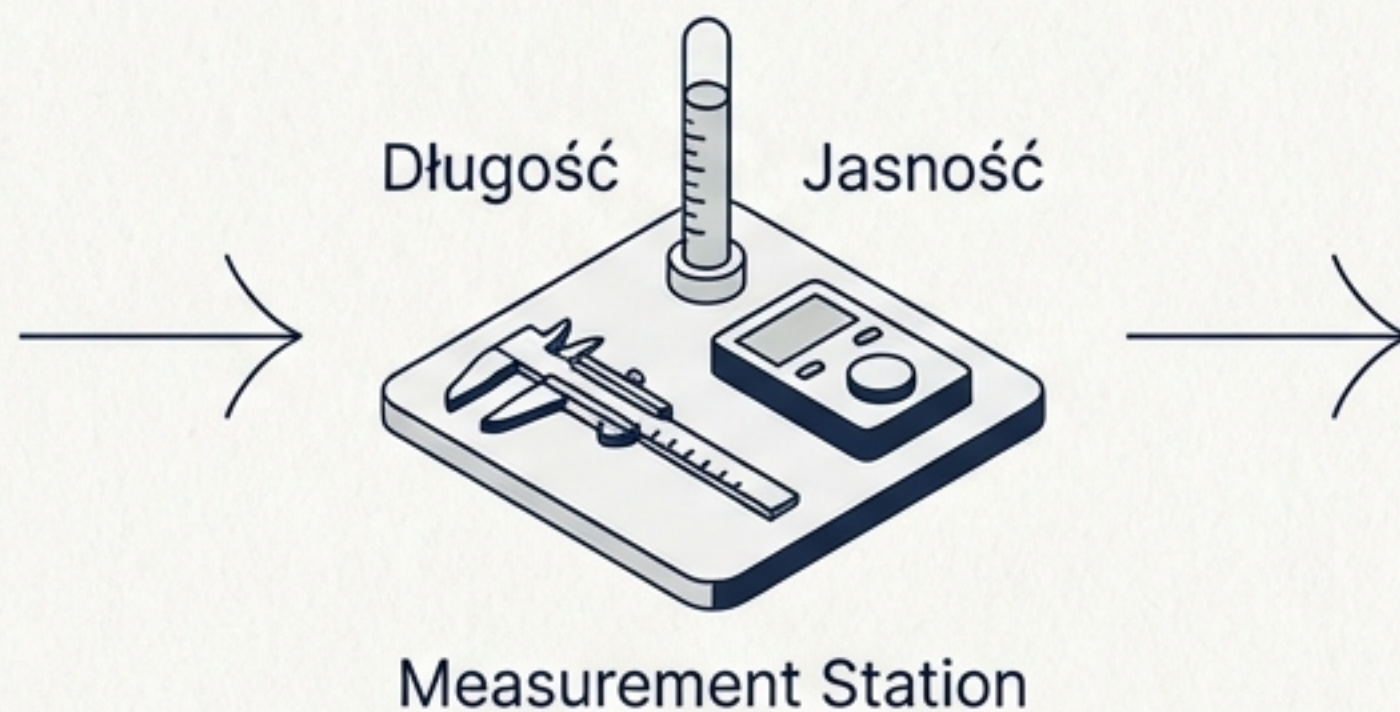
# Paradoks diagnostyczny: Martwić się czy nie?

Założmy, że w Polsce 1 na 1000 osób ma wirusa grypy. Wprowadzono nowy test o dokładności 99%.

Twój wynik jest pozytywny.  
Jakie jest prawdopodobieństwo,  
że jesteś chory?



# Obiekty w przestrzeni cech



**Atrybut (Attribute):**

np. Wiek

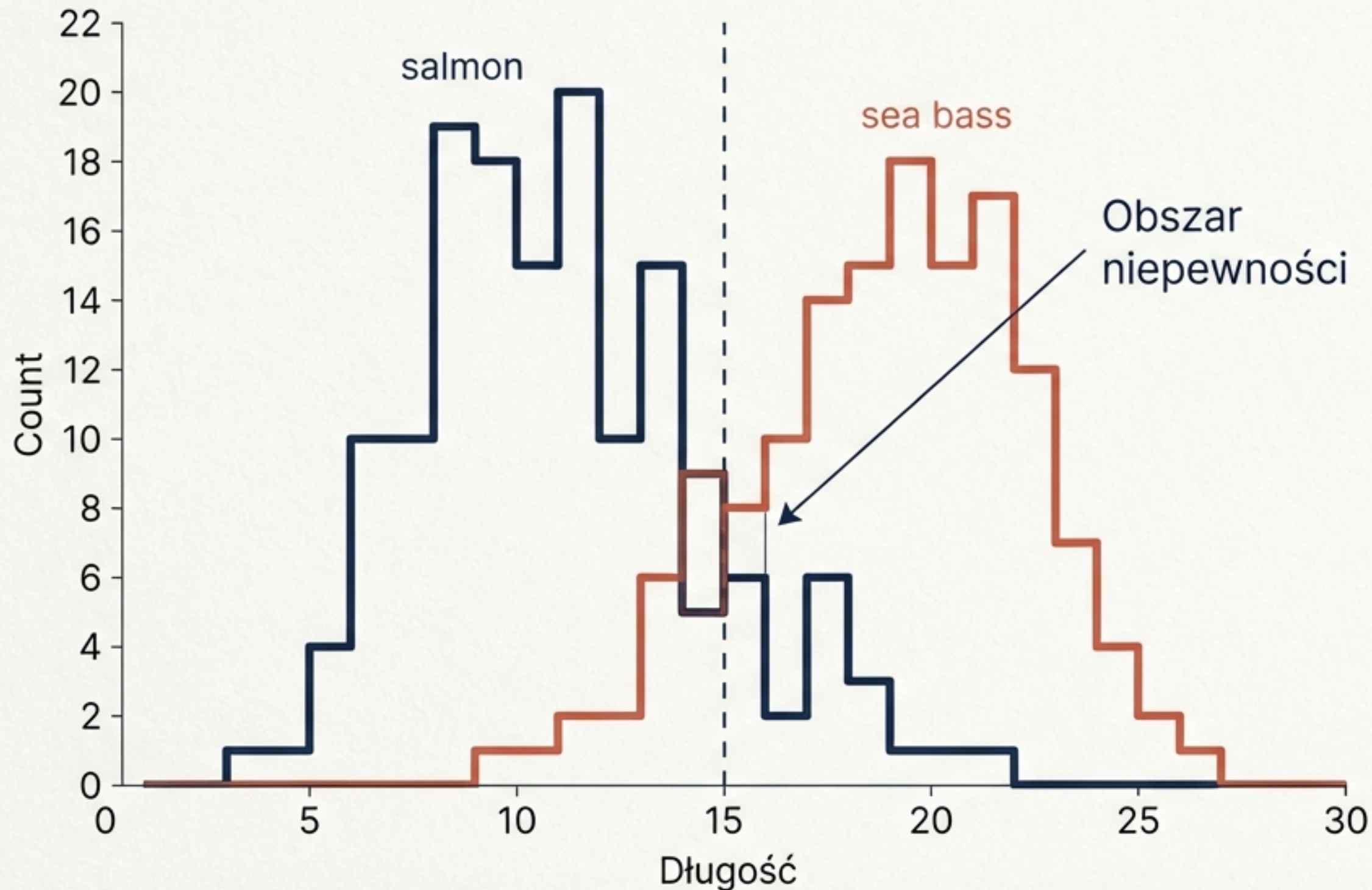
**Cecha (Feature):**

wartość atrybutu, np. Młody

**Wektor cech:**

$X = (x_1, x_2, \dots, x_d)$

# Od zliczeń do prawdopodobieństwa (Histogramy)



1. Wybieramy liczbę przedziałów ( $n = 20$ )

2. Zliczamy wystąpienia  $N(C, r_i)$

3. Prawdopodobieństwo łączne:

$$P(C, r_i) = \frac{N(C, r_i)}{N}$$

# Prawdopodobieństwo łączne a warunkowe

Cecha x

Klasa C

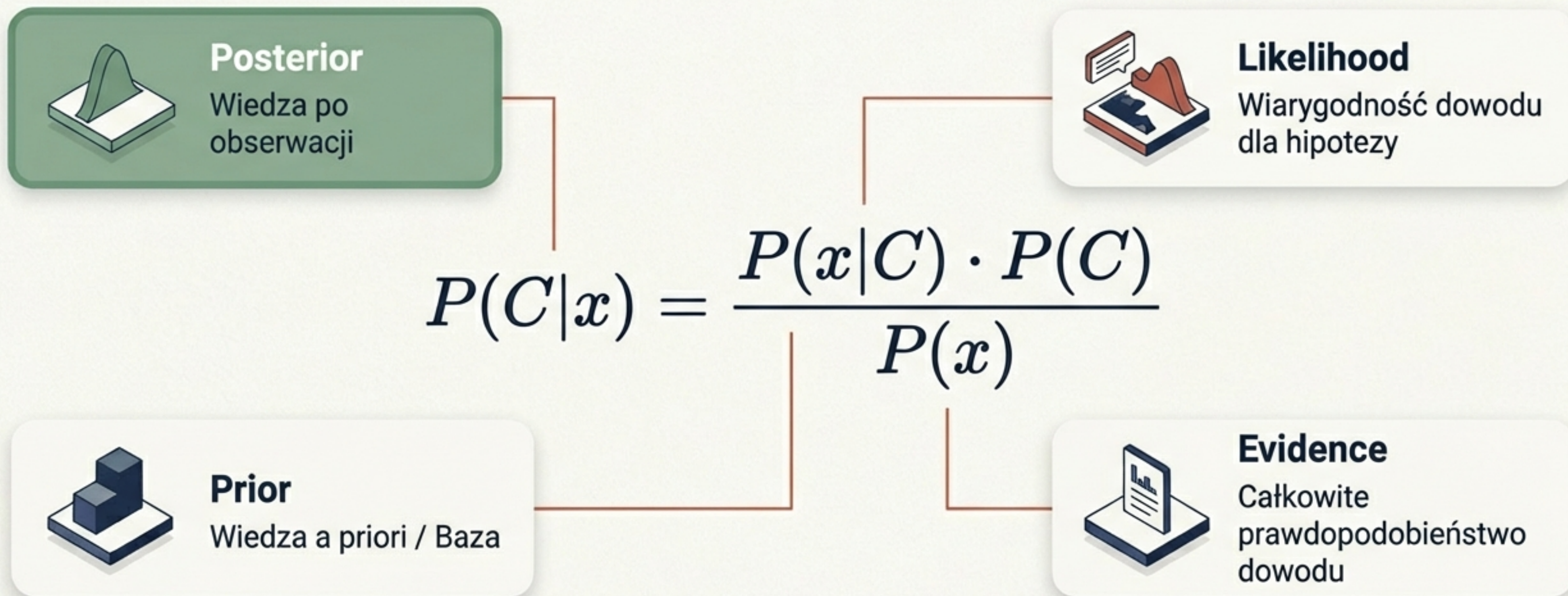

Prawdopodobieństwo dowodu  $P(x)$   
(Suma kolumny)

Prawdopodobieństwo  
*a priori*  $P(C)$   
(Suma rzędu)

**Prawdopodobieństwo warunkowe:**

- $P(x|C)$  - Jak prawdopodobna jest cecha dla tej klasy?  
(Normalizacja wiersza)
- $P(C|x)$  - Jak prawdopodobna jest klasa dla tej cechy?  
(Normalizacja kolumny)

# Twierdzenie Bayesa



Aktualizacja przekonań: Nowa Wiedza = (Wiarygodność × Stara Wiedza) / Dowód

# Rozwiązanie paradoksu: Dlaczego 99% to nie 99%

Dane:

- $P(H+) = 0.001$  (Choroba rzadka)
- $P(T+|H+) = 0.99$  (Test dokładny)

Obliczenie:

$$\begin{aligned} P(H+|T+) &= \frac{P(T+|H+) \cdot P(H+)}{P(T+)} = \frac{0.99 \cdot 0.001}{P(T+|H+)P(H+) + P(T+|H-)P(H-)} \\ &= \frac{0.99 \cdot 0.001}{(0.99 \cdot 0.001) + (0.01 \cdot 0.999)} \\ &= \frac{0.00099}{0.00099 + 0.00999} = \frac{0.00099}{0.01098} \approx 0.09 \end{aligned}$$

Wniosek:

**Tylko 9% szans na chorobę.**

