

|  |   |
|--|---|
| 01 PL Inferential Statistics                                 | 2 |
| 02 PL Inferential Statistics Confidence<br>Intervals         | 3 |
| 03 PL Inferential Statistics Sales Exceeding<br>a Threshold  | 4 |
| 04 PL Inferential Statistics versus Predictive<br>Statistics | 5 |

# STATYSTYKA WNIOSKOWANIA

Tworzenie prognoz dotyczących populacji na podstawie próby

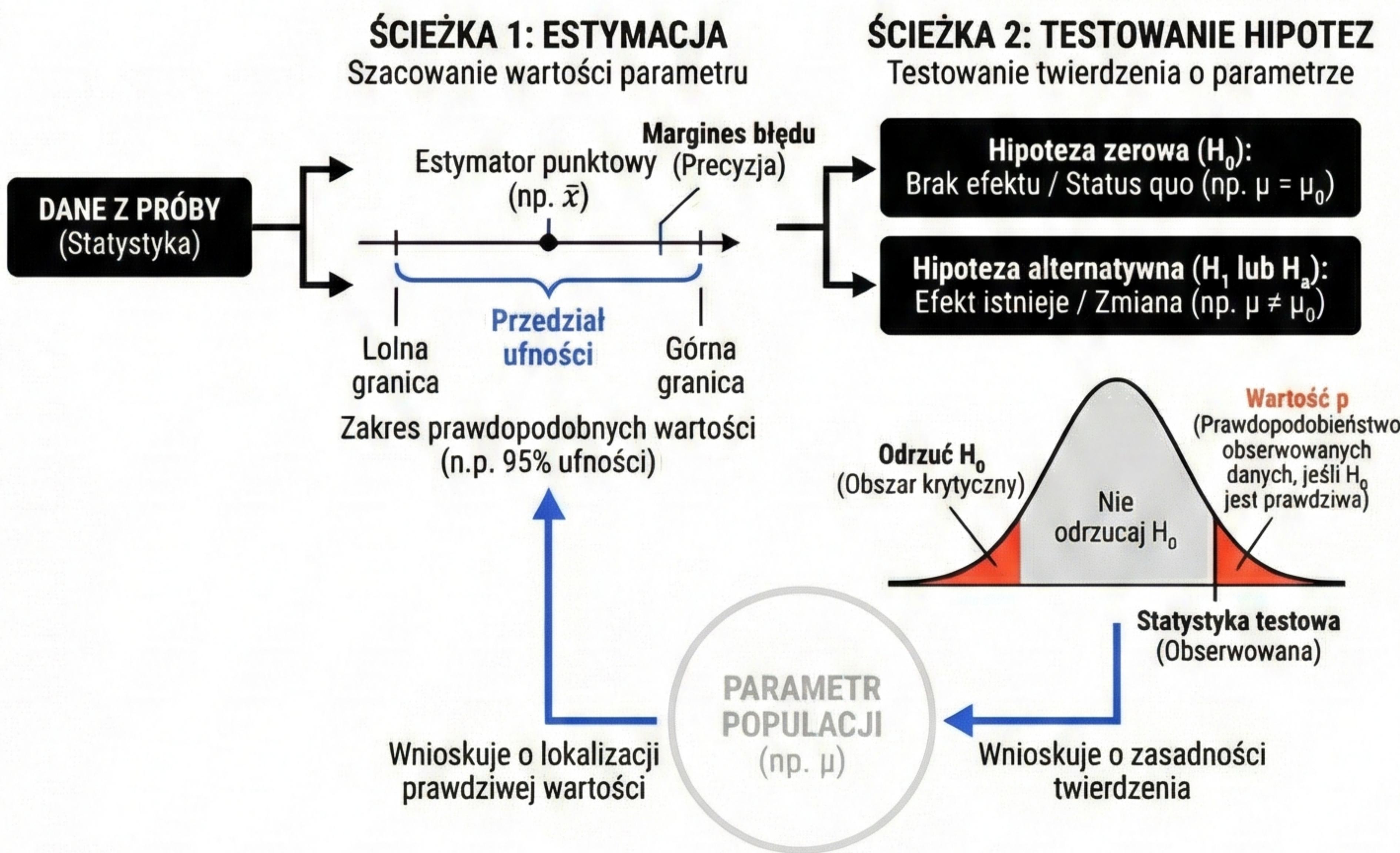
## MODUŁ 1: PODSTAWA (Populacja vs. Próba)

GŁÓWNY PROBLEM: POPULACJA vs. PRÓBA



## MODUŁ 2: PROCES (Estymacja i Testowanie)

PROCES WNIOSKOWANIA: DWIE GŁÓWNE ŚCIEŻKI



## MODUŁ 3: WYNIK (Podejmowanie decyzji)

CEL: ŚWIADOME PODEJMOWANIE DECYZJI

### WYNIK ESTYMACJI

- Zapewnia prawdopodobny zakres dla nieznanego parametru, kwantyfikując niepewność

### WYNIK TESTOWANIA HIPOTEZ

- Dostarcza dowodów na odrzucenie lub nieodrzucenie określonego twierdzenia o parametrze

# PRZEDZIAŁY UFNOŚCI

Szacowanie nieznanego parametru populacji z określonym poziomem ufności

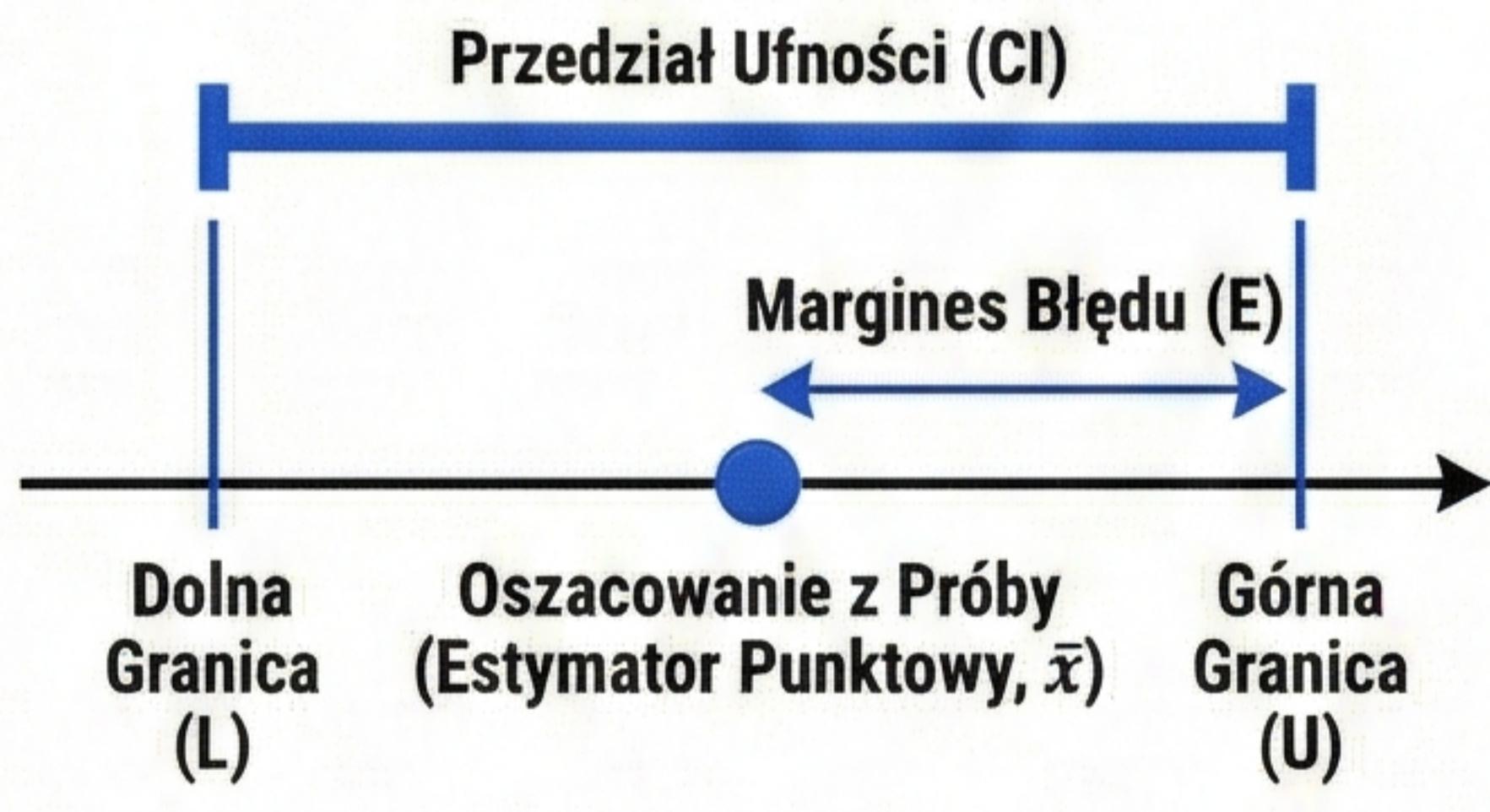
## MODUŁ 1: PROBLEM (Populacja vs. Próba)

### LUKA WNIOSKOWANIA: SZACOWANIE NIEZNAWEGO



## MODUŁ 2: ROZWIAZANIE (Przedział)

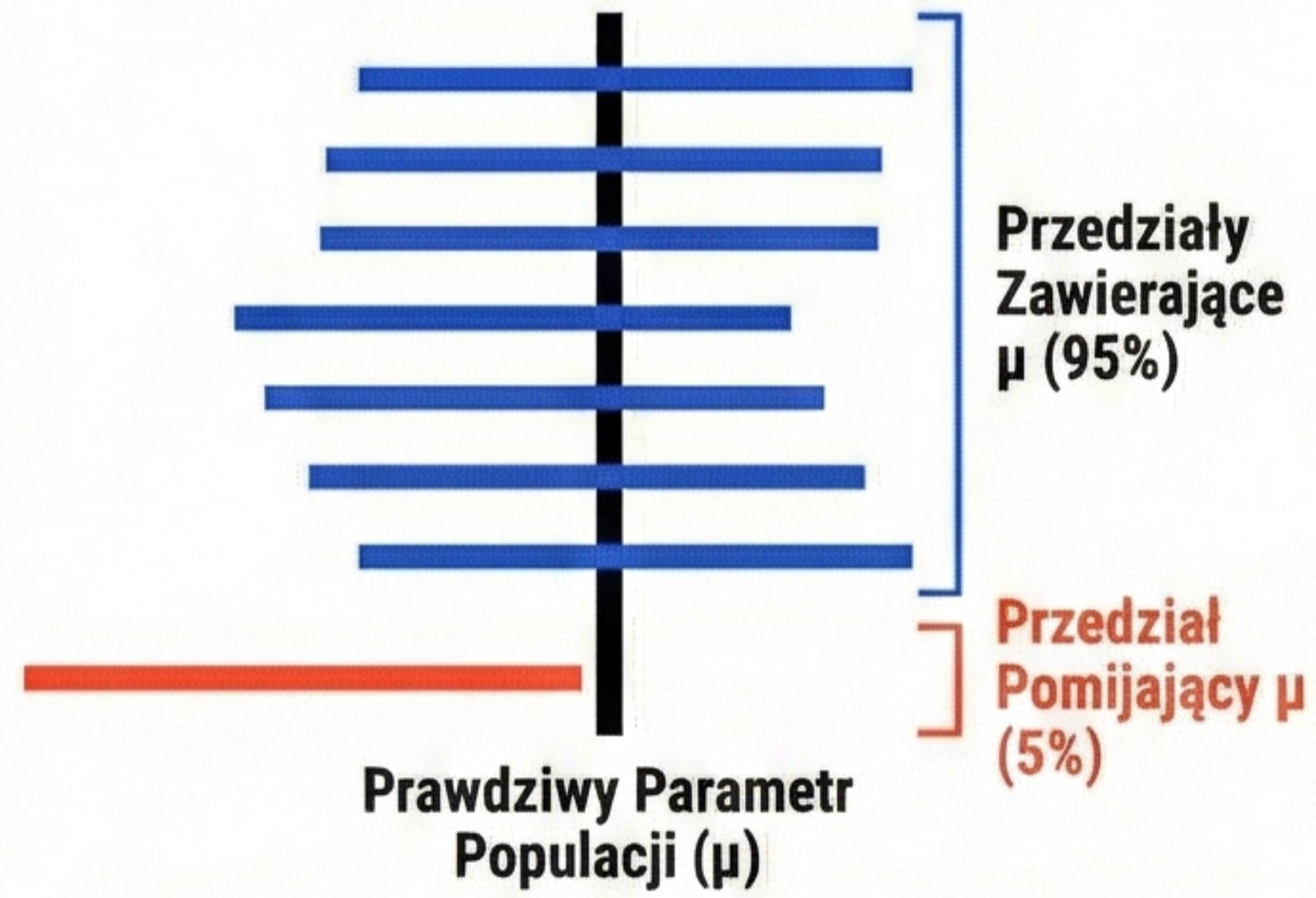
### PRZEDZIAŁ UFNOŚCI: ZAKRES PRAWDOPODOBNYCH WARTOŚCI



$$CI = \text{Estymator Punktowy} \pm \text{Margines Błędu}$$

## MODUŁ 3: INTERPRETACJA (Poziom Ufności)

### INTERPRETACJA UFNOŚCI (np. Poziom 95%)



**Prawidłowa Interpretacja:** Jeśli powtórzylibyśmy proces doboru próby wiele razy, 95% obliczonych przedziałów zawierałoby prawdziwy parametr populacji. To **NIE** jest prawdopodobieństwo, że konkretny przedział zawiera  $\mu$ .

## MODUŁ 4: CZYNNIKI (Czynniki Precyzji)

### CZYNNIKI WPŁYWAJĄCE NA SZEROKOŚĆ PRZEDZIAŁU (PRECYZJĘ)

#### WIELKOŚĆ PRÓBY (n)

Mała Wielkość Próby (n)



Większe n = Węższy Przedział (Bardziej Precyzyjny)

Duża Wielkość Próby (n)

#### ZMIENNOŚĆ ( $\sigma$ lub s)

Niska Zmienność ( $\sigma$  lub s)



Wysoka Zmienność ( $\sigma$  lub s)  
Wyższa Zmienność = Szerszy Przedział (Mniej Precyzyjny)

Wysoka Zmienność ( $\sigma$  lub s)

# TESTOWANIE HIPOTEZ: SPRZEDAŻ PRZEKRACZAJĄCA PRÓG

Ilościowe określanie prawdopodobieństwa, że pojedyncza sprzedaż przekroczy określoną wartość docelową, przy użyciu modelu rozkładu Gamma.

## 1. PYTANIE BADAWCZE I HIPOTEZY

Jakie jest prawdopodobieństwo, że pojedyncza sprzedaż przekroczy próg (np. 250 USD)?

### HIPOTEZA ZEROWA ( $H_0$ : STATUS QUO)

Kwota sprzedaży (X) jest mniejsza lub równa wartości progowej.

$$H_0: X \leq 250 \text{ USD}$$

### HIPOTEZA ALTERNatywna ( $H_1$ : TEST)

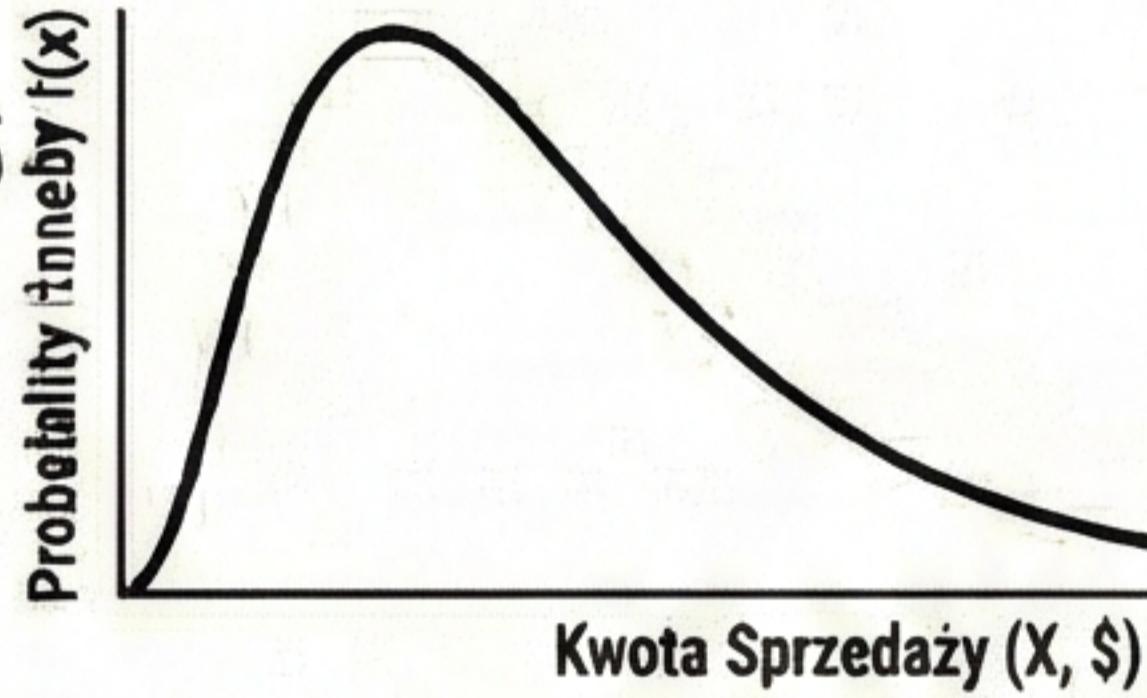
Kwota sprzedaży (X) jest ścisłe większa niż wartość progowa.

$$H_1: X > 250 \text{ USD}$$

## 2. MODEL STATYSTYCZNY (MODEL (ROZKŁAD GAMMA))

SCENARIUSZ: Analityk analizuje dzienne wartości transakcji.

Średnia = 100 USD,  
Odchylenie Standardowe = 5  
(Wariancja = 2500).



### Krok 1: Oblicz Parametry:

$$\hat{\theta} = \frac{\text{Wariancja}}{\text{Średnia}} = \frac{2500}{100} = 25.$$
$$\hat{k} = \frac{\text{Średnia}}{\hat{\theta}} = \frac{100}{25} = 4$$

### PARAMETR KSZTAŁTU ( $\alpha$ )

$$\hat{\alpha} (\hat{k}) = 4$$

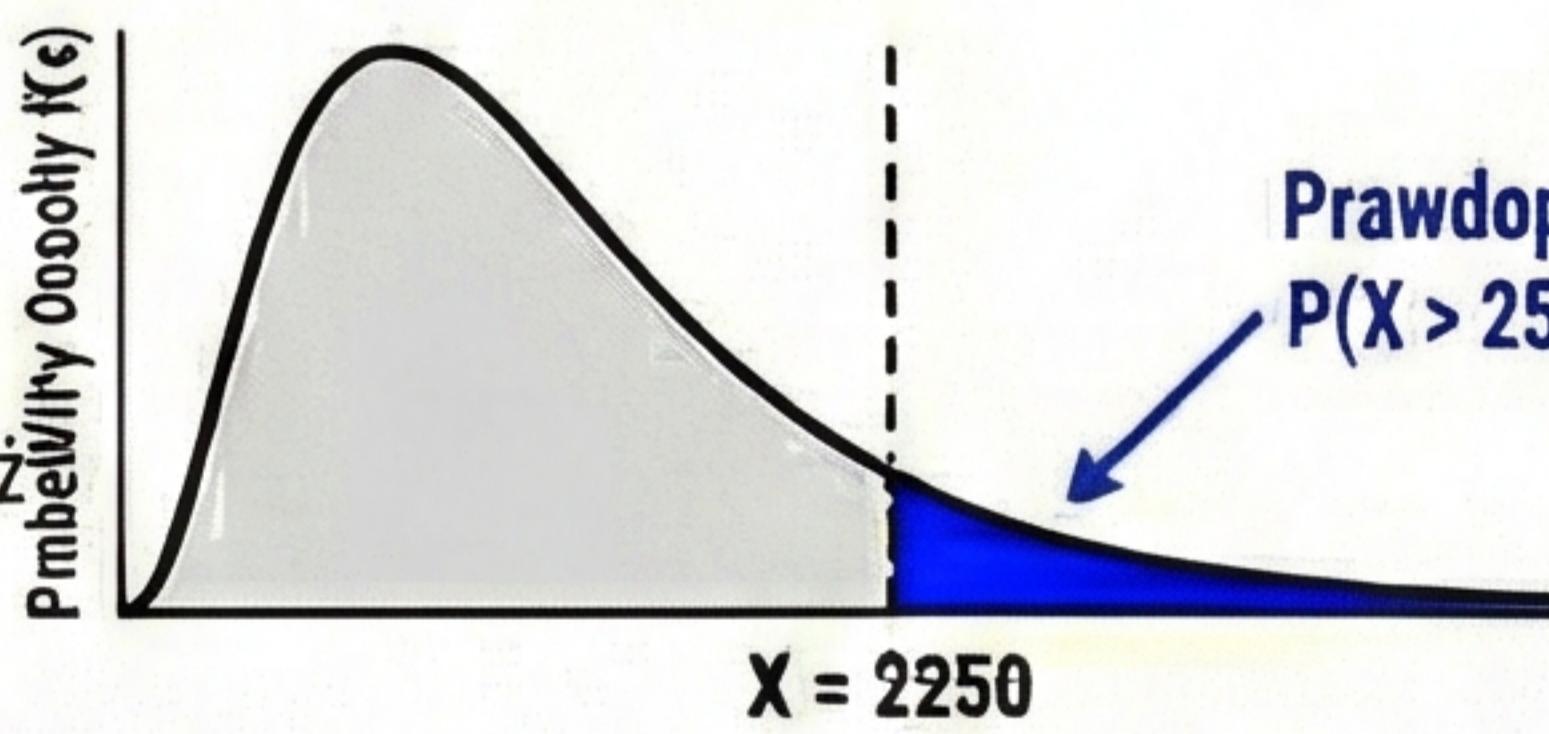
### PARAMETR SKALI ( $\beta$ )

$$\hat{\beta} (\hat{\theta}) = 25$$

## MODUŁ 2A: INTERPRETACJA BIZNESOWA PARAMETRÓW

### INTERPRETACJA BIZNESOWA: KSZTAŁT ROZKŁADU

Ponieważ  $\alpha (k) = 4$ , rozkład jest garbaty i skośny (nie wykładniczy, nie normalny). Sugeruje to zdrową mieszankę niskich i wysokich poziomów wartości, z większością sprzedaży wokół mody, ale znaczącym ogonem większych transakcji.



Prawdopodobieństwo Przekroczenia  
 $P(X > 250 \text{ USD}) = 1 - P(X \leq 250 \text{ USD})$

**WZÓR EXCEL I WYNIK (PRZYKŁAD)**  
 $= 1 - \text{GAMMA.DIST}(250, 4, 25, \text{TRUE})$   
 $\approx 1 - 0.9897 = 0.0103 \text{ (or } 1.03\%)$

Interpretacja: Istnieje 1,03% prawdopodobieństwo, że pojedyncza sprzedaż przekroczy 250 USD, w oparciu o ten model.

## 4. WNIOSKI I REGUŁA DECYZYJNA (KONTEKST ZARZĄDCZY)

Obliczone prawdopodobieństwo (kontekst p-wartości) informuje o decyzjach zarządczych dotyczących sprzedaży o wysokiej wartości.

### NISKIE PRAWDOPODOBIĘSTWO (np. < 5%)

Jeśli prawdopodobieństwo jest niskie, sprzedaż o wysokiej wartości jest rzadka, ale potencjalnie znaczącymi anomaliemi.

Działanie: Zbadaj pod kątem unikalnych czynników lub potencjalnych błędów.



### WYSOKIE PRAWDOPODOBIĘSTWO (np. > 10%)

Jeśli prawdopodobieństwo jest stosunkowo wysokie (jak 1,03%), sprzedaż o wysokiej wartości jest powtarzalną częścią procesu.

Działanie: Włącz do regularnego prognozowania i planowania zasobów.

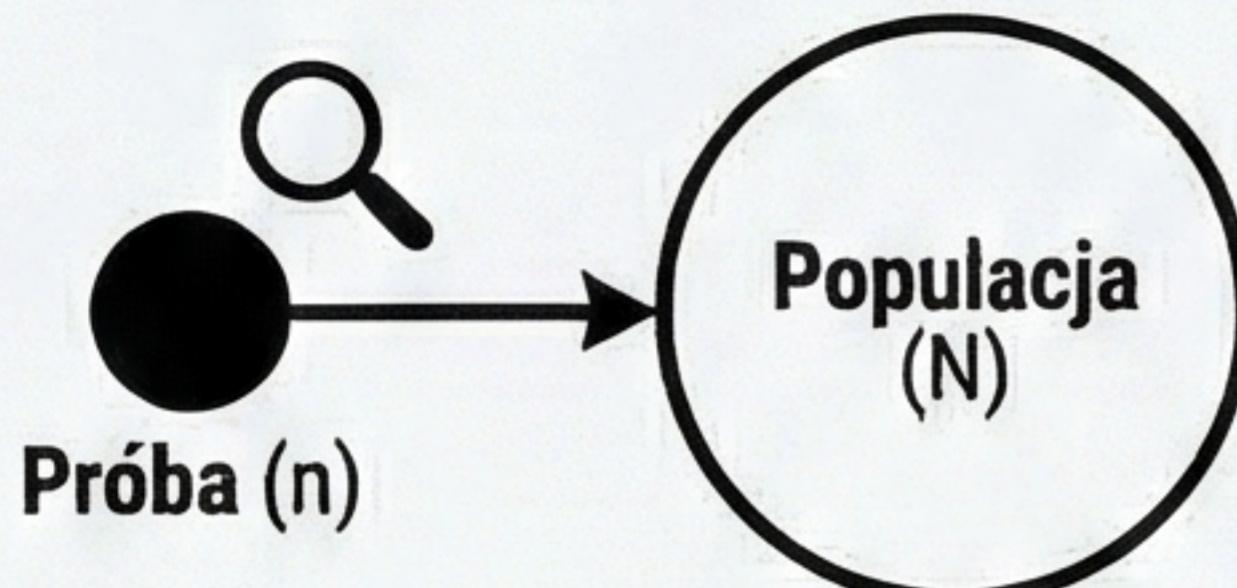


Uwaga: Zawsze weryfikuj założenia dystrybucyjne przed poleganiem na wynikach modelu.

# **STATYSTYKA WNIOSKOWANIA KONTRA PREDYKCYJNA**

# **STATYSTYKA WNIOSKOWANIA**

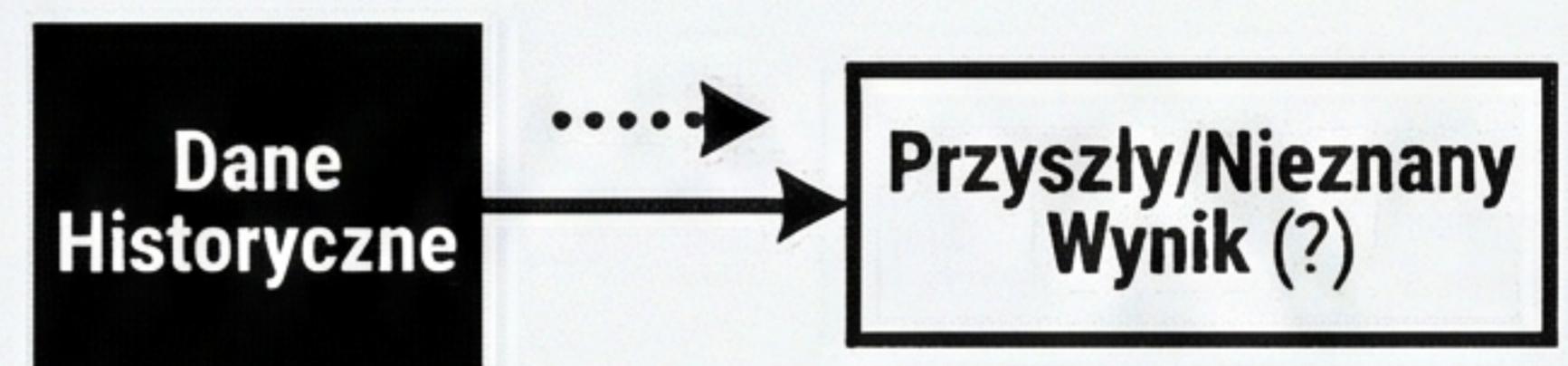
CEL | FOKUS



Zrozumienie parametrów populacji na podstawie danych z próby.  
Skupienie na zależnościach i przyczynowości.

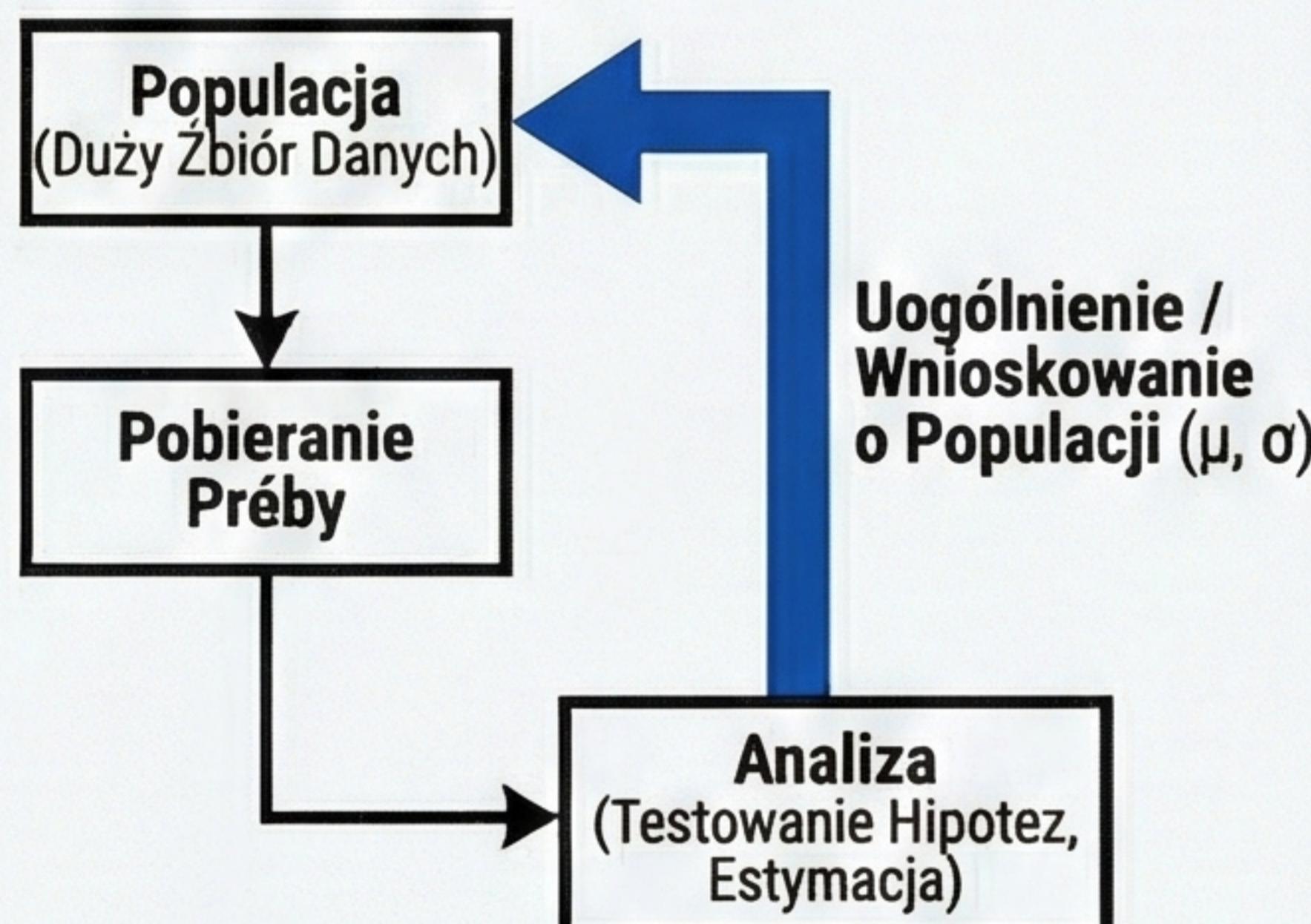
# STATYSTYKA PREDYKCYJNA

- CEL | FOKUS

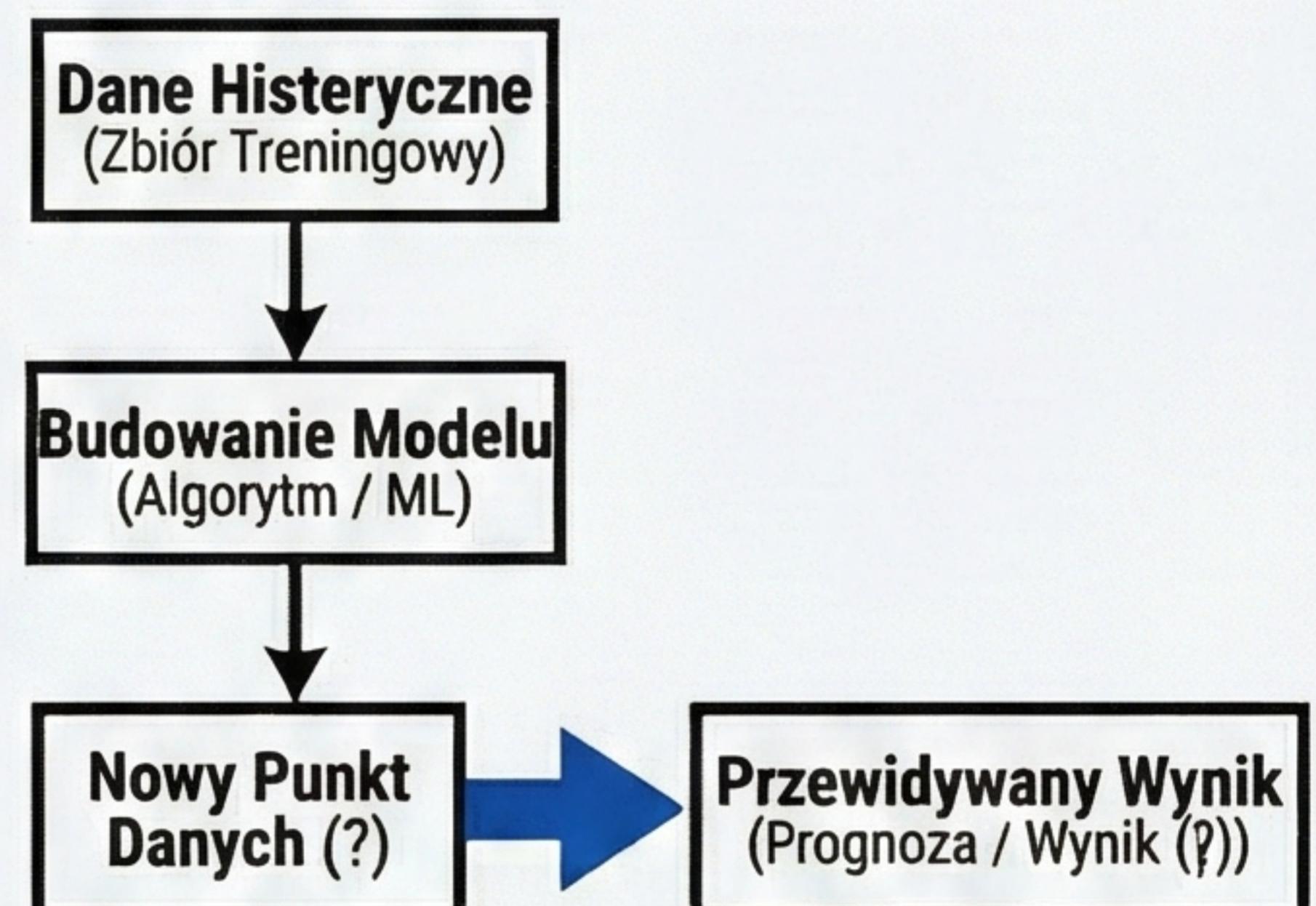


Prognozowanie przyszłych wyników dla nowych obserwacji.  
Skupienie na dokładności pojedynczego punktu danych.

# PRZEPŁYW DANYCH I PROCES



# PRZEPŁYW DANYCH I PROCES



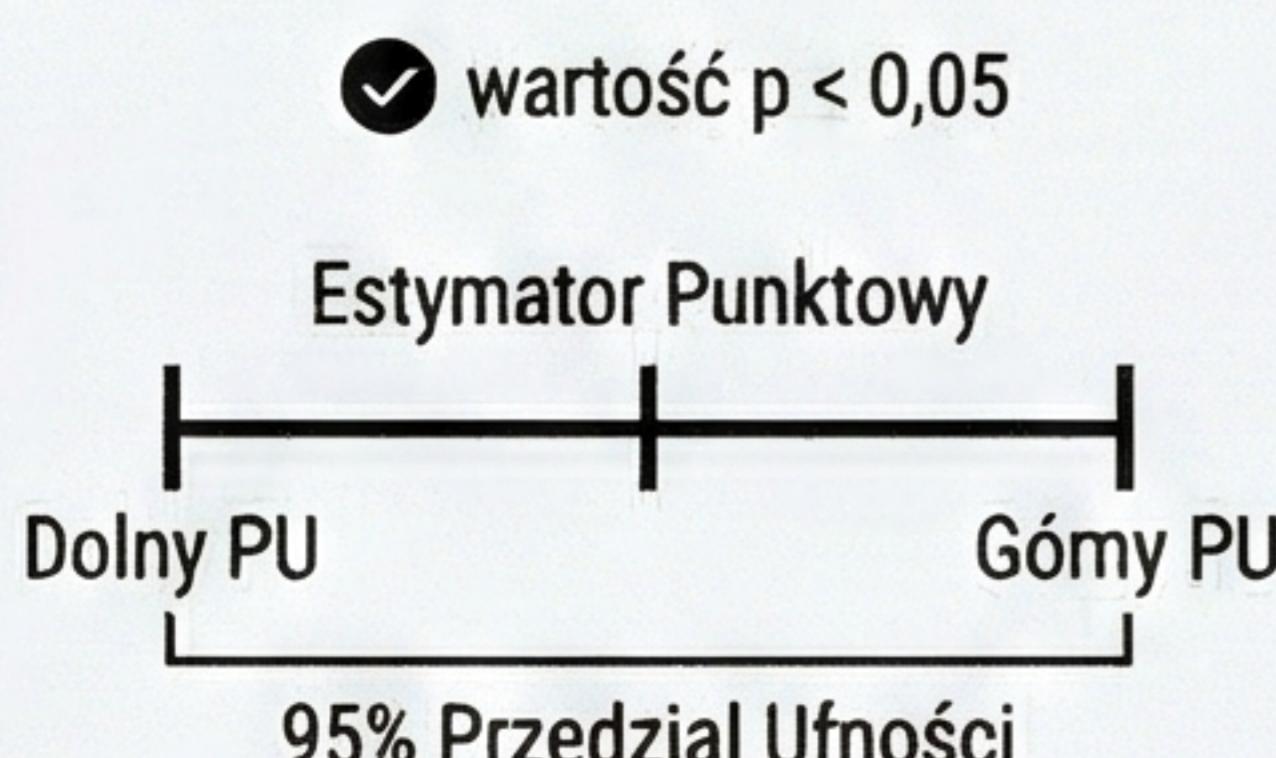
# **KLUCZOWE TECHNIKI/METODY**

- Testy t, ANOVA
  - Test Chi-Kwadrat
  - Przedziały Ufności
  - Regresja Liniowa (dla zależności)

# KLUCZOWE TECHNIKI/METODY

- Regresja Logistyczna (do klasyfikacji)
  - Drzewa Decyzyjne
  - Sieci Neuronowe
  - Prognozowanie Szeregów Czasowych

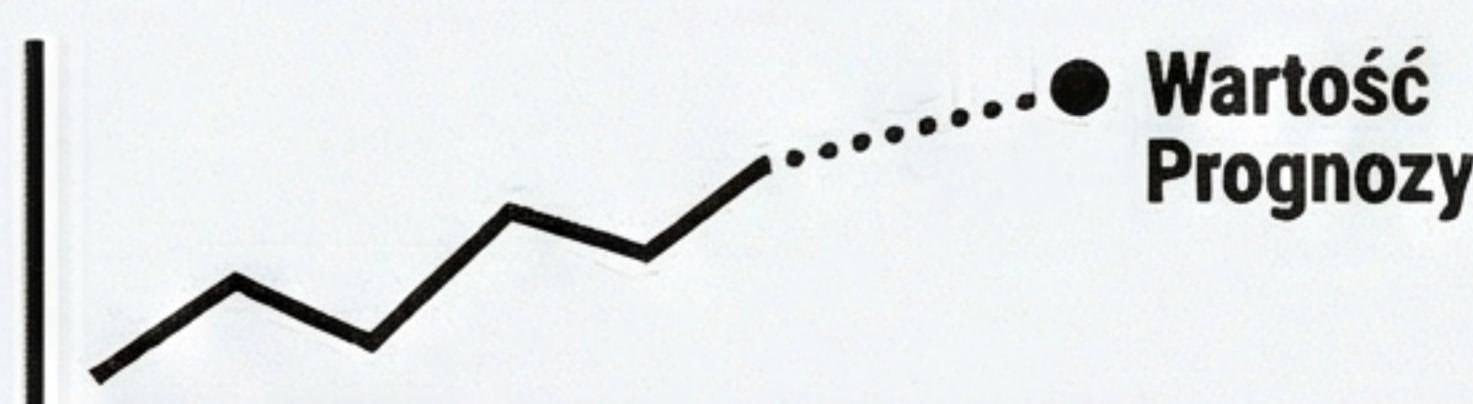
# **WYNIK I INTERPRETACJA**



Szacunki z niepewnością.  
Wyjaśnia **\*dlaczego\*** zjawisko zachodzi.

## WYNIK I INTERPRETACJA

# Macierz Pomyłek



Konkretna wartość lub klasa.  
Metryki: Dokładność, RMSE.  
Skupia się na tym, \*co\* się wydarzy.