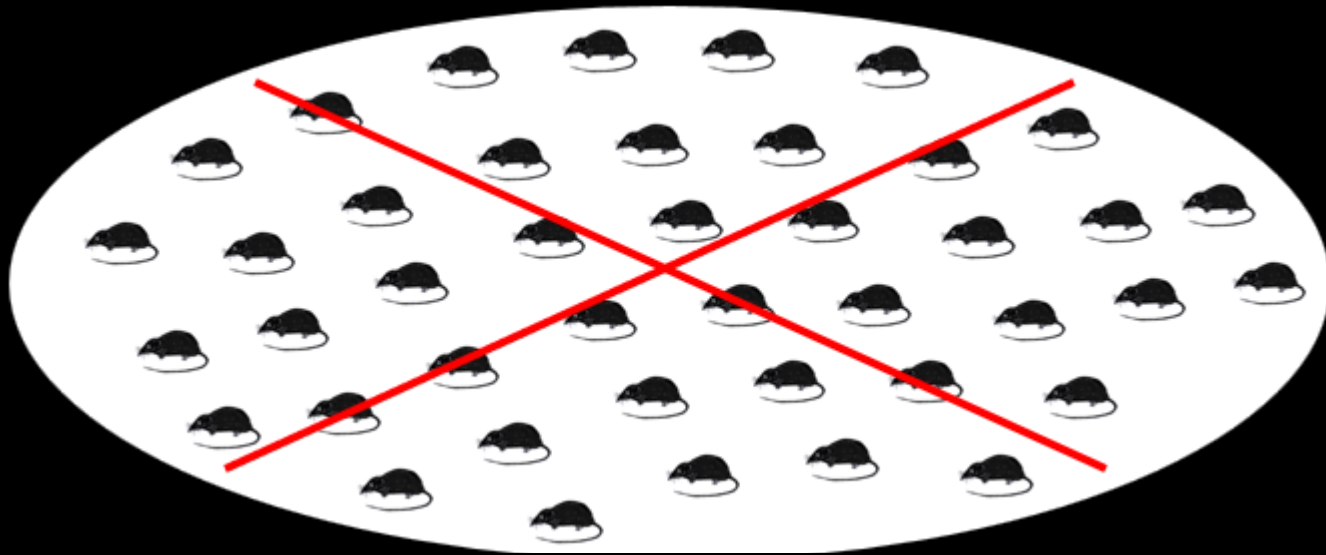




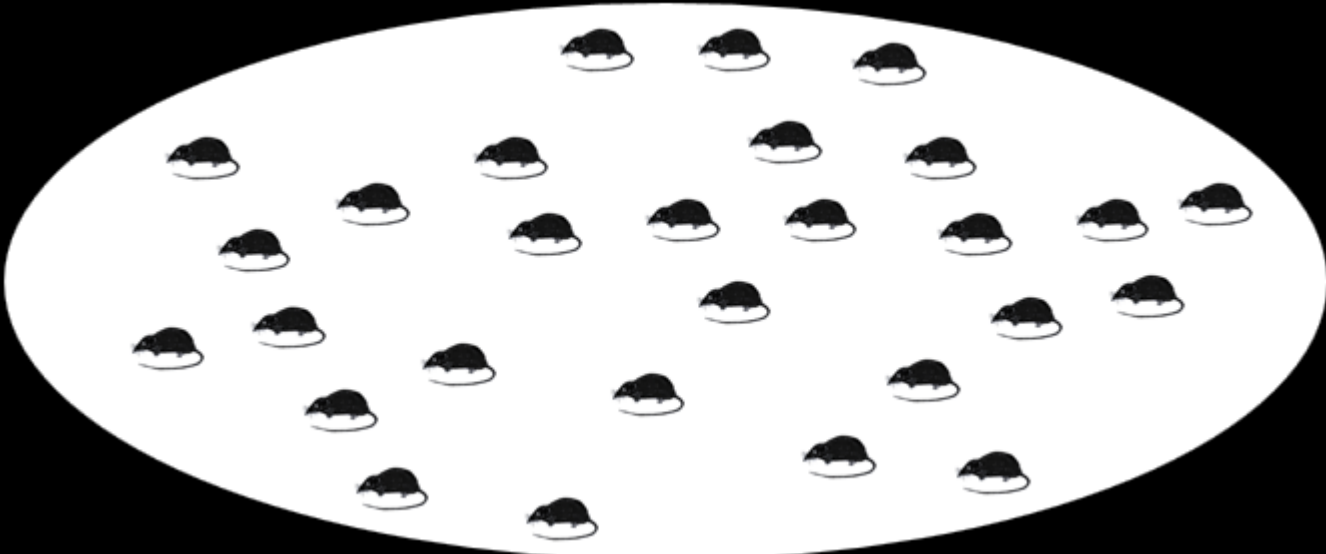
POPULACJE I PRÓBY DANYCH

Dawid Słomian

POPULACJA I PRÓBA DANYCH



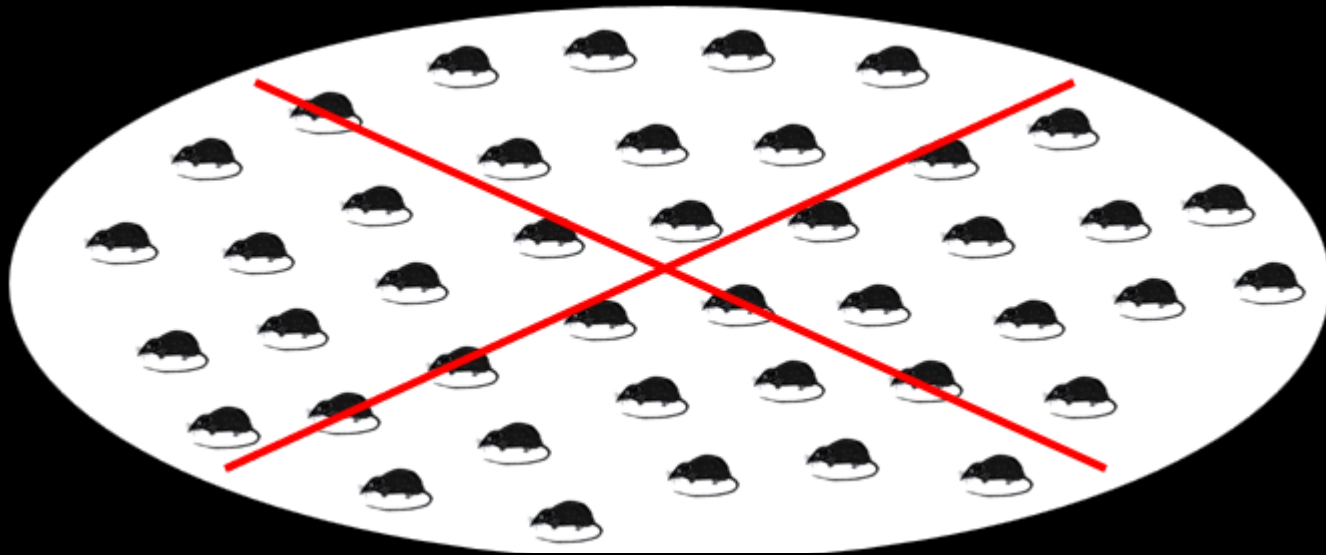
POPULACJA
(w rozumieniu statystycznym)
Obserwacje dla wszystkich
osobników danego gatunku / rasy



PRÓBA DANYCH

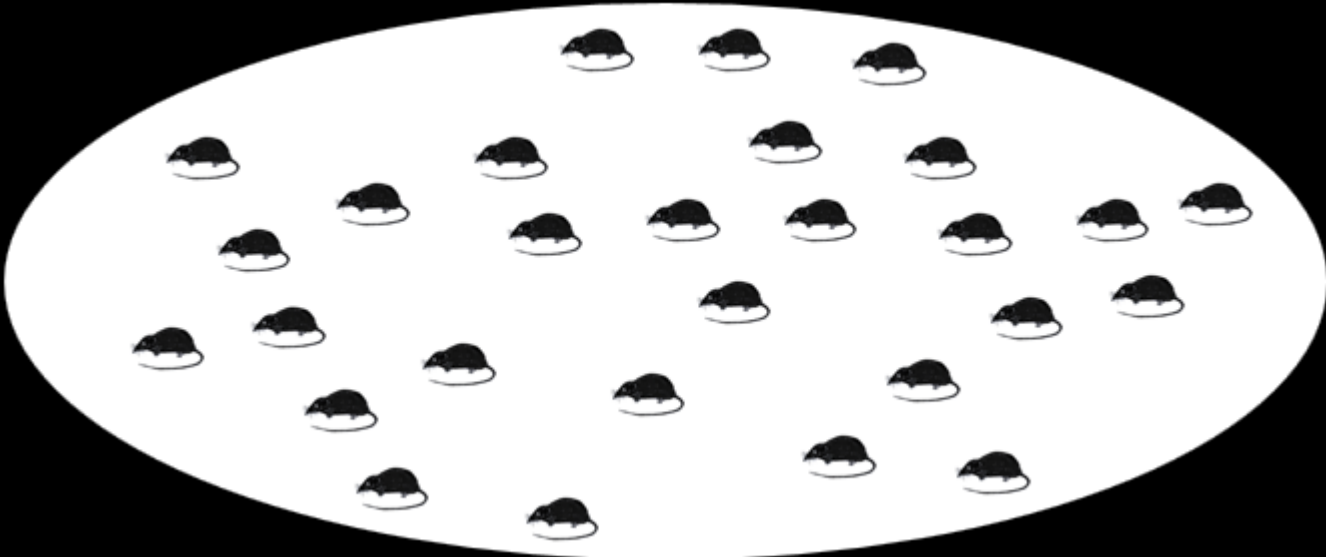
Obserwacje dotyczące wycinka
populacji

POPULACJA I PRÓBA DANYCH



Populacja

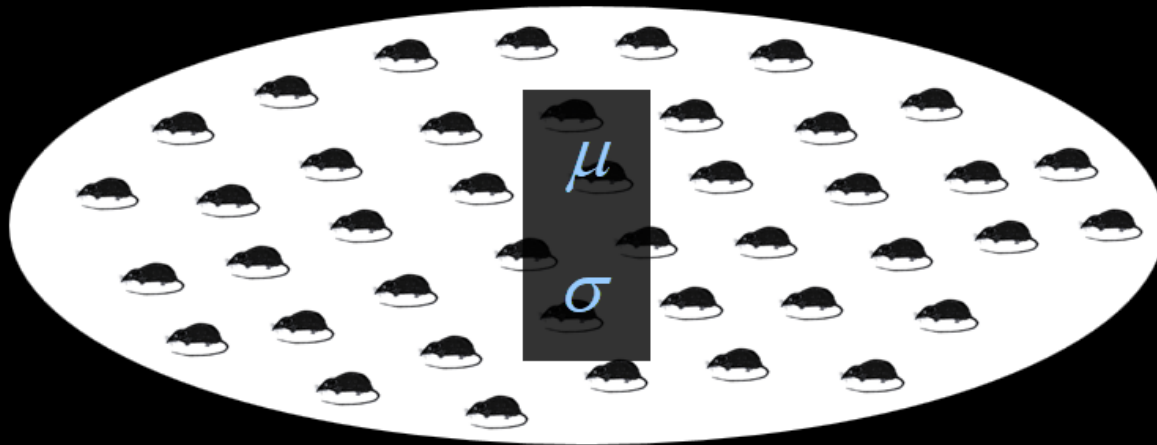
Należy ją zdefiniować przed rozpoczęciem eksperymentu ...



Próba danych

... ponieważ wielkości obliczone na podstawie uzyskanej próbie danych odnoszą się do populacji, z której została pozyskana

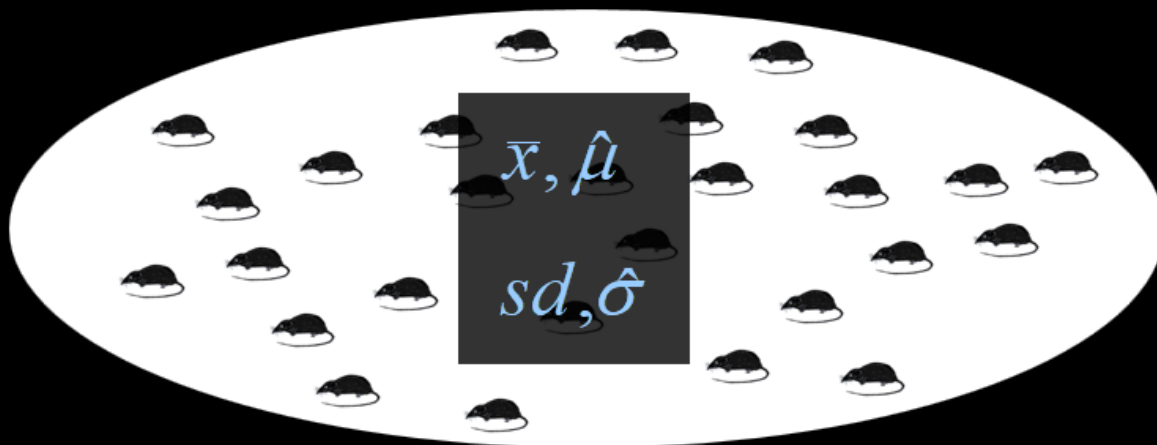
PARAMETR POPULACJI I ESTYMATOR



Populacja → Parametr

- nieznany
- dotyczy całej populacji

np. rzeczywista średnia masa ciała i jej odchylenie standardowe w populacji ryjówek

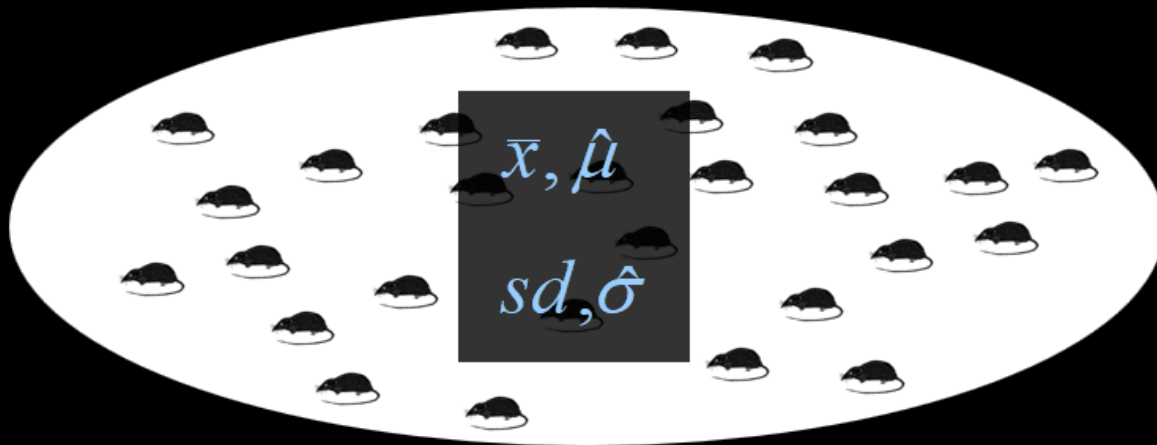
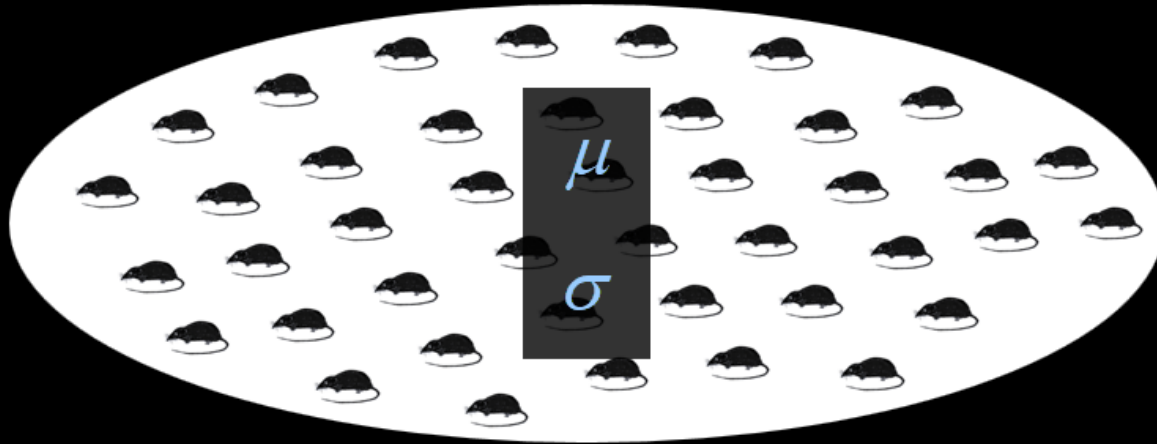


Próba danych → Estymator

- znany
- wartość obliczona dla danej próby

np. obliczona średnia masa ciała i jej odchylenie standardowe w próbie danych

PARAMETR POPULACJI I ESTYMATOR



estymator

... powinien być

- nieobciążony $E(\text{estymator}) = \text{parametr}$

- zgodny

ze wzrostem N estymator $\rightarrow$ parametr

- efektywny

najmniejsza wariancja

najlepszy estymator

LICZEBNOŚĆ PRÓBY DANYCH

populacja	próba 1	próba 2	próba 3	próba 4		średni wzrost		różnica
156	20	50	100	130		w. prawdziw	173.04	
178						próba 1	20.00	-153.04
179						próba 2	50.00	-123.04
156						próba 3	100.00	-73.04
155						próba 4	130.00	-43.04
189								
178								
167								
154								
178								
179								
190								
178								
176								
189								
180								
184								
160								
170								
175								
176								

LICZEBNOŚĆ PRÓBY DANYCH

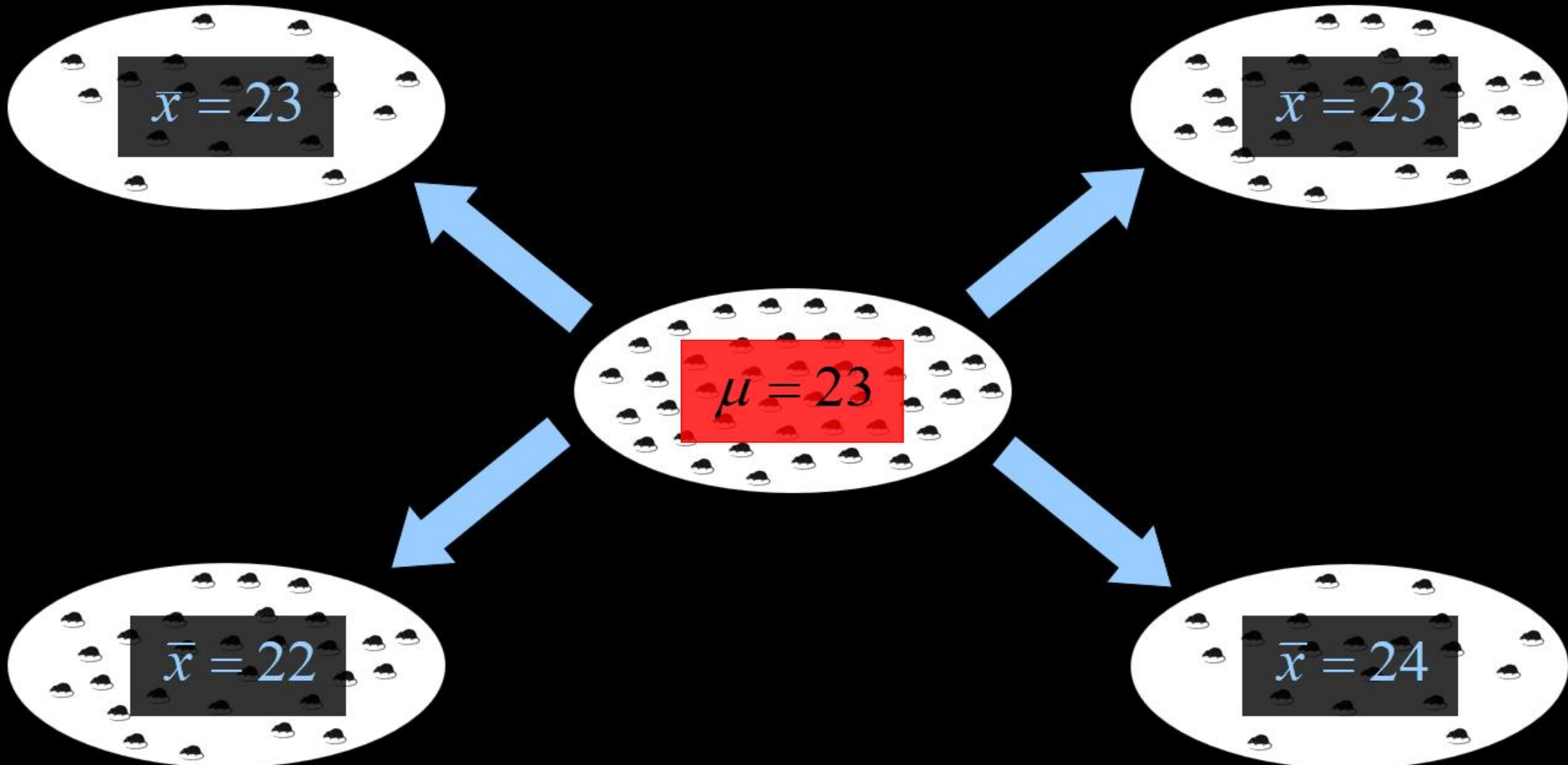
- Liczba obserwacji w próbie
- Im większa próba danych tym dokładniej można oszacować wartości parametrów populacji
- Wielkość próby danych zależy od: łatwości zdobywania obserwacji, kosztów, dostępności obserwacji
- Czynniki wpływające na wymaganą wielkość próby danych – zmienność obserwacji = odchylenie standardowe, dokładność pomiarów, hipoteza badawcza

LICZEBNOŚĆ PRÓBY DANYCH

Analiza mocy - technika statystyczna pozwalająca na określenie minimalnej wielkości próby danych dla danego eksperymentu

- np. moc 80% prawdopodobieństwo przyjęcia prawdziwej hipotezy alternatywnej np. wykrycie istniejącej różnicy = 0.8
- metody empiryczne: symulacje Monte Carlo
- metody numeryczne

REPLIKACJE



REPLIKACJE

- Ten sam rodzaj pomiaru / eksperymentu wykonany na różnych próbach danych
- Stosowanie replikacji znacznie zwiększa wiarygodność wyciągniętych wniosków
- Wybór osobników do każdej replikacji powinien być losowy
- Idealnie → każda próba danych powinna zawierać różne osobniki
- Jeżeli technicznie trudno pobrać osobniki do replikacji i niektóre powtarzają się w poszczególnych próbach danych → pseudoreplikacje

TWORZENIE PRÓBY DANYCH



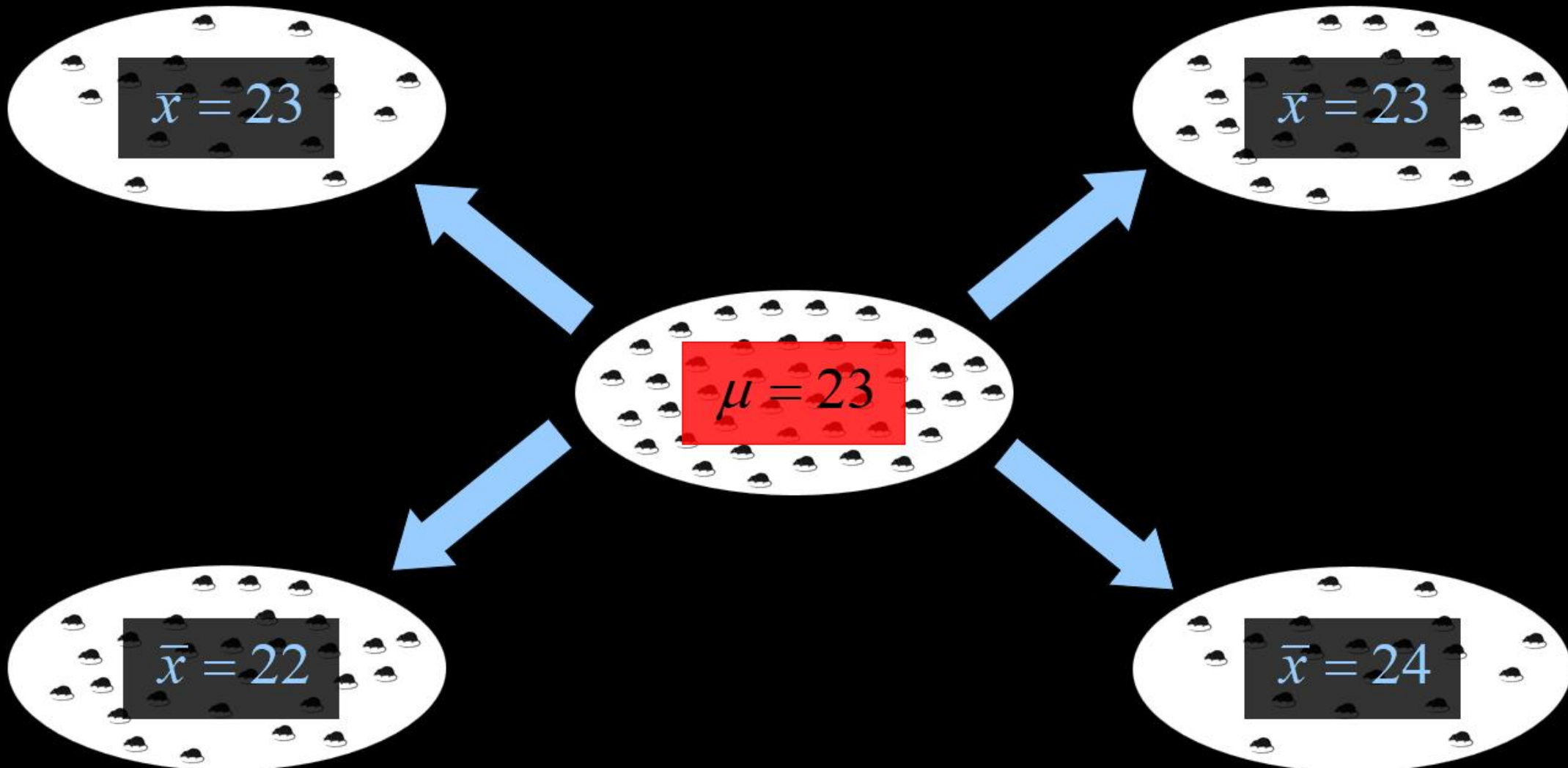
losowy

- Osobniki wybierane są bez stosowania selekcji
- Osobniki muszą reprezentować wszystkie warianty efektów działających na populację np. lokalizacja, wiek, płeć, ekspozycja na warunki środowiskowe, spokrewnienie

nielosowy

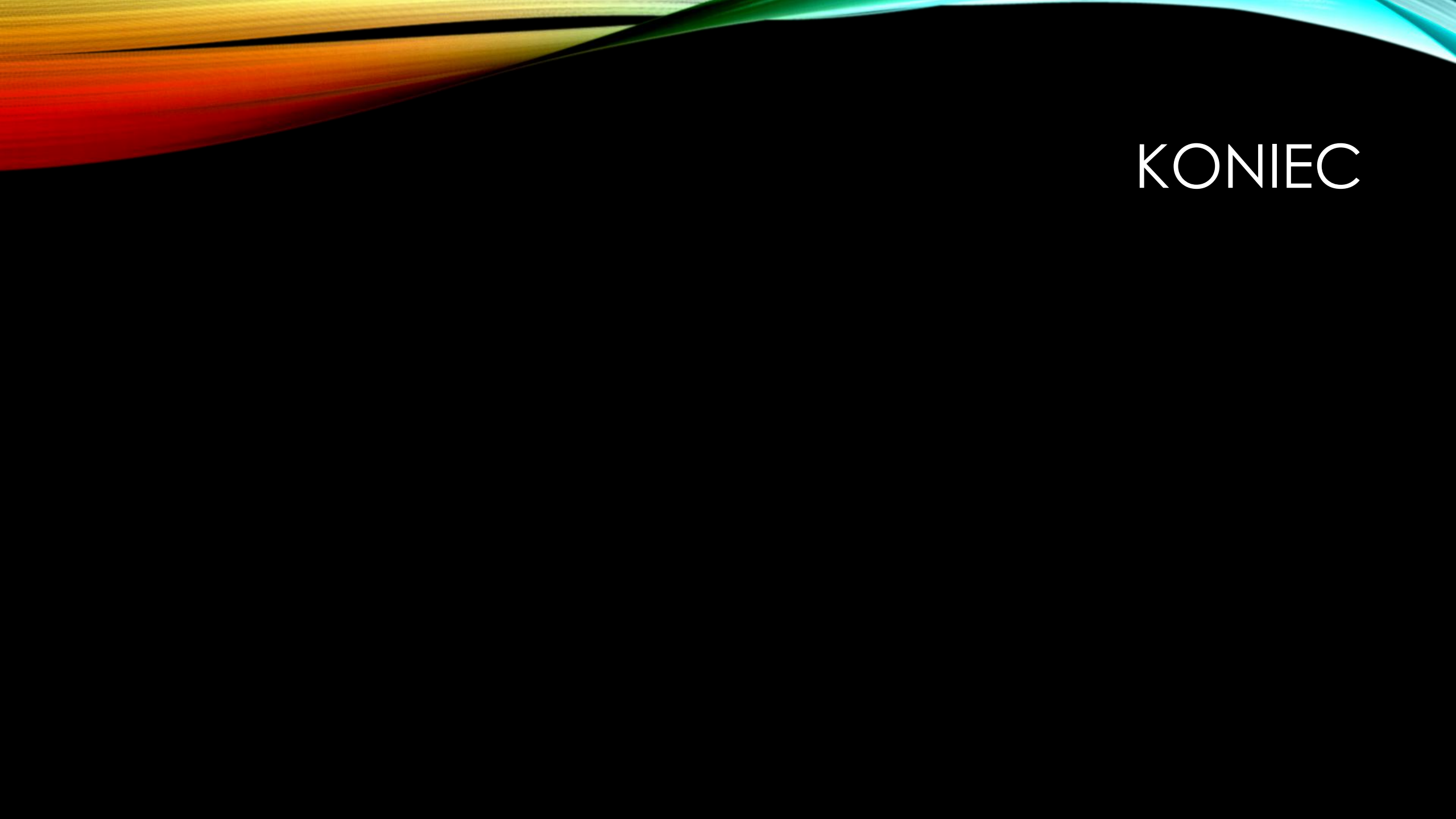
- Wybór wg spokrewnienia
- Wybór wg wartości cechy
- Eksperyment

BŁĄD PRÓBKOWANIA



BŁĄD PRÓBKOWANIA

- Błąd próbkowania / błąd estymacji wynika z faktu, że obserwujemy próbę danych zamiast populacji
- Nie możemy precyzyjnie określić wielkości błędu próbkowania
- Możemy w przybliżeniu określić dokładność estymatora
 - Błąd standardowy estymatora
 - Przedziały ufności estymatora



KONIEC