Spis treści

[Statystyka opisowa 1](#_Toc106131034)

[Definicje i notacja: 1](#_Toc106131035)

[Miary tendencji centralnej: 1](#_Toc106131036)

[Rozrzut / dyspersja: 1](#_Toc106131037)

# Statystyka opisowa

## Definicje i notacja:

Populacja – zbiór elementów, np. ludzi lub gwoździ. Chcemy wypowiedzieć się na temat populacji.

Próba – podzbiór populacji, to właśnie ten zbiór badamy.

Cecha lub zmienna – cecha populacji, np. wzrost ludzi lub długości gwoździ, (oznaczana przez X)

Rozkład populacji – rozkład cechy X na elementach populacji, np.

Rozkład normalny, rozkład wykładniczy, rozkład Poissona, rozkład dwumianowy itd.

Dane – zaobserwowane wartości zmiennej X na elementach próby

Rozkład próby (empiryczny) – rozkład zmiennej X na elementach próby

Notacja:

μ – średnia populacji

σ2 – wariancja populacji

x̄ - średnia próby

s2 – wariancja próby

## Rodzaje zmiennych:

1. Jakościowa, kategoryczna: istnieje tylko kilka możliwych wartości zmiennej. Cechy te mogą być mierzone w skali:
   1. Nominalnej:
      1. nie ma porządku
      2. nie można wykonywać operacji arytmetycznych
      3. np. kolor oczu, płeć
   2. Porządkowej:
      1. jest porządek
      2. nie można wykonywać operacji arytmetycznych
      3. np. wykształcenie, kolejność zawodników na podium
2. Ilościowa:
   1. Dyskretna:
      1. istnieje przeliczalna ilość wartości zmiennej.
      2. Wartości mogą się powtarzać.
      3. Np. liczba zgłoszeń w centrali telefonicznej.
   2. Ciągła:
      1. Istnieje nieskończona, nieprzeliczalna ilość wartości zmiennej.
      2. W próbie wartości zmiennej nie powinny się powtarzać.
      3. Np. czas, waga, wysokość, zarobki.

## Miary tendencji centralnej:

Średnia arytmetyczna -

Mediana – zwykle używamy jej gdy dane są „zepsute” przez duże wartości odbiegające od normy, np. wzrost koszykarza w zbiorze Japończyków

Kwantyl to punkt w wybranym miejscu uporządkowanej listy obserwacji, np.

Kwantyl Q1 – punkt pomiędzy 25% obserwacji a 100%

Kwantyl Q2 (mediana) – punkt pomiędzy 50% obserwacji a 100%

Kwantyl Q3 – punkt pomiędzy 75% obserwacji a 100%

Kwantyle mogą mieć dowolną wartość, np. kwantyl 1/20

## Rozrzut / dyspersja / miara zróżnicowania rozkładu:

Odchylenie standardowe – jak szeroko wartości jakiejś wielkości (np. wieku, inflacji, kursu walutowego) są rozrzucone wokół jej średniej. Liczymy je poprzez policzenie pierwiastka z wariancji

Wariancja

Współczynnik zmienności

## Reguła trzech sigm