

# Podstawy statystyki praktycznej

## Lista 2

### Laboratoria

1. Korzystając ponownie ze zbioru danych `grades.txt` ustal czy rozkład wyników testu IQ jest normalny:
  - (a) sprawdzając ile obserwacji mieści się w przedziałach  $[\bar{x} - s, \bar{x} + s]$ ,  $[\bar{x} - 2s, \bar{x} + 2s]$ ,  $[\bar{x} - 3s, \bar{x} + 3s]$ , gdzie  $\bar{x}$  jest średnią obserwacji, a  $s$  odchyleniem standardowym;
  - (b) rysując wykres kwantylowy (QQ-plots)
2. Korzystając ponownie ze zbioru danych `income.txt` ustal rysując wykres kwantylowy (QQ-plots) czy rozkład zarobków jest normalny. Jeśli nie, sprawdź czy dla pewnej transformacji zmiennej (np.  $\sqrt{X}$ ,  $\ln X$ ) otrzymamy rozkład normalny.
3. Powtórz dziesięciokrotnie następujący eksperyment w odniesieniu do zbioru danych `income.txt`:
  - (a) Pobierz 1000-elementową próbę losową z tego zbioru. Wykreśl trajektorie statystyk: średnia, mediana, odchylenie standardowe, IQR oraz procent osób mających co najmniej licencjat, liczone w oparciu o  $t$  pierwszych elementów z tej próby ( $t \in [1, 1000]$ ). Porównaj z wartościami tych statystyka dla całego zbioru danych.
  - (b) Pobierz dziesięć 100-elementowych grup z tego zbioru, stosując próbkowanie warstwowe ze względu na sektor zatrudnienia (tzn. liczba osób wylosowanych z danego sektora odpowiada procentowi osób z tego sektora w całym zbiorze danych). Następnie wykreśl trajektorie odpowiednich statystyk stopniowo zwiększając rozmiar próby o kolejne wylosowane grupy. Porównaj z wartościami odpowiednich statystyk w całym zbiorze danych.
4. Zasymuluj dziesięciokrotnie 1000 rzutów monetą. Dla każdego z 10 eksperymentów wykreśl trajektorie liczby frakcji (w procentach) orłów uzyskanych w kolejnych  $k$  rzutach ( $k \in [1, 1000]$ ).