

Podstawy statystyki praktycznej

laboratorium – lista 3

1. Rozważamy standardowy rozkład normalny $N(0, 1)$.
 - (a) Wygeneruj 100-elementową próbę z tego rozkładu i skonstruuj przedział ufności dla wartości oczekiwanej na poziomie ufności 95%.
 - (b) Doświadczenie z punktu (a) powtórz 1000 razy i oblicz jak często tak skonstruowane przedziały ufności zawierają rzeczywistą wartość oczekiwaną.
 - (c) Powtórz doświadczenie z punktu (a) dla próby 200-elementowej. Wyznacz prawdopodobieństwa pokrycia rzeczywistej wartości oczekiwanej i porównaj średnią szerokość przedziałów ufności wyznaczonych w oparciu o próby 100- i 200-elementowe.
2. Niech D będzie zmienną losową oznaczającą dochód dla populacji zawartej w zbiorze danych `income.dat`. Narysuj histogram zmiennej D i zbadaj jej normalność zarówno przy pomocy reguły 68%–95%–99.7%, jak i przy pomocy wykresu kwantylowego.
3. Niech D będzie zmienną losową oznaczającą dochód dla populacji zawartej w zbiorze danych `income.dat`.
 - (a) Skonstruuj nową zmienną $U = \sqrt{D}$ (zastanów się jak poradzić sobie z ujemnymi wartościami D). Narysuj histogram tej zmiennej i wyznacz średnią arytmetyczną μ_U tej zmiennej dla całego zbioru danych. Porównaj $(\mu_U)^2$ z μ_D . Wyznacz frakcję osób z wyższym wykształceniem p_W .
 - (b) Pobierz 200-elementową próbę losową z tego zbioru danych i w oparciu o tę próbę wyznacz estymatory dla μ_U , μ_D oraz p_W . Skonstruuj 95% przedziały ufności dla tych parametrów i sprawdź, czy zawierają one rzeczywiste wartości.
 - (c) Punkt (b) powtórz 200 razy i narysuj histogramy rozkładów powyższych estymatorów oraz wyznacz jak często przedziały ufności zawierały rzeczywistą wartość estymowanego parametru.
4. Korzystając ze zbioru danych `grades.txt`, zakładając, że zbiór danych jest prostą próbą losową z pewnej populacji, skonstruuj przedziały ufności dla średniego ilorazu inteligencji i średniego wyniku testu psychologicznego w tej populacji, przy $\alpha = 0.05$.