

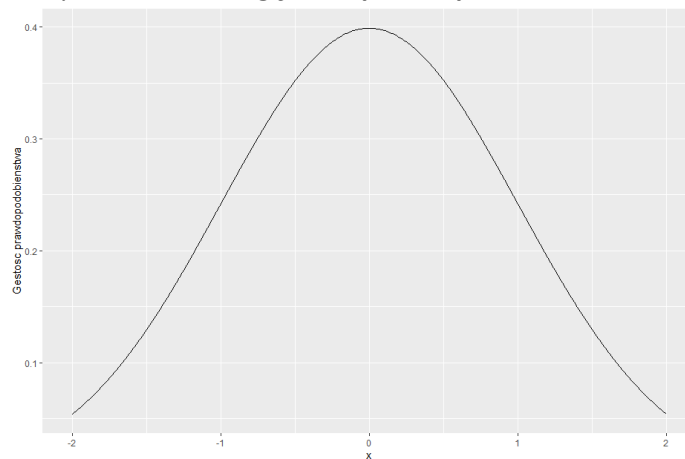
Analiza statystyczna w środowisku R. Eksploracyjna analiza danych nieprzestrzennych.

Odpowiedz na pytania:

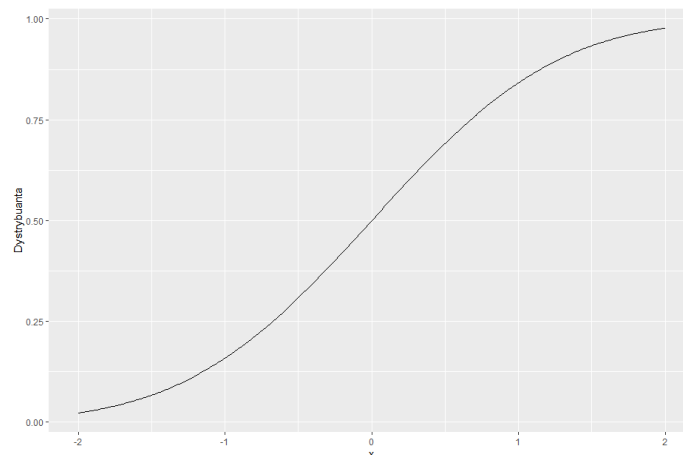
- Czym jest hipoteza statystyczna? Zdefiniuj hipotezę H_0 oraz H_1
- Czym jest poziom istotności α ?
- Co oznacza, że rozkład jest normalny?
- Czym jest dystrybucja?

Ćwiczenia:

- Utwórz wektor x o wartościach od -2 do 2 z krokiem 0.01. Dla tak zdefiniowanego wektora punktów, wyznacz i wyświetl wartości **gęstości prawdopodobieństwa**.



- Dla utworzonego wektora z ćw.1 oblicz wartości **dystrybucji** rozkładu normalnego. Wyświetl wynik.



- Zainstaluj pakiet **sp**. Wczytaj dane **ca_geo.rds** zawierające informacje na temat kwasowości gleb pewnego obszaru w Kanadzie (zmienna pH). Sprawdź klasę wczytanego zestawu danych i zmień na obiekt dataframe.
- Używając pakietu **base** wyświetl następujące statystyki: zakres wartości zmiennej pH, wartość średnią, medianę, odchylenie standardowe oraz rozstęp międzykwartyłowy.
- Wyświetl histogram zmiennej pH.
- Przedstaw estymator jądrowy gęstości dla zmiennej pH.

7. Wyświetl wykres pudełkowy dla zmiennej pH. Obserwacje mają być uznane za odstające, gdy odstają od kwartyli bardziej niż 1.5 raza wartości rozstępu międzykwartylowego.
8. Wykonaj standaryzację zmiennej pH i wyświetl jej histogram.
9. Wykonaj test normalności rozkładu zmiennej pH za pomocą testu Shapiro-Wilka używając funkcji **shapiro.test()**. Odrzuć hipotezę zerową gdy p-value < 5%.