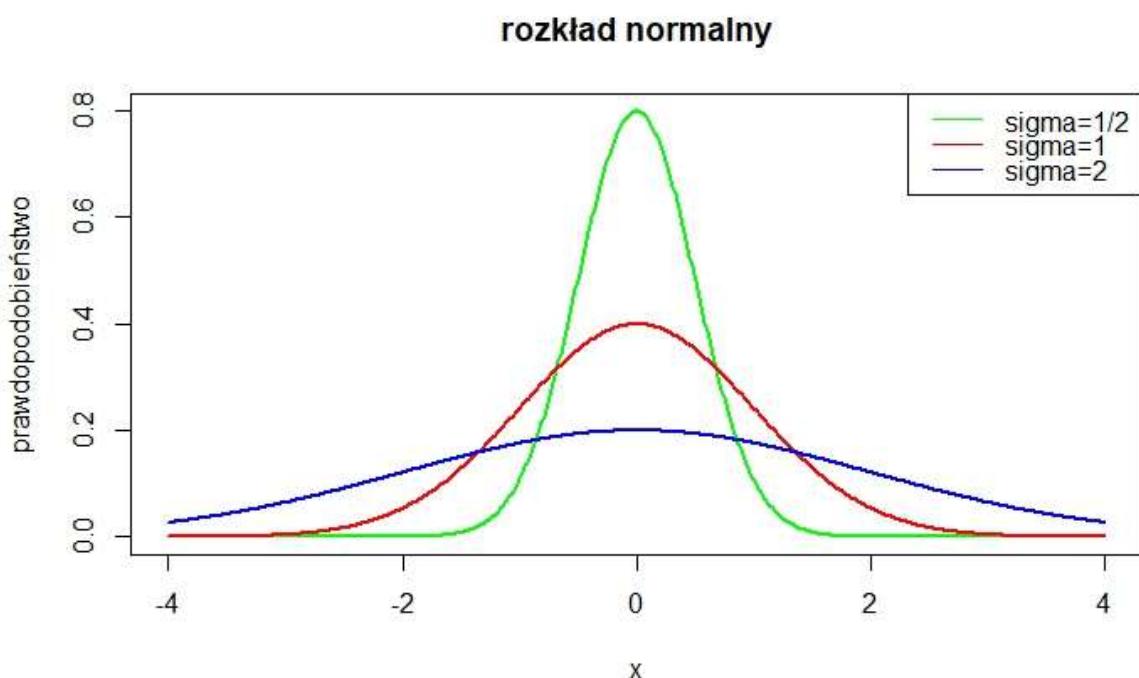


GLM-zadania

Zadanie 1. a) Wykonać poniższy rysunek rozkładu normalnego:



b) Narysować krzywe logistyczne $y=e^x/(1+e^x)$ oraz $y=e^{-x}/(1+e^{-x})$ dla x z przedziału od -10 do 10.

Zadanie 2. Dla „nadciśnienie.xlsx” zbadać wpływ wieku oraz palenia dla poziomu nadciśnienia.

Zadanie 3. Dane: „mtcars”.

A. Zbadać jednoczesny wpływ cechy „disp” i cechy „hp” na „mpg” wykorzystując model liniowy i uogólniony model liniowy z rozkładem gaussowskim.

B. Zbadać jednoczesny wpływ „disp” i „hp” na „am” stosując model logistyczny oraz model Poissona.

Zadanie 4. Dla danych „days-students.txt” (liczba zakażonych studentów w zależności od liczby dni liczących od początku pandemii):

A. Przedstawić graficznie dane

B. Wyznaczyć modele: liniowy, logistyczny oraz Poissona. Wybrać „najlepszy”.

Zadanie 5. Dla danych „PlantGrowth” z pakietu datasets:

A. Przedstawić graficznie dane

B. Wyznaczyć uogólniony model liniowy z rozkładem dwumianowym

Zadanie 6. Mamy następujące dane

Lp	ck	ha	ok
1	20	2	88
2	60	13	26
3	100	30	8

4	140	30	5
5	180	21	0
6	220	19	1
7	260	18	1
8	300	13	1
9	340	19	1
10	380	15	0
11	420	7	0
12	460	8	0

ck – poziom enzymu kinazy kreatyninowej we krwi

ha – liczba pacjentów, którzy przeszli zawał serca

ok – liczba pacjentów, którzy nie mieli zawału serca

- A. Wykonać wykres punktowy, gdzie OX - poziom ck, oś OY - prawdopodobieństwo ataku serca.
- B. Wyznaczyć model liniowej regresji logistycznej z rozkładem dwumianowym.
- C. Wyświetlić informacje o tym modelu oraz zsumaryzowane informacje
- D. Wyznaczyć model kwadratowej regresji logistycznej z rozkładem dwumianowym.
- E. Wyświetlić informacje o tym modelu oraz zsumaryzowane informacje
- F. Ocenić, który model jest „lepszy”
- G. Wykonać rysunek zawierający wykres punktowy danych