Karol Pichurski

321663

# Zadanie domowe nr 1 z Analizy Danych

Zadanie A

Zadanie A.a

Średnia: 3.5388

Mediana: 3.56

Wariancja: 0.08088656

Odchylenie standardowe: 0.281547205278262

Kwartyl dolny: 3.2775

Kwartyl górny: 3.77

Odstęp międzykwartylowy: 0.4925

Odchylenie ćwiartkowe: 0.24625

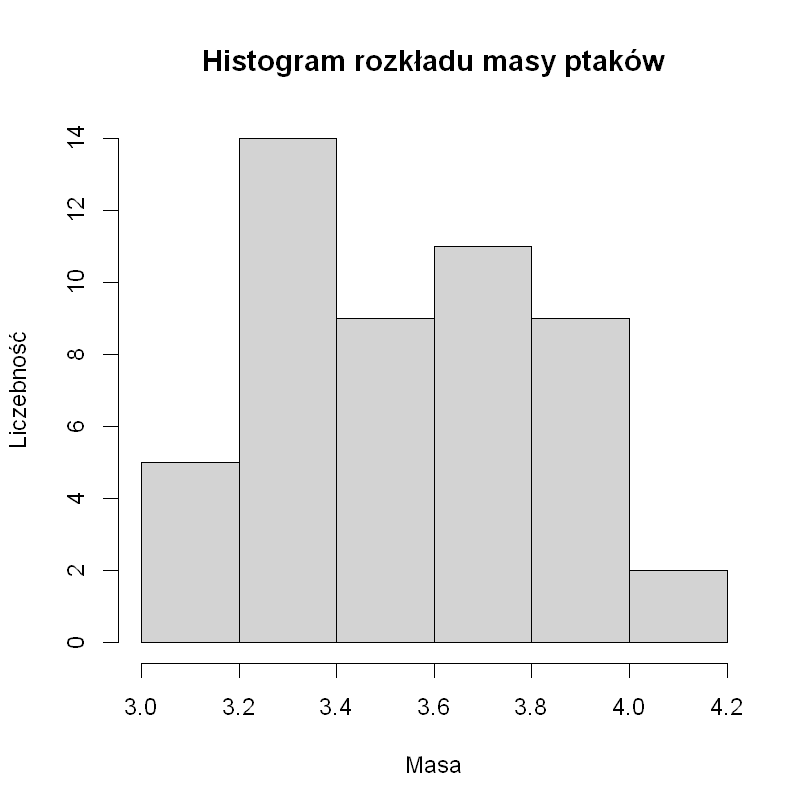
Współczynnik skośności: 0.0298407084372191

Współczynnik skośności jest > 0 więc mamy asymetrie prawostronną.

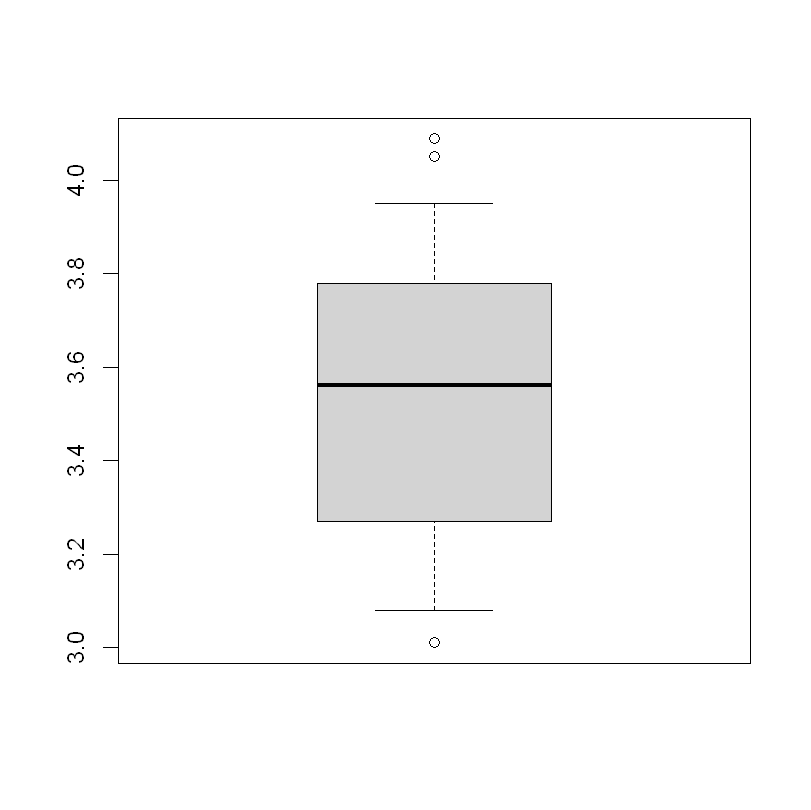
Kurtoza jest dodatnia, intensywność wartości skrajnych jest większa niż dla rozkładu normalnego („ogony“ rozkładu są „grubsze“).

Użyteczne dane to średnia, mediana i kwartyle, współczynnik skośności.

Zadanie A.b

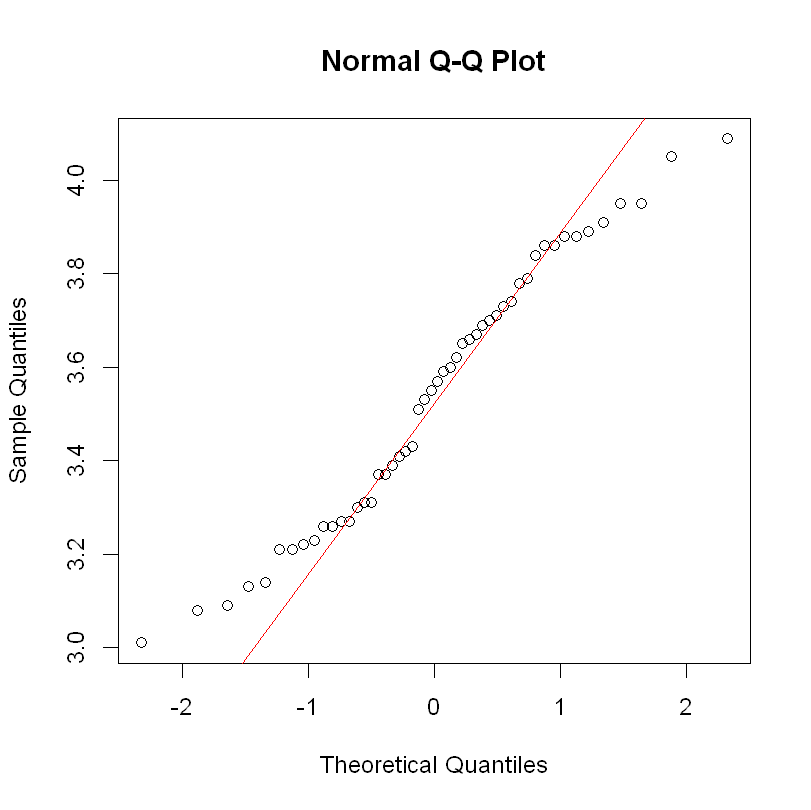


Zadanie A.c



Na wykresie ramka-wąsy zauważyłem 3 odstające obserwacje.

Zadanie A.d



Na wykresie możemy zauważyć spore ogony a więc cecha nie ma rozkładu normalnego i ma odstające obserwacje.

1. #dane
2. dane = read\_excel('Zad\_domowe\_nr\_1\_2022-2023\_KP.xlsx')
3. masa = dane$Masa\_ptaka
4. t = ((length(masa) - 1) / length(masa))
6. #A
8. #średnia
9. my\_print("Średnia", mean(masa))
11. #mediana
12. my\_print("Mediana", median(masa))
14. #wariancja
15. my\_print("Wariancja", var(masa) \* t)
17. #odchylenie standardowe
18. my\_print("Odchylenie standardowe", sd(masa) \* t)
20. #kwartyle
21. my\_print("Kwartyl dolny", quantile(masa, 0.25))
22. my\_print("Kwartyl górny", quantile(masa, 0.75))
24. #odstęp międzykwartylowy
25. my\_print("Odstęp międzykwartylowy", IQR(masa))
27. #odchylenie ćwiartkowe
28. my\_print("Odchylenie ćwiartkowe", IQR(masa) / 2)
30. #współczynnik skośności
31. my\_print("Współczynnik skośności", skewness(masa))

34. #kurtoza
35. my\_print("Kurtoza", kurtosis(masa))

38. #B
40. #histogram
41. hist(masa, main='Histogram rozkładu masy ptaków', xlab='Masa', ylab='Liczebność')
43. #C
45. #wykres ramka-wąsy
46. boxplot(masa, range=0.5)

49. #D
51. #wykres kwantyl kwantyl
52. qqnorm(masa)
53. qqline(masa, col='red')

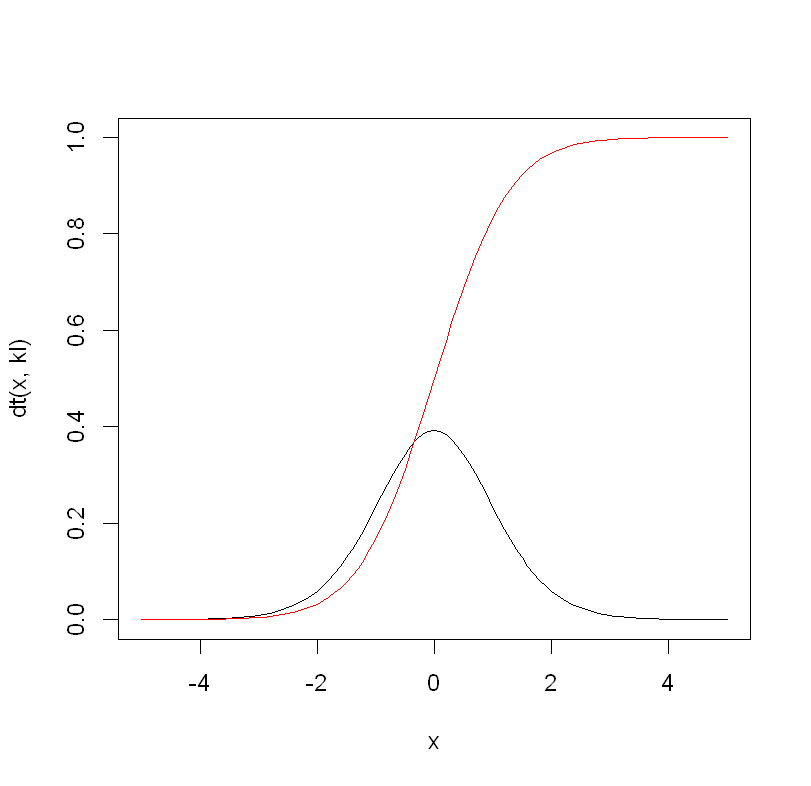
Zadanie B

Zadanie B.1

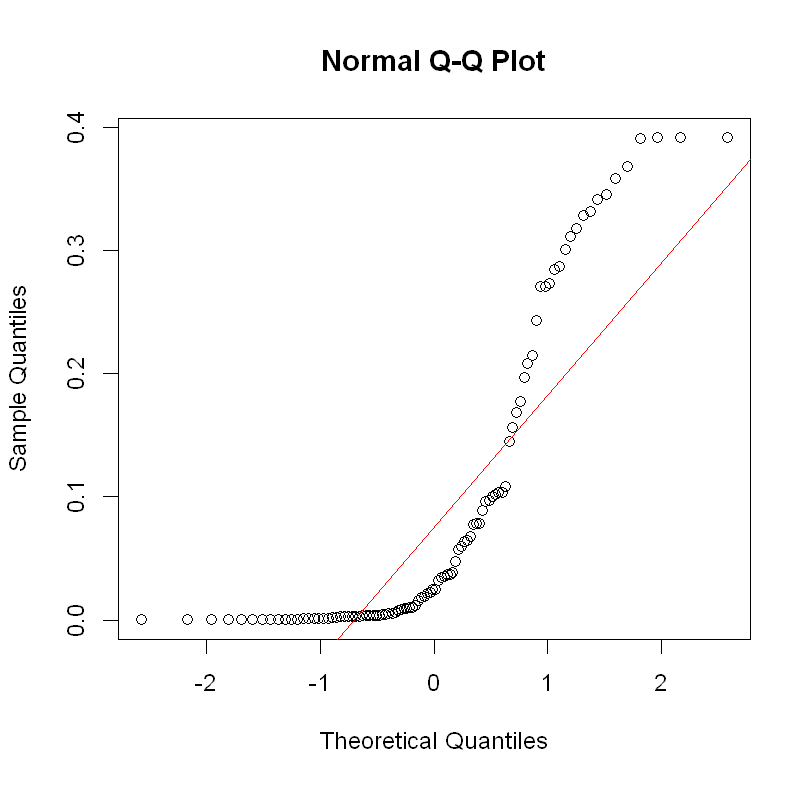
k: 14.5517245922238

kl: 14

Zadanie B.2



Zadanie B.3



Rozkład danych bardzo odbiega od rozkładu normalnego ze względu na spore ogony, które możemy zauważyć na wykresie.

Zadanie B.4

Wylosowane wartości: -0.622762571554631, 1.18726636748761

Prawdopodobieństwo dla rozkładu studenta: 0.600835207171609

Prawdopodobieństwo dla rozkładu normalnego: 0.615718436352135

Zadanie B.5

Kwantyl 0.001 dla rozkładu studenta: 1.03973864391885e+96

Kwantyl 0.001 dla rozkładu normalnego: 0.295254906651271

Kwantyl 0.005 dla rozkładu studenta: 1296880265862021120

Kwantyl 0.005 dla rozkładu normalnego: 0.299254906651271

Kwantyl 0.01 dla rozkładu studenta: 303753010.660534

Kwantyl 0.01 dla rozkładu normalnego: 0.304254906651271

Kwantyl 0.05 dla rozkładu studenta: 10.3127414723942"

Kwantyl 0.05 dla rozkładu normalnego: 0.344254906651271

Kwantyl 0.95 dla rozkładu studenta: 0.331683090193465

Kwantyl 0.95 dla rozkładu normalnego: 1.24425490665127

Kwantyl 0.99 dla rozkładu studenta: 0.32856384119809

Kwantyl 0.99 dla rozkładu normalnego: 1.28425490665127

Kwantyl 0.995 dla rozkładu studenta: 0.328191592013738

Kwantyl 0.995 dla rozkładu normalnego: 1.28925490665127

Kwantyl 0.999 dla rozkładu studenta: 0.327896478452956

Kwantyl 0.999 dla rozkładu normalnego: 1.29325490665127

Kod:

1. #dane
2. k = runif(1, 1, 29)
3. my\_print('k', k)
5. #2
7. #część całkowita
8. kl = as.integer(k)
9. my\_print('kl', kl)
11. #zakres funkcji
12. x = seq(from=-5, to=5, by=0.01)
13. xx = dt(x, kl)
14. xxlim = c(-5, 5)
15. yylim = c(0, 1)
17. #wykres gęstości rozkładu studenta
18. curve(dt(x, kl), xlim=xxlim, ylim=yylim)
20. #wykres dystrybuanty rozkładu
21. curve(pt(x, kl), xlim=xxlim, ylim=yylim, col='red', add=TRUE)
23. #3
25. #losowanie 100 liczb z rozkładu studenta z punktu 2
26. kk = sample(xx, 100)
28. #wykres kwantyl-kwantyl
29. qqnorm(kk)
30. qqline(kk, col='red')
32. #4
34. #2 liczby z rozkładu jednostajnego
35. kkk = sort(runif(2, -3, 3))
36. my\_print('Wylosowane wartości', kkk)
38. #prawdopodobieństwo rozkładu studenta
39. p1 = pt(kkk[2], kl) - pt(kkk[1], kl)
40. my\_print('Prawdopodobieństwo dla rozkładu studenta', p1)
42. #prawdopodobieństwo rozkładu normalnego
43. p2 = pnorm(kkk[2]) - pnorm(kkk[1])
44. my\_print('Prawdopodobieństwo dla rozkładu normalnego', p2)
46. #5
48. #porównanie wartości poszczególnych kwantyli
49. my\_print('Kwantyl 0.001 dla rozkładu studenta', qt(p1, 0.001))
50. my\_print('Kwantyl 0.001 dla rozkładu normalnego', qnorm(p2, 0.001))
52. my\_print('Kwantyl 0.005 dla rozkładu studenta', qt(p1, 0.005))
53. my\_print('Kwantyl 0.005 dla rozkładu normalnego', qnorm(p2, 0.005))
55. my\_print('Kwantyl 0.01 dla rozkładu studenta', qt(p1, 0.01))
56. my\_print('Kwantyl 0.01 dla rozkładu normalnego', qnorm(p2, 0.01))
58. my\_print('Kwantyl 0.05 dla rozkładu studenta', qt(p1, 0.05))
59. my\_print('Kwantyl 0.05 dla rozkładu normalnego', qnorm(p2, 0.05))
61. my\_print('Kwantyl 0.95 dla rozkładu studenta', qt(p1, 0.95))
62. my\_print('Kwantyl 0.95 dla rozkładu normalnego', qnorm(p2, 0.95))
64. my\_print('Kwantyl 0.99 dla rozkładu studenta', qt(p1, 0.99))
65. my\_print('Kwantyl 0.99 dla rozkładu normalnego', qnorm(p2, 0.99))
67. my\_print('Kwantyl 0.995 dla rozkładu studenta', qt(p1, 0.995))
68. my\_print('Kwantyl 0.995 dla rozkładu normalnego', qnorm(p2, 0.995))
70. my\_print('Kwantyl 0.999 dla rozkładu studenta', qt(p1, 0.999))
71. my\_print('Kwantyl 0.999 dla rozkładu normalnego', qnorm(p2, 0.999))