

EXERCISE 4

Intelligent data analysis DV1597

April 1, 2024

This second exercise consists of three different tasks. The tasks are extracted from Steven Skiena's material [1].

The exercise should be performed **individually**, i.e. no group cooperation. Please hand in your written as a **PDF** file via Canvas. The report should be in **English** and include your name and the answers to the four tasks below. The exercise is grade with G/Ux/U.

1. Suppose that $P(A) = 0.3$ and $P(B) = 0.7$.
 - (a) Can you compute $P(A \text{ and } B)$ if you only know $P(A)$ and $P(B)$?
 - (b) Assuming that events A and B arise from independent random process:
 - What is $P(A \text{ and } B)$?
 - What is $P(A \text{ or } B)$?
 - What is $P(A | B)$?
2. Compare each pair of distributions to decide which one has the greater mean and the greater standard deviation.
 - i. 1,7,7,7,9,12,12,1,14.
ii. 1,7,7,7,9,12,12,12,21.
 - i. -10,0,0,0,17,27,40,40.
ii. -30,0,0,0,17,27,40,40.
3. Consider the following 50 observed data:
-1.10, -0.51, 1.19, 0.29, -0.54, -1.05, -0.33, 0.82, -0.44, 0.93, 1.01, 1.81,
-1.59, 1.59, 1.27, -0.59, -0.62, -1.07, -0.33, 0.31, -1.07, -1.68, -0.60, -1.25,
0.43, 1.05, 0.47, 0.98, -0.93, 0.85, -0.36, -0.85, -0.40, 0.37 -0.82, 0.05, 1.12,
-1.28, 0.36, 0.42, -0.47, -0.05, 0.13, 0.35, -0.10, 1.07, 0.27, 0.46, 1.23, -0.53,

0.26, 0.63, -0.74, 0.84, -0.63, 0.25, 1.26, -2.22, -0.61, 0.94, -0.41, -2.09,
0.85, 1.20, -0.84, -0.28, -1.38, -1.19, -0.17, -0.13, -0.08, 0.07, -0.12, -1.35,
-1.26, 0.57, 0.21, -0.25, -0.79, -0.89, -0.51, -1.16, -0.90, 0.65, -2.61, 1.01,
-0.36, -0.13, -1.36, -2.16, 1.54, 0.42, 0.62, -0.40, -0.64, 1.77, -0.69, 0.28,
2.47, -1.35

Is it suitable to assume normality in the above data?

Reference:

1. Skiena, Steven S. The data science design manual. Springer, 2017.