Exercise Set 7

Give your results with 2 significant digits precision e.g. 0.95 or 0.15%, or as a fraction e.g. 1/3

1 Exam-style Python questions [normal-advanced]

You can find a "python cheathsheet" on the Moodle. This will not be provided in the exam, but you are free to copy any parts of it into your "personal cheatsheet"

a) Which output will the following code give?

```
import numpy as np

myArray = np.array([3,4,2])

for i in range(3):
    val = myArray[i]-np.mean(myArray)
    print(val)
```

b) Which output will the following code give?

```
import numpy as np
def myFilter(a):
    if a<0:
        return 0
    else:
        return a

myNumber = -1
print(myFilter(myNumber))
```

c) Continuing in the code above you write:

```
myArray = np.array([1,2,-1])
print(myFilter(myArray))
```

and you receive the error message:

```
ValueError: The truth value of an array with more than one element is ambiguous.
Use a.any() or a.all()
```

What went wrong? How could you solve it so that you get an array `[1,2,0]` in the end? There are several possible solutions. *You can write code, without paying attention to exact syntax, or describe the approach in words.*

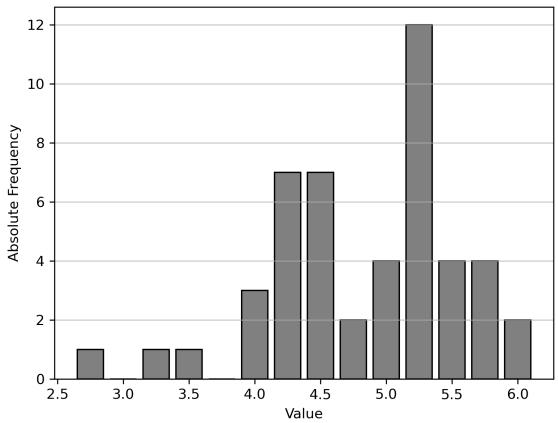
2 Exam grades [basic]

Below is a histogram of the grades of the 2024 exam.

a) Compute the mean and the median of these data.

b) Compute the unbiased estimator for the variance of this distribution.

c) Would a normal distribution give a good description of this data? If not, why not (and what kind of distribution could give a better description)?



3 Novel diet for a healthy lifestyle [normal]

You can find all necessary statistics tables in the "StatTables2024Exam.pdf" file on the Moodle.

Two groups, with twelve people in each group that initially had the exact same weight of 105kg were selected to test 2 new diets. They followed their diets strictly for 4 months and their weights after the diet phase are shown in the table below. For each diet separately, using the appropriate test, determine if the diets were effective on the $\alpha = 0.05$ level of significance. Specify the "null hypothesis."

Diet 1	86	96	82	103	91	88	94	90	97	87	97	105
Diet 2	106	118	106	91	96	102	100	100	108	105	118	93

4 Error of type 1 (false discovery/false positive) and 2 (false negative), power of a test [optional, advanced]

Let Y_1, Y_2, ..., Y_n be independent and identically distributed (i.i.d) random variables that follow the Gaussian/normal $\mathcal{N}(\mu = \theta, \sigma^2 = 1)$ distribution. We are interested to test if $\theta = 0$. There are many tests possible, and they differ in the amount and relative ratio of the type 1 and type 2 errors.

We propose a certain test for for $\theta = 0$ is proposed: The test accepts $\theta = 0$ if the measured average times the square-root of the number of measurements, $|\sqrt{n}\bar{Y}| \leq 2$, and rejects is otherwise (hence implying $\theta \neq 0$).

a) Lets assume that the true value of θ is zero, i.e. the real distribution is the standard (centered and normalized) normal law $\mathcal{N}(\mu = 0, \sigma^2 = 1)$. What is the probability of the test to incorrectly state $\theta \neq 0$ when only one single measurement is performed?

b) How does that probability change if 9 measurements are made?

c) We now assume that $\theta = 1$ and still take 9 measurements. Compute the type 1 (false discovery/positive) and type 2 errors.

d) Compute the, depending on which is appropriate, the type 1 or the 2 error (as well as the power of the test, corresponding to 1 minus the probability for a type 2 error), when $\theta = 0$, 0.5, 1, 1.5 and 2.