Exercise 1 Calc. : ✓

In 2002 in Luxembourg the average temperatures per month have been recorded. It is known that January 2002 was the coldest month measured as 1.6°C and the highest average temperature was measured in June 2002 as 18.6°C.

1. **Justify**, that in Europe the monthly average temperatures for some consecutive years can be modelled with a periodic model. — 2 marks
2. **Give** the amplitude and the period of this model. — 2 marks
3. **Determine** the parameters a, b, c and d in the model of the type: — 5 marks

$$T(x) = a \cdot \sin(b \cdot (x - c)) + d$$

   that describes the given data where T is the average Temperature and x is the month, starting with $x = 1$ for January 2002.

On one specific day in March 2002 the rainfall was observed. The rainfall on that day can be modelled by the function

$$R(t) = 0.002t^3 - 0.064t^2 + 0.512t, \qquad 0 \leq t \leq 24$$

where $R(t)$ is the rate of rainfall in mm/h and t is the time in hours.

4. **Describe**, using a short text description, this day in terms of rainfall. Your answer should focus on the times with the most and the least rainfall. — 3 marks

An empty glass cylinder was placed outside during this day to help see how much rain had fallen.

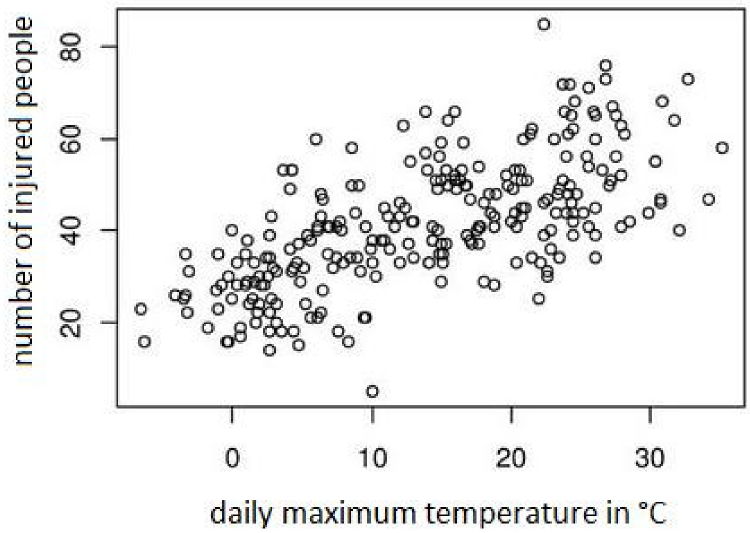
5. **Sketch** the graph of a function, that shows the height of water in this glass cylinder. — 3 marks
6. **Calculate** the total amount of rain on that day in mm. — 2 marks

The year 2002 in Luxembourg turned out to have 195 rainy days and 170 days without rain. It can be assumed, that all days have the same chance of being a rainy day. One year later, meteorologists want to investigate, if there was more rain in 2003. Unfortunately, some data were lost, so they took only a small sample of 30 consecutive days.

7. **Calculate** the probability that it rains on a random day, if we assume, that the total number of rainy days in both years remains constant and the rainy days are equally distributed over the whole year. — 1 mark

8. **Use** a NHST procedure to find out how many days it must rain so the meteorologists can say that there was more rain in 2003 compared to 2002 when the significance level is at 5%. — 5 marks

The following diagram shows the maximum temperature and the number of injured people caused by traffic accidents in Berlin on a long-term base.



9. **Describe** the correlation between the two values. — 1 mark

10. **Explain**, why the number of injured people possibly correlates in such a way with the maximum temperature. — 1 mark

Exercise 2 Calc. : ✓

In a Covid-19 testing station, 19 people with symptoms were tested on a specific day and 6 of them had a positive result. On the same day, 87 people without symptoms were tested of which 85 were tested negative.

1. **Show** that the probability of getting a positive result depends on whether a person has symptoms or not. — 2 marks

To protect personal data, the test probes are labelled with a code, that contains 2 letters (out of an alphabet with 26 letters) and 4 digits (0–9). The same letters and digits may be chosen more than once.

2. **Calculate** the total number of different codes, that can be created by this system. — 2 marks

After several months, statistics have shown, that 1.7% of the people without symptoms are tested positive. A company, with 20 employees (all without symptoms), instructs everyone to get tested.

3. **Give** two assumptions, that need to be made to model this situation with a binomial distribution. — 2 marks
4. **Calculate** the probability, that at least one of the employees is tested positive. — 3 marks

A different company in another country also sends all their employees for a Covid-19 test. Assuming that the situation can be modelled by a binomial distribution given by the formula

$$B(84; 0.02; k) = \binom{84}{k} \cdot 0.02^k \cdot 0.98^{84-k}.$$

5. **Interpret** the values 84, 0.02 and 0.98 in the given context. — 3 marks

On March 5 in 2020 a man who returned from Italy is the first person in Luxembourg who was tested positive with COVID-19. So, this day is marked as day 0 in the statistic. The following table shows the total number of registered infected people in Luxembourg in the days after the first case appeared.

Day	0	1	2	3	4	5	6
Number	1	3	4	5	5	7	7

6. **Draw** a scatter graph of these values together with a linear and an exponential regression model. — 3 marks
7. **Give** the equations that describe the two regression models in part 6. — 2 marks
8. **Explain**, why it is so difficult to decide, if the spread out of the virus is best modelled with a linear or an exponential model in this early stage. — 2 marks

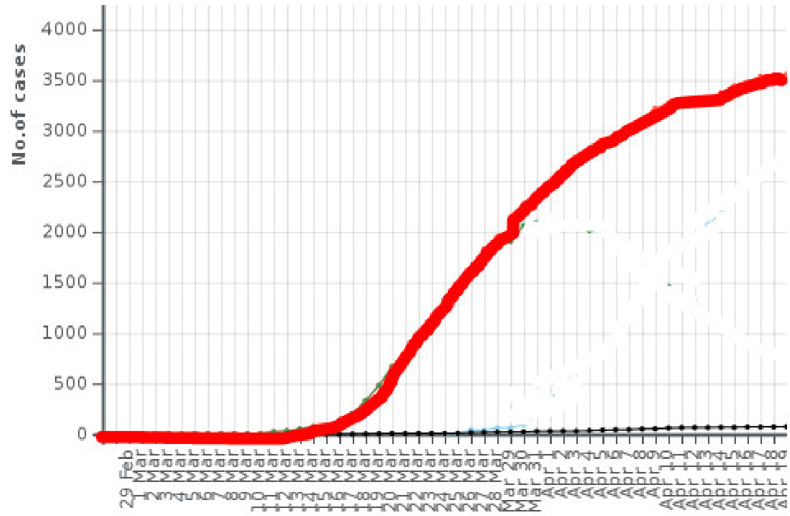
After seven more days other models were made to make better predictions, where t is given in days:

$$A(t) = 1.35567 \cdot 1.46977^t \qquad\qquad B(t) = 12.4396 \cdot t - 34.8571$$

On day 16, there were 670 registered cases of COVID-19 in Luxembourg.

9. **Calculate** the predicted number of infected people on day 16 with model A and model B and **compare** it with the true number. **Decide**, which model obviously works better for this situation and **reason** your answer. — 2 marks

The following diagram shows the graph of the total number of registered infections for the first 4 weeks in Luxembourg.



10. **Give** two possible reasons, why the curve flattens in a later stage. — 2 marks

The curve can be modelled by the function

$$C(t) = \frac{3404}{1 + 193 \cdot e^{-0.233 \cdot t}}$$

11. **Determine** the day with the highest infection rate by calculation. — 2 marks