

Exercise 4: Statistical Analysis & Mathematical Models

Submission Deadline: December 01 2025, 07:00 UTC

University of Oldenburg

Winter 2025/2026

Instructors: [Jannik Schröder](#), [Wolfram "Wolle" Wingerath](#)

Submitted by: <your names here>

Part 1: Statistical Distributions

/ 25

1.) Assuming that the relevant distribution is normal, estimate the probability of the following events:

a) That there will be 70 or more heads in the next hundred flips of a fair coin?

Solution:

<your answer here>

b) That a randomly selected person will weight over 150 kg, assuming mean and standard deviation of 55/9 kg for women and 70/11 kg for men?

Solution:

<your answer here>

2.) The average on a history exam was 85 out of 100 points, with a standard deviation of 15. Was the distribution of the scores on this exam symmetric? If not, what shape would you expect this distribution to have? Explain your reasoning.

Solution:

<your answer here>

3.) Facebook data shows that 50% of Facebook users have a hundred or more friends. Further, the average user's friend count is 190. What do these findings say about the shape of the distribution of number of friends of Facebook users?

Solution:

<your answer here>

4.) Assume that the European electricity prices for non-household consumers in the first half of 2025 are above €0.20 per kWh on 50% of all days. Furthermore, assume that the average electricity price per day is €0.25 per kWh in the same time period. What do these results tell us about the way electricity prices are distributed?

Solution:

<your answer here>

Part 2: Significance & Permutation Tests

/ 30

4.) Which of the following events are likely independent and which are not? Explain your reasoning.

a) Coin tosses

Solution:

<your answer here>

b) Goals in soccer

Solution:

<your answer here>

c) Party success rates in presidential elections

Solution:

<your answer here>

d) Electricity price fluctuations on consecutive days

Solution:

<your answer here>

5.) The 2010 American Community Survey estimates that 47.1% of women aged 15 years and over are married. (Assume that marriages happen independently from one another.)

a) Randomly select three women between these ages. What is the probability that the third woman selected is the only one that is married?

Solution:

<your answer here>

b) What is the probability that all three women are married?

Solution:

<your answer here>

6.) Obtain data on the heights of m men and w women.

a) Use a t-test to establish the significance of whether the men are on average taller than the women.

Solution:

<your answer here>

b) Perform a permutation test to establish the same thing: whether the men are on average taller than the women.

Solution:

<your answer here>

Part 3: Building Models

/ 25

7.) Quantum physics is much more complicated than Newtonian physics. Which model is preferable according to the Occam's Razor, and why?

Solution:

<your answer here>

8.) Name 2 models that predict something that you are personally interested in. For each of these, decide (and briefly explain) which properties these models have:

a) Are they discrete or continuous?

Solution:

<your answer here>

b) Are they linear or non-linear?

Solution:

<your answer here>

c) Are they blackbox or descriptive?

Solution:

<your answer here>

d) Are they data driven or first principle?

Solution:

<your answer here>

9.) Give at least 1 example of a first-principle and at least 1 example of a data-driven model used in practice.

Solution:

<your answer here>

Part 4: Evaluating Models

/ 20

10.) Suppose you build a classifier that answers yes on every possible input. What precision and recall will this classifier achieve?

Solution:

<your answer here>

11.) Is it better to have too many false positives, or too many false negatives? Explain.

Solution:

<your answer here>

Finally: Submission

Save your notebook and submit it (as both **notebook and PDF file**). And please don't forget to ...

- ... choose a **file name** according to convention (see Exercise Sheet 1, but please **add your group name as a suffix** like `_group01`) and to
- ... include the **execution output** in your submission!