HOMEWORK 1 (Due 17/11/2025 – 17:30)

Show all calculations clearly; round probabilities to 2 decimal places. For numerical parts, you may use programming or a calculator, but all steps must be identified symbolically first. Interpret results in plain language; numerical work without context will lose credit. You are supposed to upload your scanned solutions to ODTÜClass before the due date.

1) "The Soundtrack of Student Life"

During METU's annual "Music and Mind" week, a psychology club decided to explore whether listening habits relate to mood and focus. They asked 40 students how many hours of music they listened to in a typical week.

The (unsorted) data collected were:

```
10, 5, 13, 8, 15, 7, 9, 11, 14, 16, 21, 19, 18, 23, 17, 12, 6, 25,
20, 22, 9, 14, 15, 18, 27, 24, 28, 32, 30, 26, 33, 36, 35, 31, 40,
39, 38, 42, 37, 41
```

When the group tried to make a bar chart for their social media post, some argued the chart "looked wrong" as the bars seemed uneven, and they couldn't tell what "typical" listening looked like. The team reached out to you for help summarizing the data properly. Your report to the club should include:

a) Decide on a reasonable number of classes for grouping (defend your choice in plain language).
b) Construct the frequency, relative frequency, and cumulative frequency table.
c) Use midpoints to estimate the mean listening time and identify which class contains the median.
d) Describe the shape of the distribution (symmetric, right-skewed, or left-skewed).
e) Write two sentences interpreting the story behind the numbers: what kind of listening pattern does the typical student have, and what might explain the extremes?

2) "The Myth of the 20-Hour Week"

Every midterm season, METU students exchange advice about study habits. One popular claim circulating on social media reads:

"Everyone who passes the first midterm studies around 20 hours a week. That's the magic number."

To test this myth, a teaching assistant surveyed 20 randomly chosen students from different departments about their weekly study hours:

`8, 9, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 26, 28, 35, 40`

When the assistant shared the summary, some students protested: "You can't just use averages that hides how different we are!"

You are asked to prepare a balanced analysis for the department newsletter.

- **a)** Compute and present the mean, median, variance, standard deviation, and mean absolute deviation (MAD).
- **b)** Discuss whether the slogan "everyone studies about 20 hours" is statistically fair.
- **c)** Recalculate the mean and SD after removing the two highest values (35 and 40). Compare and interpret the changes.
- **d)** Explain why certain measures react strongly to extreme values while others remain stable, using your results as evidence.
- **e)** In three sentences, summarize the story you would tell a reader who doesn't know statistics: what do these numbers really say about student study behavior?

3) "Rush Hour Reliability"

A local delivery start-up, AnkaExpress, partners with METU statistics students to analyze their delivery times (in minutes) for orders completed last week.

The company promises "fast and consistent" service, but drivers often complain that heavy traffic and bad weather make this claim unrealistic.

To investigate, 25 delivery times were recorded randomly during various hours of the day:

```
22, 25, 27, 29, 31, 28, 33, 35, 26, 37, 30, 32, 34, 36, 38, 40,
42, 44, 41, 43, 45, 49, 52, 55, 60
```

Management wants to know: *Are we really "consistent"?*

You are asked to prepare a short data brief for the company dashboard:

a) Compute Q_1, Q_2, Q_3, and the IQR for the data.

b) Use the 1.5×IQR rule to detect possible outliers, and identify any deliveries that could be considered unusually slow.

c) Sketch a boxplot and describe the shape of the distribution (symmetric, right-skewed, or left-skewed).

d) Explain, in plain language, what the quartiles and whiskers mean for the company's performance.
For example: "25% of orders are delivered within ___ minutes."

e) The operations manager argues that the few long deliveries "don't matter" because they occur during rainstorms.

f) How would you respond statistically?

g) Suppose the company expands to a new district, where five additional deliveries took 70, 72, 74, 78, and 80 minutes. Predict qualitatively how these would alter the boxplot and what they would imply about service reliability.

4) "Searching for a Match"

Two researchers at a medical statistics lab are discussing different situations they've encountered and wonder which ones can be modeled using Bernoulli, Binomial, or Geometric distributions. They ask you to help formalize and compute a few probabilities.

Part A: Faulty or Fine?

A small electronics manufacturer produces batches of sensors used in hospital monitoring equipment. Based on past data, each sensor has a 4% chance of being defective. Each morning, a technician tests 50 sensors before shipment.

One day, another technician claims, *"If we ever get a perfect batch with zero defects I'll treat everyone to breakfast."*
How likely is that to happen? Compute the probability that no defects occur, find the expected number of defectives, and comment on whether the promise is generous or safe.

Part B: The Blood Bank Dilemma

A regional blood bank searches for a donor with a rare antigen that occurs in 25% of people. Nurses test volunteers one by one, stopping when the first suitable donor is found. Find the probability that the first match appears within the first four tests, and the expected number of tests needed.

Briefly explain, in your own words, how this kind of model captures "waiting time until success," and what practical insight the expected value gives the medical team.

Part C: A Single Toss for Justice

A university council debates whether to approve a new student-led research grant that ultimately comes down to a simple vote: *"Approve"* (success) or *"Reject"* (failure).

Based on previous funding rounds, the probability of approval for any such proposal is about **0.6**.

- **a)** Write the probability distribution of X.
- **b)** Compute $E(X)$ and $\mathrm{Var}(X)$, and interpret $E(X)$ in plain language.
- **c)** If five independent proposals of similar type are submitted during the semester, what would you expect the average number of approvals to be?