

STATISTICAL LABORATORY

Lecture 3: Exploratory Data Analysis (Practice)

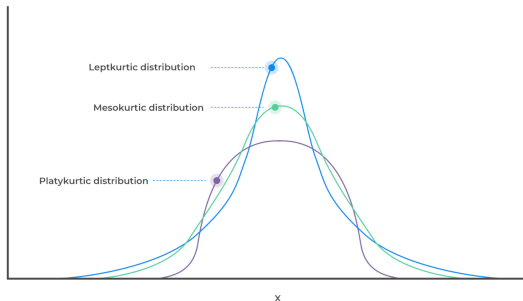
Yiu Lim LUI, Allen

Lisbon School of Economics & Management
Universidade de Lisboa

`allenlui@iseg.ulisboa.pt`

Applied Mathematics for Economics and Management
1st Year / 1st Semester – 2026/2027

Kurtosis Graph



- **Leptokurtic Distribution** (high peak, heavy tails): Most data concentrated near the mean. More extreme values (outliers) likely.
- **Platykurtic Distribution** (low peak, light tails): Data more evenly spread. Fewer extreme values.
- **Mesokurtic Distribution** (moderate peak, normal tails): Moderate concentration of data around the mean.

Exercise 2.1

For quality-control purposes, the product is shipped in boxes of six units. A simple random sample without replacement of 40 boxes is drawn from the lot. For each box, the number of defective units is recorded. The observed sample is:

2, 0, 1, 3, 0, 3, 6, 4, 5, 3, 0, 1, 0, 3, 2, 0, 1, 0, 2, 3, 1, 0, 3, 0, 1, 0, 0, 1, 5, 6, 0, 2, 5, 1, 0, 3, 0, 2, 3, 0

- a) Define the statistical unit and the variable under analysis, and classify the variable.
- b) Present the frequency table.
- c) Provide a graphical representation of the data.
- d) Compute the range and the interquartile range (IQR).
- e) Assess the symmetry (skewness) of the distribution.

Exercise 2.2

The following table shows the number of weekly study hours outside classes observed in a sample of MAEG students:

6.7	8.2	8.5	5.7	9	8	18.2	9.5	5.6	11.2	9.8	3.1	7.6	8.8	6.3
-----	-----	-----	-----	---	---	------	-----	-----	------	-----	-----	-----	-----	-----

- a) Define the statistical unit and the variable under study.
- b) Compute the five-number summary for this dataset.
- c) Analyse the symmetry of the data using a boxplot.

Exercise 2.3

The statistical information on the monthly wages of workers in a textile company is presented below:[cite: 3]

Salary (euros)	Workers (%)
[600 ; 900]	9
]900 ; 1000]	18
]1000 ; 1100]	35
]1100 ; 1200]	22
]1200 ; 1300]	16

- a) Define and classify the variable under study.
- b) Provide a graphical representation of the data.
- c) Assess the symmetry of the data based on the graphical representation.

Exercise 2.4

Consider the grades of 50 students in this course from the previous academic year:

93, 77, 67, 72, 52, 83, 66, 84, 59, 63,
75, 97, 84, 73, 81, 42, 61, 51, 91, 87,
34, 54, 71, 47, 79, 70, 65, 57, 90, 83,
58, 69, 82, 76, 71, 60, 38, 81, 74, 69,
68, 76, 85, 58, 45, 73, 75, 42, 93, 65

- a) Represent the data in an ordered stem-and-leaf plot.
- b) Compute the five-number summary from the plot.
- c) Construct a boxplot and assess the symmetry (or asymmetry) of the data.
- d) Group the data into 5 classes of equal width, with the lower bound of the first class set to 25 and construct a histogram.

Exercise 2.5

The grades of 10 students in Statistics Laboratory (SL) and Linear Algebra (LA) are presented below:

SL	15	12	18	14	13	10	15	16	17	17
LA	14	15	17	13	15	10	14	15	18	17

- a) Provide a graphical representation of these data.
- b) Assess the possibility of a linear relationship between the grades in these two courses.

Exercise 2.6

The statistical information on the heights of students in a school is presented below:

Height (cm)	Number of Students
[150 ; 158]	5
]158 ; 166]	18
]166 ; 174]	42
]174 ; 182]	27
]182 ; 190]	8

- a) Identify the statistical unit and the variable, and classify the variable.
- b) Briefly assess the data based on its graphical representation.