107.369 Statistics and Probability Theory

UE Exercises, 2.0 h, 3.0 EC

Winter semester 2023/24

Coordinator: Univ.Ass.in Tijana Levajković, Priv.-Doz.in, Ph.D.
Email: tijana.levajkovic@tuwien.ac.at
Contact: Institute of Statistics and Mathematical Methods in Economics
DB 04 B24 (Freihaus, 4.floor, yellow)
https://tiss.tuwien.ac.at/person/321365.html?suchtext=levajkovic/

Supervisors:

Nadja Azzouz	nadja.azzouz@tuwien.ac.at
Alejandra Avalos Pacheco	alejandra.avalos@tuwien.ac.at
Stefanie Groger	stefanie.groeger@tuwien.ac.at
Manuel Hölzl	manuel.hoelzl@tuwien.ac.at
Stefan Laszlo	stefan.laszlo@tuwien.ac.at
Tijana Levajković	tijana.levajkovic@tuwien.ac.at
Julian Silan	julian.silan@tuwien.ac.at
Manuel Schranzhofer	manuel.schranzhofer@tuwien.ac.at
Jan Ole Schürmann	jan.schuermann@tuwien.ac.at

Course registration: In TISS

Course Pages: All course material and further course information can be found in TUWEL.

Time and place: All exercise sessions are on on Tuesdays according to the following schedule.

Group	Hours	Room	Supervisor
1	9:00 – 11:00	Hörsaal 14	Manuel Hölzl
2	9:00 – 11:00	Seminarraum AE U1-5	Jan Ole Schürmann
3	11:00 – 13:00	Seminarraum AE U1-5	Jan Ole Schürmann
4	11:00 – 13:00	Hörsaal 14	Manuel Hölzl
5	13:00 – 15:00	Seminarraum AE U1-5	Julian Silan
6	13:00 – 15:00	Hörsaal 14	Tijana Levajković
7	15:00 – 17:00	Hörsaal 14	Stefan Laszlo
8	15:00 – 17:00	Seminarraum AE U1-5	Julian Silan
9	17:00 – 19:00	Seminarraum AE U1-5	Tijana Levajković
10	17:00 – 19:00	Hörsaal 14	Stefanie Gröger
11	9:00 – 11:00	Zeichensaal 1	Manuel Schranzhofer
12	11:00 – 13:00	Sem.R.DB gelb. 03	Alejandra Avalos Pacheco
13	9:00 – 11:00	Seminarraum AE U1-4	Nadja Azzouz
14	11:00 – 13:00	Seminarraum AC 02-1-UIW	Nadja Azzouz

On **Tuesday 10.10.2023, 8:00 - 9:15** in zoom, a brief **introduction to** `R` will be given to **all** course participants. The meeting details will be available in TUWEL. The **first exercise session** is on **Tuesday 17.10.2023** and it takes place **in presence**.

Exercise dates: There are 14 study weeks. The course activities are planned as follows:

Week	Date	Exercises session	Sheet
II	October 10	Introduction to R	-
III	October 17	HW1 presented	1
IV	October 24	HW2 presented	2
V	October 31	HW3 presented	3
VI	November 7	HW4 presented	4
VII	November 14	HW5 presented	5
VIII	November 21	HW6 presented	6
IX	November 28	HW7 presented	7
X	December 5	HW8 presented	8
XI	December 12	HW9 presented	9
XII	December 19	HW10 presented	10
	Dec. 22 - Jan. 5	Christmas Holidays	
XIII	January 9	HW11 presented	11
XIV	January 16	HW12 presented	12
XV	January 23	HW13* bonus	13
	January 31	Grades	

Description: This is a one-semester exercise course covering methods and tools of descriptive and inferential statistics, core concepts of probability theory and implementation of the statistical analysis methods. The contents of the course are presented in the accompanying lectures and then practiced and deepened in these exercise sessions.

The focus is on methods and tools of descriptive and inferential statistics, core concepts of probability theory and implementation of the statistical analysis methods.

- Probability theory: counting (multiplication rules, permutations, combinations), calculate probabilities, Bayes theorem, discrete and continuous random variables, distributions (Bernoulli, binomial, geometric, Poisson, uniform, normal and exponential distributions), quantiles, mean and variance, independence, conditional probability, covariance, correlation, independence, Law of large numbers, Central limit theorem.
- Descriptive Statistics: elementary statistics, empirical distribution, graphical representations (frequency table, diagrams, histograms, scatterplots).
- Inferential statistics: Significance tests and confidence intervals (one- and two-sample z-tests and t-tests), p-value, significance level, α- and β-errors (also called errors of Type I and Type II), non-parametric tests (goodness-of-fit tests, chi-square tests for homogeneity and independence), analysis of variance, multiple comparisons.
- Linear regression (correlation, simple linear model, regression line, coefficient of determination).

Particular attention is paid to the use of the statistical software R to implement the statistical analysis methods covered in the course.

Alternative in summer term: The course UE 107.370 Statistics and Probability Theory offered in summer semester is usually less frequented.

Prerequisites:

- Successful completion of STEOP.
- Basic knowledge of linear Algebra and Calculus.
- Registration in the exercises in TISS within the registration period (from 04.09.2023, until 02.10.2023). Deregistration is possible until 31.10.2023.
- Access to a computer where the statistic software `R` (https://cran.r-project.org) and the Interface `RStudio` (https://www.rstudio.com) are installed.

Procedure:

- **Ticking problems and uploading solution paths:**
  Problem sheets are posted online weekly on Monday afternoon in TUWEL. Each homework assigment will have **five** problems except the last bonus one which will have **two** problems. Up until **Monday 20:00h** of the following week students are expected to **tick** in TUWEL **problems** and **upload the solution paths** of those problems that they solved, i.e., there are 7 days from when the problems are posted until they are to be ticked and solutions to be submitted.

  There will be **13 sheets** in total, **12 regular sheets with five problems** and **one bonus** sheet with **two bonus problems**. In total there will be **62 problems** to be ticked (**60** will be used for averaging when computating the grades).

  The first set of problems is posted on Monday, 09.10.2023. (they are to be ticked and solution paths uploaded in TUWEL until 20:00h on Monday 16.10.2023.) and are to be presented on Tuesday 17.10.2023.

- **Weekly submissions:**
  Weekly submission of solution paths is mandatory. Up to **two** missing submissions are allowed. For an exercise sheet at least **two** problems need to be ticked, and otherwise the submission is marked as missing. If problems are ticked but not all associated solution paths uploaded, then all ticks for the corresponding sheet will be deleted (and the sheet is marked as missing). Delayed submissions are **not accepted** (no ticks, and thus a missing submission).

- **Exercises:**
  For the exercise sessions, please prepare your homework for presentation.
  Attendance is **mandatory**. By signing, you confirm your presence in each exercise session. From all submissions, the supervisor asks a randomly chosen student to present a problem they ticked. The performance will be assessed as a percentage (0-100%). If you are clearly unprepared for a ticked example (Achievement = 0%), all ticked examples of this exercise will be deleted. If you do not attend a class, all ticked examples of this exercise will be deleted.

Attendance: Attendance is required, up to two absences are allowed.

Academic integrity: Students in this course are expected to comply with the academic integrity code and to behave with academic honesty. Academic dishonesty is defined as cheating of any kind, including misrepresenting one's own work, taking credit for the work of others without crediting them and without appropriate authorization, and the fabrication of information.

Plagiarism: All your submitted work should be result of your own thought and research.
We take plagiarism seriously; students who are caught plagiarizing will fail the course.

Grading: For a positive grade, the following requirements must meet:

(I) at most two missing attendances,

(II) at least 20 problems solved and

(III) at least one presentation given.

If **all three** requirements (I)-(III) are met, then the final grade is derived from the `score` which is defined as the **mean** of the two percentages:

- the percentage of ticked problems, computed as

$$P_1 = \frac{\text{the total sum of ticks (including the bonus problems solved)}}{60} \cdot 100\%$$

- the average score of graded presentations, computed as

$$P_2 = \frac{\text{the sum of graded presentations (in percentage)}}{\text{the number of presentations given}}$$

according to the following key:

grade	`score` (%)
5	[0%, 50%)
4	[50%, 62,5%]
3	(62,5%, 75%]
2	(75%, 87,5%]
1	(87,5%, 100%]

Example: There are 60 problems in total. Consider that a student solved and submitted solutions of 48 problems, and has presented two solutions with achievements of 70% and 85%. These lead to

$$P_1 = \frac{48}{60} = 80\%, \quad P_2 = \frac{70\% + 85\%}{2} = 77.5\% \quad \text{and} \quad \texttt{score} = \frac{80\% + 77.5\%}{2} \approx 78.75\%.$$

and the final grade 2 since $\texttt{score} \in (75\%, 87.5\%]$.

If **at least one** of the three requirements (I)-(III) is not satisfied, the final grade is 5 (failing the course).

Help: If you have any questions concerning the exercises, please contact your supervisor by e-mail.