COURSE OF STUDY: STATISTICS
ACADEMIC YEAR: 2024 - 2025
ACADEMIC SUBJECT: Statistical Inference and Linear Models (Inferenza Statistica e Modelli Lineari)

General information	
Year of the course	Second year
Academic calendar (starting and ending date)	Second semester (17 February – 6 June 2025)
Credits (CFU/ETCS):	10
SSD	SECS-S/01 (Statistics)
Language	Italian
Mode of attendance	Optional but strongly recommended

Professor/ Lecturer	
Name and Surname	Nunziata Ribecco
E-mail	nunziata.ribecco@uniba.it
Telephone	+390805049251
Department and address	Largo Abbazia Santa Scolastica, 53 (room n. 17)
Virtual room	For all students, Microsoft Teams code: **rfl3ft0** Thesis students, Microsoft Teams code: **xiz8pwr**
Office Hours (and modalities: e.g., by appointment, on line, etc.)	Tuesday: 9.30 - 11.30; Wednesday: 9.30 - 11.30. During the period of the lessons: Tuesday: 9.00 - 11.00 (first and second semester); Wednesday: 9.00 - 11.00 (first semester); Wednesday: 11.30 - 13.30 (second semester).

Work schedule			
Hours			
Total	Lectures	Hands-on (laboratory, workshops, working groups, seminars, field trips)	Out-of-class study hours/ Self-study hours
250	49	21	180
CFU/ETCS			
10	7	3	

Learning Objectives	The course aims to prepare professionals capable of analyzing real phenomena using the tools of classical statistical inference that are presented both from a theoretical and an applied point of view. It also wants to lay the foundations for Bayesian and decision-making approaches, for multivariate analyzes and non-parametric techniques. The topics in the program are those of classical statistical inference and are treated with the aim of deepening both the statistical-mathematical and the application part, developing the ability to identify a problem and solve it with a professional approach. The inferential problems are addressed, initially, under the general aspect of the methods (estimation of the parameters and verification of the hypotheses) which see, mainly, their application to one or more normal populations with a look, also, to the dichotomous populations, the whole is presented with reference to both small and large samples. Furthermore, some problems of non-parametric statistics, some resampling techniques and decision theory are analysed. The purpose with which these topics are presented is to make the student master of the

	mathematical statistical aspect and able to be autonomous in the inferential analysis of sample data concerning phenomena that occur in various application contexts (experimental, economic, social, etc. ...).
Course prerequisites	The student must have passed the exam of Institutions of Statistics, moreover he must have knowledge of Mathematical Analysis (Institutions), Mathematical analysis and Elements of Linear Algebra and Probability.

Teaching strategie	The face-to-face teaching concerns theoretical topics and exercises during which problems are exposed to develop the student's ability to apply the theory in the context of real phenomena. Furthermore, laboratory activities are carried out with the R software, for a number of hours equal to approximately one third of the entire duration of the course, with applications of statistical inference methods to real phenomena.
Expected learning outcomes in terms of	The expected learning outcomes are contained in these points: - know the methods to estimate the parameters and be able to apply them (mainly maximum likelihood estimators) both for the parameters of normal populations and for parameters of a binomial model; - know and be able to verify the properties of an estimator and face the problem of determining its precision (interval estimation); - be able to deal with the problem of verifying both parametric and non-parametric hypotheses with reference to one or more populations; - knowing how to interpret the results obtained and being able to make decisions for operational purposes; - be familiar with inferential methods and acquire mastery of the tools in order to be able to operate in total autonomy when dealing with concrete problems; - acquire inferential logic both in methodological and applicative terms.
Knowledge and understanding on:	At the end of the course the student must know the theory of classical inference and its applications: - Estimation theory: Determination of the main estimators using the maximum likelihood method. Know the properties of these estimators and be able to verify the properties of a generic linear estimator. Know the methods for the accuracy of an estimate and be able to determine it. - Parametric hypothesis testing theory: Knowing what a parametric statistical hypothesis is and all the problems connected to it. Knowing how to apply hypothesis testing tools for one or more normal and binomial model. - Linear model: Knowing what is meant by the relationship between two variables and how to deal with the inferential problems connected to it both in the case of a non-quantitative antecedent variable (comparison between several means of normal populations) and in the case of a quantitative antecedent variable (linear regression). - Nonparametric inference: Know the methods of nonparametric inference and their application.
Applying knowledge and understanding on:	At the end of the course, the student must be able to apply the methods of statistical inference presented to concrete phenomena and, in particular, be able to interpret the results obtained from the application of these methods. Teaching is one of the three methodologies included in the training programme. Therefore, at the end of the course the student has acquired the tools necessary to deal with the analysis of sample data in all application contexts and is also able to interpret the results and justify the decisions to which the results obtained will lead.
Soft skills	The methodological tools of inferential analysis are constantly presented with reference to real phenomena. Therefore given the transversal nature of the discipline, through exercises and laboratory activities, the student analyzes phenomena such as social, economic, biomedical, etc., through the use of

	inferential tools and comments on the results of the analysis. Therefore, the student not only acquires the methodological skills but is also able to acquire the mental flexibility to apply the tools in a transversal way. Laboratory tests and exercises are systematically carried out in the classroom and involve the students.
Syllabus	
Content knowledge	• Introduction to the course: purposes and logic of statistical inference. • Principal elements of probability; random variables; discrete and continuous principal random variables; some limit theorems. • Introduction to statistical inference: fundamental concepts; population and sample. • Statistical inference techniques: parameter estimation, estimators and properties, some estimation methods, confidence intervals; testing of parametric and non-parametric hypotheses, Neyman-Pearson lemma, likelihood ratio. • Sample size. • Problems of inference on the parameters of a normal population: inference on the mean and on the variance. • Problems of inference on the parameters of two or more normal populations. • Comparison of the means of two populations: independent samples and dependent samples (paired data). • Comparison of the variances of two populations. • Comparison of the means of several populations (ANOVA). • The linear model: linear regression. • Other linear and non-linear models. • ANOVA in terms of linear model. • Inference on the correlation coefficient. • Problems of inference on the parameter of a dichotomous population. • Problems of inference on the parameters of two dichotomous populations. • Nonparametric inference problems. **- Insights** • Robustness and robust estimators. • Resampling methods: Jackknife and Bootstrap. • Introductory concepts of decision theory.
Texts and readings	<u>Theory</u> PICCOLO D., Statistica per le decisioni, Terza edizione, Il Mulino, 2020 (capp. 8 – 19) NEWBOLD P., CARLSON W. L., THORNE B., Statistica, seconda edizione, 2015 AGRESTI A., FINLAY B., Statistical Methods for the Social Sciences, Prentice Hall, New Jersey 2007 CHIEPPA M., RIBECCO N., VITALE C., Teoria e metodi statistici, Edizioni scientifiche italiane, Napoli 1994 G. CICCHITELLI, Statistica: Principi e Metodi- seconda edizione, Pearson, 2012 (capp.12 - 21 e Appendici B e C). FRED N., JONES S., BERGQUIST T., BONNINI S., Statistica per le scienze economiche e azienali, Isedi, Torino, 2019 (capitoli 4 - 9) FREUND R.J., WILSON W.J., Metodi statistici, Piccin, Padova 2000 PIERACCINI L., Fondamenti di inferenza statistica, G. Giappichelli, Torino 2007 <u>Applications</u> MURRAY R. SPIEGEL, Statistica, Collana Schaum, McGraw Hill Italia, Milano 2003 BERNSTEIN S., BERNSTEIN R., Statistica Inferenziale, Collana Schaum, McGraw Hill Italia, Milano 2003. <u>Laboratory</u> IACUS S.M., MASAROTTO G., Laboratorio di Statistica con R, MacGraw-Hill Informatica, 2007

	COCCARDA R., FRASCATI F., Manuale interattivo di statistica con R con MyLab, Pearson 2015 RACUGNO W., VENTURA L., 2017, Biostatistica, casi di studio in R, EGEA, Milano.
Notes, additional materials	The student is free to choose any of the texts indicated both for the theoretical and for the practical part
Repository	The didactic material related to the in-depth study is available on the e-learning platform (dief.osel.it)

Assessment	
Assessment methods	• During the course tests are carried out which have no validity for the purposes of the profit exam but have the purpose of evaluating and verifying the skills acquired by the students on the topics covered up to the moment of the test. These tests can include group work on topics of statistics-mathematics, followed by a power point presentation by the students, and individual tests in which problems are posed that require the application of inferential techniques suitable for the context. • The exam includes a written test, at the end of the course, in which the student is asked to solve problems concerning real cases using suitable statistical methods. This test is followed by an oral discussion which can be held on one of the dates indicated in the exam calendar. • Any negative outcome of the test taken at the end of the course does not affect admission to the exam, which will be carried out according to the following methods: a written test which contains questions to be resolved through the application of inferential methods and oral exam. • The oral test includes the discussion of the results obtained and the verification of knowledge on additional topics that are not covered by the written test.
Assessment criteria	The written and the oral contribute to the final evaluation. In particular, contributes to the final evaluation, the correct application of the methods to the concrete problems, that are submitted during the written, and the mental flexibility and the logical ability of the student in interpreting the results and, above all, in the presentation of the methods .
Final exam and grading criteria	The evaluation is out of thirty and the evaluation of the written and of the oral contribute equally to determining the final grade. The exam is passed when the student has achieved a minimum of 18/30.
Further information	**Statistical Laboratory** (**3 ECTS**) It is possible to acquire the 3 credits for the Statistical Laboratory within the course of Statistical Inference and Linear Models, upon request to the Coordinator of the Interclass of Statistical Sciences. To this end, students who request it will have to carry out an additional 21 hours (equal to 7 hours for ECTS) of Laboratory in R. The laboratory materials are published on the e-learning platform (dief.osel.it) which can be accessed after registration. To become eligible, you must: 1. Attend the laboratory. Those who have accumulated more than three absences will not be able to participate in the final evaluation. 2. Having achieved an average rating (multiple choice questions and R) of no less than 18/30. 3. Pass the final test of R (minimum rating 18/30) in which it will be possible to participate only if the requirements set out in points 1. and 2. are met.
	.