Syllabus

Course Description

Course Title	Statistical Methods
Course Code	27523
Course Title Additional	
Scientific-Disciplinary Sector	
Language	English
Degree Course	Master in Data Analytics for Economics and Management
Other Degree Courses (Loaned)	
Lecturers	Prof. Davide Ferrari, Davide.Ferrari2@unibz.it https://www.unibz.it/en/faculties/economics-management/academic-staff/person/39001 Prof. Alessandro Casa, Alessandro.Casa@unibz.it https://www.unibz.it/en/faculties/economics-management/academic-staff/person/46549 dr. Greta Goracci, Greta.Goracci@unibz.it https://www.unibz.it/en/faculties/economics-management/academic-staff/person/46136
Teaching Assistant	
Semester	All semesters
Course Year/s	1
CP	12
Teaching Hours	M1: - 24 hours of in-person lectures - 12 hours of video lectures (counted as 24 hours to account for re-watching) M2: - 24 hours of in-person lectures - 12 hours of video lectures (counted as 24 hours to account for

	re-watching)
Lab Hours	-
Individual Study Hours	-
Planned Office Hours	M1: 18, including 6 hours offered as virtual classroom sessions M2: 18, including 6 hours offered as virtual classroom sessions
Contents Summary	The first module (M1) provides a mathematically grounded introduction to statistical inference, with an emphasis on the theoretical principles underlying many modern statistical and data science methods. Building on basic coursework in undergraduate statistics and econometrics, the course develops a rigorous framework for understanding how statistical procedures are constructed, justified, and evaluated. Core topics include probability review, sampling distributions, point estimation, interval estimation, hypothesis testing, and likelihood-based methods. The module also introduces asymptotic theory and basic Bayesian techniques, highlighting their role in modern statistical and data science applications. Throughout, the focus is on connecting mathematical reasoning, including derivations, properties of estimators, and probabilistic arguments, to practical statistical and data science methodology. Students will develop both a deeper theoretical understanding of inference and the ability to apply these ideas in contexts relevant to data science and econometrics. The second module (M2) provides a broad introduction to statistical learning, covering the core methods and principles that underlie modern data analysis and predictive modelling. Building on foundational knowledge of statistics and regression, the course develops a unified framework for understanding supervised and unsupervised learning, with model selection serving as a connecting thread throughout. Supervised methods range from regression extensions and generalized linear models to modern classification techniques, while unsupervised learning encompasses dimensionality reduction and clustering. Practical implementation in R on real-world business datasets ensures that theoretical understanding is consistently grounded in applied data science practice.
Course Topics	M1 - Review of Probability: Probability spaces, random variables, distributions; expectation, variance, covariance; common discrete

	and continuous families. - Multivariate Distributions and Transformations: Joint, marginal, and conditional distributions; independence; functions of random variables. - Sampling Distributions: Random samples; sample mean and sample variance; normal-theory results; chi-square, t, and F distributions. - Principles of Estimation: Point estimators; bias, variance, and mean squared error; method of moments; introduction to sufficiency. - Maximum Likelihood Estimation: Likelihood functions; MLEs for common models; invariance property; numerical optimization ideas. - Properties of Estimators: Unbiasedness and consistency; efficiency; Cramér–Rao lower bound; Rao–Blackwell theorem (introductory treatment). - Confidence Intervals: Pivotal quantities; exact and approximate confidence intervals; confidence intervals based on asymptotic normality. - Hypothesis Testing: Null and alternative hypotheses; Type I and Type II errors, power; Neyman–Pearson lemma; likelihood ratio tests. - Bayesian and Likelihood Perspectives: Likelihood principle; basic Bayesian inference; posterior distributions and credible intervals; comparison with classical inference. - Asymptotic Inference: Law of large numbers; central limit theorem; delta method; asymptotic distribution of MLE. - Modern and Computational Topics: Bootstrap methods; permutation tests; Monte Carlo methods; extension of the likelihood function (penalized, M-estimation) M2 - Flexible Linear Regression: Recap of linear regression and diagnostics; polynomial regression, interactions, and categorical predictors; splines and local regression; flexibility-interpretability trade-off. - Classification: Logistic regression; linear discriminant analysis and comparison with logistic regression; quadratic discriminant analysis and k-nearest neighbours; classifier evaluation through confusion matrix, ROC curves, and AUC.

	- Model Selection and Assessment: Training and test error; bias-variance trade-off; validation set approach, k-fold and leave-one-out cross-validation; best subset and stepwise selection; information criteria. - Clustering: k-means and hierarchical clustering; linkage criteria and dendrograms; introduction to model-based clustering and mixture models; notes on density-based approaches. - Dimensionality Reduction: Principal component analysis; exploratory factor analysis as an alternative perspective; considerations on high-dimensional settings. - Generalized Linear Models: Exponential family and link functions; GLM formulation; logistic regression revisited as a GLM; Poisson regression for count data; estimation, deviance, and goodness of fit.
Keywords	M1 statistical inference, maximum likelihood, hypothesis testing, confidence intervals, Bayesian inference M2 statistical learning, model selection, classification, clustering, dimensionality reduction
Recommended Prerequisites	M1 An undergraduate-level course on statistical inference or closely related subject M2 Knowledge of probability and statistical inference, including random variables, common distributions, estimation, hypothesis testing, and confidence intervals. Familiarity with linear regression, including model specification, estimation by least squares, and interpretation of coefficients.
Propaedeutic Courses	
Teaching Format	Recorded lectures, in-person teaching, exercises. The course adopts a blended, student-centred approach that emphasises problem-based learning and active engagement. A portion of the lecture content is made available online in advance, allowing students to explore key concepts independently and at their own pace before attending class. This preparatory work enables inperson sessions to focus on the application of knowledge

	through real-world problems, collaborative activities, and guided discussions — fostering critical thinking and deeper learning. The course is fully aligned with the principles of the Italian Universities Digital Hub (EDUNEXT) initiative (https://edunext.eu), which promotes the integration of digital resources and active learning strategies within university teaching.
Mandatory Attendance	Attendance is recommended, but not mandatory.
Specific Educational Objectives and Learning Outcomes	Intended Learning Outcomes (ILO) M1: ILO 1 Knowledge and understanding: ILO 1.1 The student acquires knowledge of the analytical techniques and tools required to understand and quantitatively analyse economic and business phenomena in order to support decision-making processes. ILO 1.2 The student consolidates knowledge of statistical inference, linear models and their generalisations, linear algebra, and optimisation techniques. ILO 1.3 The student acquires an in-depth knowledge of the main techniques of supervised and unsupervised statistical learning, which are instrumental in the development of analysis and visualisation of economic and business data. ILO 2 Applying knowledge and understanding: ILO 2.1 Ability to apply and implement analysis techniques focusing on different types of datasets such as streaming data, tabular data, documents and images and analysis on joint datasets. ILO 2.2 Ability to apply supervised and unsupervised learning, and knowledge modelling, extraction, integration, analysis and exploitation; these skills are declined in various application domains of interest to companies and public and private organisations. ILO 3 Making judgements: ILO 3.1 The student acquires the ability to apply acquired knowledge to interpret data in order to make directional and operational decisions in a business context.

	ILO 3.2 The student acquires the ability to apply acquired knowledge to support processes related to production, management and risk promotion activities and investment choices through the organisation, analysis and interpretation of complex databases. ILO4 Communication skills: ILO 4.1 The student acquires the ability to communicate effectively in oral and written form the specialised content of the individual disciplines, using different registers, depending on the recipients and the communicative and didactic purposes, and to evaluate the formative effects of his/her communication. ILO 5 Learning skills: ILO 5.1 The student acquires knowledge of scientific research tools. He/she will also be able to make autonomous use of information technology to carry out bibliographic research and investigations both for his/her own training and for further education. Furthermore, through the curricular teaching and the activities related to the preparation of the final thesis, she will be able to acquire the ability - to identify thematic connections and to establish relationships between methods of analysis and application contexts; - to frame a new problem in a systematic manner and to implement appropriate analysis solutions; - to formulate general statistical-econometric models from the phenomena studied. M2: ILO 1 Knowledge and understanding: ILO 1.1 The student acquires knowledge of the analytical techniques and tools required to understand and quantitatively analyse economic and business phenomena in order to support decision-making processes. ILO 1.2 The student consolidates knowledge of statistical inference, linear models and their generalisations, linear algebra, and optimisation techniques. ILO 1.3 The student acquires an in-depth knowledge of the main

	techniques of supervised and unsupervised statistical learning, which are instrumental in the development of analysis and visualisation of economic and business data. ILO 2 Applying knowledge and understanding: ILO 2.1 Ability to apply and implement analysis techniques focusing on different types of datasets such as streaming data, tabular data, documents and images and analysis on joint datasets. ILO 2.2 Ability to apply supervised and unsupervised learning, and knowledge modelling, extraction, integration, analysis and exploitation; these skills are declined in various application domains of interest to companies and public and private organisations. ILO 3 Making judgements: ILO 3.1 The student acquires the ability to apply acquired knowledge to interpret data in order to make directional and operational decisions in a business context. ILO 3.2 The student acquires the ability to apply acquired knowledge to support processes related to production, management and risk promotion activities and investment choices through the organisation, analysis and interpretation of complex databases. ILO4 Communication skills: ILO 4.1 The student acquires the ability to communicate effectively in oral and written form the specialised content of the individual disciplines, using different registers, depending on the recipients and the communicative and didactic purposes, and to evaluate the formative effects of his/her communication. ILO 5 Learning skills: ILO 5.1 The student acquires knowledge of scientific research tools. He/she will also be able to make autonomous use of information technology to carry out bibliographic research and investigations both for his/her own training and for further education. Furthermore, through the curricular teaching and the activities related to the preparation of the final thesis, she will be

	able to acquire the ability - to identify thematic connections and to establish relationships between methods of analysis and application contexts; - to frame a new problem in a systematic manner and to implement appropriate analysis solutions; - to formulate general statistical-econometric models from the phenomena studied.
Specific Educational Objectives and Learning Outcomes (additional info.)	
Assessment	The overall exam mark will be determined by the assessment of the two modules (M1+M2). M1 Final exam (60%)(ILO 1.1–1.3, 2.1, 2.2, 4.1) and 4 homework assignments (10% each, 40% total)(ILO 2.1, 2.2, 3.1, 3.2, 5.1). Students who do not submit the homework will be assessed entirely on the final exam (100%). M2 Final exam (60%)(ILO 1.1-1.3, 2.1, 2.2, 5.1) and homeworks/data analysis assignments to be handed in (40%)(ILO 2.1, 2.2, 3.1, 3.2, 4.1). Students who do not complete the homeworks/assignment will be assessed entirely on the final exam (100%)
Evaluation Criteria	In both modules the exam modalities are the same for the attending and the non-attending students. Project work (40% of the final grade) and written exam (60% of the final grade). - Written exam: M1: understanding of probabilistic and inferential concepts, correct derivation and interpretation of estimators and tests, clarity and precision of explanations M2: understanding of statistical concepts, correct interpretation of the results of statistical analyses, clarity and precision of explanations - Project work: clarity of presentation, ability to gain useful and novel insights from data, creativity and critical thinking, adherence to reproducible research practices, proficiency in using R and other software to perform data preparation, statistical analyses, and

	graphical representations appropriate to the data at hand.
Required Readings	**M1** • Rice, J. A. (2007) *Mathematical Statistics and Data Analysis*. Cengage Learning. • Hogg, R. V., McKean, J. W., and Craig, A. T. (2019) *Introduction to Mathematical Statistics*. Pearson. • Casella, G., & Berger, R. (2024). *Statistical inference*. Chapman and Hall/CRC. **M2** • Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning.* Springer • James, G., Witten, D., Hastie, T., Tibshirani, R. (2013) *An Introduction to Statistical Learning with Applications in R*. Springer.
Supplementary Readings	**M2** • Bishop, C. M. (2006). *Pattern recognition and machine learning.* New York: Springer. • Slides and lecture notes provided
Further Information	M1 & M2 - Software Students may use R, Python or another computer program of their choice for computational assignments. Example code and demonstrations may be provided in R.
Sustainable Development Goals (SDGs)	Quality education

Course Module

Course Constituent Title	M1 - Statistical Inference
Course Code	27523A
Scientific-Disciplinary Sector	STAT-01/A
Language	English

Lecturers	Prof. Davide Ferrari, Davide.Ferrari2@unibz.it https://www.unibz.it/en/faculties/economics-management/academic-staff/person/39001 dr. Greta Goracci, Greta.Goracci@unibz.it https://www.unibz.it/en/faculties/economics-management/academic-staff/person/46136
Teaching Assistant	
Semester	First semester
CP	6
Responsible Lecturer	
Teaching Hours	- 24 hours of in-person lectures - 12 hours of video lectures (counted as 24 hours to account for re-watching)
Lab Hours	-
Individual Study Hours	-
Planned Office Hours	18
Contents Summary	This module provides a mathematically grounded introduction to statistical inference, with an emphasis on the theoretical principles underlying many modern statistical and data science methods. Building on basic coursework in undergraduate statistics and econometrics, the course develops a rigorous framework for understanding how statistical procedures are constructed, justified, and evaluated. Core topics include probability review, sampling distributions, point estimation, interval estimation, hypothesis testing, and likelihood-based methods. The module also introduces asymptotic theory and basic Bayesian techniques, highlighting their role in modern statistical and data science applications. Throughout, the focus is on connecting mathematical reasoning, including derivations, properties of estimators, and probabilistic arguments, to practical statistical and data science methodology. Students will develop both a deeper theoretical understanding of inference and the ability to apply these ideas in contexts relevant to data science and econometrics.
Course Topics	- Review of Probability: Probability spaces, random variables, distributions; expectation, variance, covariance; common discrete

	and continuous families. - Multivariate Distributions and Transformations: Joint, marginal, and conditional distributions; independence; functions of random variables. - Sampling Distributions: Random samples; sample mean and sample variance; normal-theory results; chi-square, t, and F distributions. - Principles of Estimation: Point estimators; bias, variance, and mean squared error; method of moments; introduction to sufficiency. - Maximum Likelihood Estimation: Likelihood functions; MLEs for common models; invariance property; numerical optimization ideas. - Properties of Estimators: Unbiasedness and consistency; efficiency; Cramér–Rao lower bound; Rao–Blackwell theorem (introductory treatment). - Confidence Intervals: Pivotal quantities; exact and approximate confidence intervals; confidence intervals based on asymptotic normality. - Hypothesis Testing: Null and alternative hypotheses; Type I and Type II errors, power; Neyman–Pearson lemma; likelihood ratio tests. - Bayesian and Likelihood Perspectives: Likelihood principle; basic Bayesian inference; posterior distributions and credible intervals; comparison with classical inference. - Asymptotic Inference: Law of large numbers; central limit theorem; delta method; asymptotic distribution of MLE. - Modern and Computational Topics: Bootstrap methods; permutation tests; Monte Carlo methods; extension of the likelihood function (penalized, M-estimation)
Teaching Format	Recorded lectures, in-person teaching, exercises. The module adopts a blended, student-centred approach that emphasises problem-based learning and active engagement. A portion of the lecture content is made available online in advance, allowing students to explore key concepts independently and at their own pace before attending class. This preparatory work enables inperson sessions to focus on the application of knowledge through real-world problems, collaborative activities, and guided discussions — fostering critical thinking and deeper learning. The module is fully aligned with the principles of the Italian Universities

	Digital Hub (EDUNEXT) initiative (https://edunext.eu), which promotes the integration of digital resources and active learning strategies within university teaching.
Required Readings	Rice, J. A. *Mathematical Statistics and Data Analysis*. Hogg, R. V., McKean, J. W., and Craig, A. T. *Introduction to Mathematical Statistics*. Casella, G., and Berger, R. L. *Statistical Inference*
Supplementary Readings	

Course Module

Course Constituent Title	M2 - Statistical and Machine Learning Methods for Business Analysis
Course Code	27523B
Scientific-Disciplinary Sector	STAT-01/A
Language	English
Lecturers	Prof. Alessandro Casa, Alessandro.Casa@unibz.it https://www.unibz.it/en/faculties/economics-management/academic-staff/person/46549
Teaching Assistant	
Semester	Second semester
CP	6
Responsible Lecturer	
Teaching Hours	- 24 hours of in-person lectures - 12 hours of video lectures (counted as 24 hours to account for re-watching)
Lab Hours	-
Individual Study Hours	-
Planned Office Hours	18
Contents Summary	This module provides a broad introduction to statistical learning, covering the core methods and principles that underlie modern data analysis and predictive modelling. Building on foundational knowledge of statistics and regression, the course develops a unified framework for understanding supervised and unsupervised

	learning, with model selection serving as a connecting thread throughout. Supervised methods range from regression extensions and generalized linear models to modern classification techniques, while unsupervised learning encompasses dimensionality reduction and clustering. Practical implementation in R on real-world business datasets ensures that theoretical understanding is consistently grounded in applied data science practice.
Course Topics	- Flexible Linear Regression: Recap of linear regression and diagnostics; polynomial regression, interactions, and categorical predictors; splines and local regression; flexibility-interpretability trade-off. - Classification: Logistic regression; linear discriminant analysis and comparison with logistic regression; quadratic discriminant analysis and k-nearest neighbours; classifier evaluation through confusion matrix, ROC curves, and AUC. - Model Selection and Assessment: Training and test error; bias-variance trade-off; validation set approach, k-fold and leave-one-out cross-validation; best subset and stepwise selection; information criteria. - Clustering: k-means and hierarchical clustering; linkage criteria and dendrograms; introduction to model-based clustering and mixture models; notes on density-based approaches. - Dimensionality Reduction: Principal component analysis; exploratory factor analysis as an alternative perspective; considerations on high-dimensional settings. - Generalized Linear Models: Exponential family and link functions; GLM formulation; logistic regression revisited as a GLM; Poisson regression for count data; estimation, deviance, and goodness of fit.
Teaching Format	Recorded lectures, in-person teaching, exercises. The module adopts a blended, student-centred approach that emphasises problem-based learning and active engagement. A portion of the lecture content is made available online in advance, allowing students to explore key concepts independently and at their own pace before attending class. This preparatory work enables inperson sessions to focus on the application of knowledge through real-world problems, collaborative activities, and guided discussions — fostering critical thinking and deeper learning. The module is fully aligned with the principles of the Italian Universities

	Digital Hub (EDUNEXT) initiative (https://edunext.eu), which promotes the integration of digital resources and active learning strategies within university teaching.
Required Readings	• Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning.* Springer • James, G., Witten, D., Hastie, T., Tibshirani, R. (2013) *An Introduction to Statistical Learning with Applications in R*. Springer.
Supplementary Readings	• Bishop, C. M. (2006). *Pattern recognition and machine learning.* New York: Springer. • Slides and lecture notes provided