Teaching Guide				
Identifying Data				2021/22
Subject (*)	Nonparametric Methods		Code	614493111
Study programme	Mestrado Universitario en Técnicas Estadísticas (Plan 2019)			
Descriptors				
Cycle	Period	Year	Type	Credits
Official Master's Degree	1st four-month period	First	Obligatory	5
Language				
Teaching method	Hybrid			
Prerequisites				
Department	Matemáticas			
Coordinador	Vilar Fernandez, Jose Antonio	E-mail	jose.vilarf@udc.es	
Lecturers	Vilar Fernandez, Jose Antonio	E-mail	jose.vilarf@udc.es	
Web	http://eio.usc.es/pub/mte/			
General description	Nonparametric methods to estimate the probability distribution, probability density and regression functions are introduced, paying sparticular attention to the kernel smoothing techniques. The main nonparametric goodness-of-fit test procedures, association tests in contingency tables and nonparametric rank-based location tests for one, two and more than two samples are also presented.			
Contingency plan	1. Modifications to the contents The contents will not be modified. 2. Methodologies All teaching methodologies will be maintained. 3. Mechanisms for personalized attention to students. E-mail and telematic resources (Teams) will be employed. 4. Modifications in the evaluation Both continuous evaluation activities and exams can be developed by using telematic means. 5. Modifications to the bibliography or webgraphy No modifications.			

Study programme competences	
Code	Study programme competences
A16	CE1 - Coñecer, identificar, modelar, estudar e resolver problemas complexos de estatística e investigación operativa, nun contexto científico, tecnolóxico ou profesional, xurdidos en aplicacións reais.
A17	CE2 ? Desenvolver autonomía para a resolución práctica de problemas complexos surdidos en aplicación reais e para a interpretación dos resultados cara á axuda na toma de decisións.
A18	CE3 - Adquirir coñecementos avanzados dos fundamentos teóricos subxacentes ás distintas metodoloxías da estatística e a investigación operativa, que permitan o seu desenvolvemento profesional especializado.
A19	CE4 - Adquirir as destrezas necesarias no manexo teórico-práctico da teoría de probabilidade e as variables aleatorias que permitan o seu desenvolvemento profesional no eido científico/académico, tecnolóxico ou profesional especializado e multidisciplinar.
A20	CE5 - Profundizar no coñecemento dos fundamentos teórico-prácticos especializados de modelado e estudo de distintos tipos de relacións de dependencia entre variables estatísticas.
A21	CE6 - Adquirir coñecementos teórico-prácticos avanzados de distintas técnicas matemáticas, orientadas específicamente á axuda na toma de decisións, e desenvolver a capacidade de reflexión para avaliar e decidir entre distintas perspectivas en contextos complexos.
A23	CE8 - Adquirir coñecementos teórico-prácticos avanzados das técnicas destinadas á realización de inferencias e contrastes relativos a variables e parámetros dun modelo estatístico, e saber aplicalos con autonomía suficiente nun contexto científico, tecnolóxico ou profesional.

B1	CB6 - Posuír e comprender coñecementos que acheguen unha base ou oportunidade de ser orixinais no desenvolvemento e/ou aplicación de ideas, a miúdo nun contexto de investigación
B2	CB7 - Que os estudantes saiban aplicar os coñecementos adquiridos e a súa capacidade de resolución de problemas en ámbitos novos ou pouco coñecidos dentro de contextos máis amplos (ou multidisciplinares) relacionados coa súa área de estudo
B3	CB8 - Que os estudantes sexan capaces de integrar coñecementos e enfrontarse á complexidade de formular xuízos a partir dunha información que, sendo incompleta ou limitada, inclúa reflexións sobre as responsabilidades sociais e éticas vinculadas á aplicación dos seus coñecementos e xuízos
B4	CB9 - Que os estudantes saiban comunicar as súas conclusións e os coñecementos e razóns últimas que as sustentan a públicos especializados e non especializados dun modo claro e sen ambigüidades
B5	CB10 - Que os estudantes posúan as habilidades de aprendizaxe que lles permitan continuar estudando dun modo que haberá de ser en gran medida autodirixido ou autónomo.
B17	CG1 - Coñecer, comprender e saber aplicar os principios, metodoloxías e novas tecnoloxías na estatística e a investigación operativa en contextos científico/académicos, tecnolóxicos ou profesionais especializados e multidisciplinares, así como adquirir as destrezas e competencias descritas nos objectivos generales do título.
B18	CG2 - Desenvolver autonomía para identificar, modelar e resolver problemas complexos da estatística e da investigación operativa en contextos científico/académicos, tecnolóxicos ou profesionais especializados e multidisciplinares.
B19	CG3 - Desenvolver a capacidade para realizar estudos e tarefas de investigación e transmitir os resultados a públicos especializados, académicos e xeneralistas.
B20	CG4 - Integrar coñecementos avanzados e enfrontarse á toma de decisións a partir de información científica e técnica.
B21	CG5 - Desenvolver a capacidade de aplicación de algoritmos e técnicas de resolución de problemas complexos no eido da estatística e a investigación operativa, manexando o software especializado axeitado.
C11	CT1 - Desenvolver firmes capacidades de razoamento, análise crítica e autocrítica, así como de argumentación e de síntese, contextos especializados e multidisciplinais.
C13	CT3 - Ser capaz de resolver problemas complexos en novos escenarios mediante a aplicación integrada dos coñecementos.
C14	CT4 - Desenvolver unha sólida capacidade de organización e planificación do estudo, asumindo a responsabilidade do seu propio desenvovemento profesional, para a realización de traballos en equipo e de xeito autónomo.
C15	CT5 - Desenvolver capacidades para o aprendizaxe e a integración no traballo en equipos multidisciplinais, nos ámbitos científico/académico, tecnolóxico e profesional.

Learning outcomes			
Learning outcomes	Study programme competences		
Get thorough knowledge about strengths and weaknesses of the nonparametric approach in data analysis.			
To know how present data analysis based on nonparametric techniques to both specialized and non-specialized audience.			
To become familiar with basic techniques of nonparametric estimation of the probability distribution function, the probability density function and the regression function.			
Get the know-how to apply the main nonparamteric tests for goodness-of-fit and association.			
Get thorough knowledge about strengths and weaknesses of the nonparametric approach in data analysis.			
Develop autonomus competence to apply nonparametric tools in data analysis, in complex and/or multidisciplinary scenarios.			
To know how present data analysis based on nonparametric techniques to both specialized and non-specialized audience.			
Develop autonomus competence to apply nonparametric tools in data analysis, in complex and/or multidisciplinary scenarios.			
To become familiar with basic techniques of nonparametric estimation of the probability distribution function, the probability density function and the regression function.			
Get the know-how to apply the main nonparamteric tests for goodness-of-fit and association.			
To become familiar with basic techniques of nonparametric estimation of the probability distribution function, the probability density function and the regression function.			
Get the know-how to apply the main nonparamteric tests for goodness-of-fit and association.			
Get thorough knowledge about strengths and weaknesses of the nonparametric approach in data analysis.			
Develop autonomus competence to apply nonparametric tools in data analysis, in complex and/or multidisciplinary scenarios.			
To know how present data analysis based on nonparametric techniques to both specialized and non-specialized audience.			

To become familiar with basic techniques of nonparametric estimation of the probability distribution function, the probability density function and the regression function.	AC18 AC19 AC20 AC21 AC23	BJ1 BJ3 BJ5 BJ20 BJ21	CJ13
Get the know-how to apply the main nonparamteric tests for goodness-of-fit and association.	AC18 AC19 AC20 AC21 AC23	BJ1 BJ3 BJ5 BJ20 BJ21	CJ13
Get thorough knowledge about strengths and weaknesses of the nonparametric approach in data analysis.	AC16 AC17 AC19 AC21 AC23	BJ2 BJ17 BJ20 BJ21	CJ11 CJ13
Develop autonomus competence to apply nonparametric tools in data analysis, in complex and/or multidisciplinary scenarios.	AC17	BJ18	CJ14 CJ15
To know how present data analysis based on nonparametric techniques to both specialized and non-specialized audience.		BJ4 BJ19	

Contents	
Topic	**Sub-topic**
Nonparametric distribution estimation	The empirical distribution. Properties. Moments and quantiles estimation.
Classical one-sample nonparametric tests.	Goodness-of-fit tests: Kolmogorov-Smirnov test. Normality analysis: Q-Q plot, Lilliefors test, Shapiro-Wilk test, transformations for normality. One-sample location tests: sign test, Wilcoxon signed-rank test.
Two-sample tests.	Two-sample comparison: Kolmogorv-Smirnov test for two-samples, Mann-Whitney-Wilcoxon test. Extensions for three or more samples: Kruskal-Wallis test, Friedman test.
Tests based on contingency tables.	Contingency tables analysis. Chi-squared tests for goodness-of-fit, homogeneity and independence on contingency tables.
Smoothing methods: nonparametric density estimation.	The histogram. Kernel density estimation. Assessment of density estimators. Smoothing parameter selectors in kernel density estimation: cross-validation and plug-in approaches. Multivariate kernel density estimation.
Nonparametric regression estimation.	Kernel regression. Local polynomial regression. k-nearest neighbor regression. Smoothing parameter selectors in kernel regression estimation: cross-validation and plug-in approaches. Loess algorithm. Spline regression: a brief introduction.

Planning				
Methodologies / tests	**Competencies**	**Ordinary class hours**	**Student?s personal work hours**	**Total hours**
Guest lecture / keynote speech	A16 A18 A21 A23 B1 B3 B17 B20 C11 C13	20	15	35
Seminar	A16 A17 A19 A20 A21 A23 B2 B3 B5 B17 B18 B19 B20 B21 C11 C13	7	5.25	12.25

ICT practicals	A17 A19 A20 A21 A23 B2 B5 B18 B19	7	5.25	12.25
Problem solving	A16 A17 A19 A23 B2 B3 B5 B18 B19 B20 B21 C11 C13 C14 C15	0	28.5	28.5
Case study	A16 A17 A19 A21 A23 B2 B3 B4 B5 B18 B19 B20 B21 C11 C13 C14 C15	0	21	21
Supervised projects	A17 A19 A21 A23 B2 B4 B5 B18 B19 B20 B21 C11 C13 C14 C15	0	9.5	9.5
Workshop	A16 A17 B2 B3 B4 B17 B18 B19 C11 C13 C14 C15	1	2.5	3.5
Objective test	A16 A17 A18 A19 A20 A21 A23 B20 B21 C11 C13	0	3	3
Personalized attention		0	0	0
(*)The information in the planning table is for guidance only and does not take into account the heterogeneity of the students.				

Methodologies	
Methodologies	**Description**
Guest lecture / keynote speech	The theoretical principles of the nonparametric tools and procedures for their application in practice will be introduced. Their interest in applications will be illustrated by using specific real examples from different disciplines, highlighting advantages and limitations. Students participation will be strongly encouraged.
Seminar	Specific problems and suitable approaches to get them solved will be presented in seminars. The main objective is to show how the concepts and algorithms exposed in the keynote speechs are useful to face these problems.
ICT practicals	Interactive sessions addressed to solve specific exercises by using scripts with free code from R software. The lecturer will support and supervise the right application of the knowledge and skills gathered by the students.
Problem solving	Issues, exercises and examples that can be addressed by using nonparametric techniques of inference and modeling will be provided to be individually solved by students.
Case study	Specific study cases will be proposed to be solved in group and/or individually.
Supervised projects	Solutions for exercises and study case will be supervised by the lecturer.
Workshop	Case study analyzed in detail by students will be presented and discussed.
Objective test	Written examination to assess the the acquisition of knowledge.

Personalized attention	
Methodologies	**Description**

ICT practicals Supervised projects	The ICT practicals will be conducted to solve exercises by using scripts with free code from R software. This way, students must thoroughly understand the used R-packages, particularly the main functions and the type of generated outputs. To reach this objective as soon as possible, personalized attention is desirable and will be provided during the session. Every student must complete, properly supervised, a specific project involving real data and using techniques and skills developed throughout the course. Personalized attention will consist in monitoring the different stages of the project at succesive working meetings.

Assessment			
Methodologies	Competencies	Description	**Qualification**
Workshop	A16 A17 B2 B3 B4 B17 B18 B19 C11 C13 C14 C15	The defence in the workshop of a supervised work is worth 5% of the global mark.	5
Supervised projects	A17 A19 A21 A23 B2 B4 B5 B18 B19 B20 B21 C11 C13 C14 C15	A detailed development of the proposed study case, completed either individually or in group, is worth up to 10% of the global qualification.	10
Objective test	A16 A17 A18 A19 A20 A21 A23 B20 B21 C11 C13	The final objective test is a written exam consisting of several theoretical-practical questions on the key contents of the subject, including proper interpretation of the results obtained from the R software. This exam is mandatory and the attained mark is worth up to 70% of the global qualification.	70
Problem solving	A16 A17 A19 A23 B2 B3 B5 B18 B19 B20 B21 C11 C13 C14 C15	Solving and timely delivery of exercises proposed during the course will be part of the continuous evaluation. Correct answers in this item is worth up to 7.5% of the global qualification.	7.5
Case study	A16 A17 A19 A21 A23 B2 B3 B4 B5 B18 B19 B20 B21 C11 C13 C14 C15	Solving and timely delivery of case study proposed during the course will be part of the continuous evaluation. Correct answers in this item is worth up to 7.5% of the global qualification.	7.5

Assessment comments
Presentación á avaliación: Considérase que un estudante concurre a unha convocatoria cando participa en actividades que lle permiten obter cando menos un 50% da avaliación final. A cualificación obtida conservarase entre as oportunidades (ordinaria e extraordinaria) dentro da convocatoria de cada curso. Oportunidade extraordinaria de recuperación (proba de xullo): O peso da avaliación continua na oportunidade extraordinaria de recuperación (proba de xullo) será o mesmo que na avaliación ordinaria. Na segunda oportunidade de avaliación (recuperación), realizarase un exame e a nota final será o máximo de tres cantidades: a nota da avaliación ordinaria, a nota do novo exame e a media ponderada do novo exame e a avaliación continua.

Sources of information

Basic	Fan J., Gijbels I. (1996) Local polynomial modelling and its applications. Monographs on Statistics and Applied Probability 66. Chapman & Hall/CRC.Gibbons J.D, Chakraborti S. (2010) Nonparametric Statistical Inference (5th edition). Statistics: Textbooks and Monographs. Chapman & Hall/CRC.Hollander M., Wolfe D.A., Chicken E. (2014) Nonparametric Statistical Methods (3rd edition). Wiley Series in Probability and Statistics, Wiley.Silverman, B. W. (1986) Density Estimation for Statistics and Data Analysis. Monographs on Statisticsand Applied Probability 26. Chapman & Hall/CRC. Wand M.P., Jones M.C. (1995) Kernel smoothing. Monographs on Statistics and Applied Probability 60. Chapman & Hall/CRC. Fan J., Gijbels I. (1996) Local polynomial modelling and its applications. Monographs on Statistics and Applied Probability 66. Chapman & Hall/CRC.Gibbons J.D, Chakraborti S. (2010) Nonparametric Statistical Inference (5th edition). Statistics: Textbooks and Monographs. Chapman & Hall/CRC.Hollander M., Wolfe D.A., Chicken E. (2014) Nonparametric Statistical Methods (3rd edition). Wiley Series in Probability and Statistics, Wiley.Silverman, B. W. (1986) Density Estimation for Statistics and Data Analysis. Monographs on Statisticsand Applied Probability 26. Chapman & Hall/CRC. Wand M.P., Jones M.C. (1995) Kernel smoothing. Monographs on Statistics and Applied Probability 60. Chapman & Hall/CRC.
Complementary	Bowman A.W., Azzalini A. (1997) Applied Smoothing Techniques for Data Analysis. Oxford: Oxford University Press.McKean J.W., Kloke J. (2014) Nonparametric Statistical Methods using R. The R Series. Chapman and Hall/CRC. Simonoff J.S. (1996) Smoothing methods in statistics. Springer Series in Statistics. New York: Springer.Smeeton N.C, Sprent P. (2007) Applied Nonparametric Statistical Methods (4th edition) Chapman & Hall/CRC Texts in Statistical Science. Chapman & Hall/CRC.Wasserman L. (2006) All of Nonparametric Statistics. Springer Texts in Statistics. New York: Springer.Bowman A.W., Azzalini A. (1997) Applied Smoothing Techniques for Data Analysis. Oxford: Oxford University Press.McKean J.W., Kloke J. (2014) Nonparametric Statistical Methods using R. The R Series. Chapman and Hall/CRC. Simonoff J.S. (1996) Smoothing methods in statistics. Springer Series in Statistics. New York: Springer.Smeeton N.C, Sprent P. (2007) Applied Nonparametric Statistical Methods (4th edition) Chapman & Hall/CRC Texts in Statistical Science. Chapman & Hall/CRC.Wasserman L. (2006) All of Nonparametric Statistics. Springer Texts in Statistics. New York: Springer.

Recommendations
Subjects that it is recommended to have taken before
Subjects that are recommended to be taken simultaneously
Subjects that continue the syllabus
Other comments
It is advisable that students have basic notions on probability calculus and statistic. Also it is desirable to posses regular skills to manage computers, and particularly knowledge of statistical software. To be able of understanding the practical sense of the learnt methods will allow to improve the learning process.