



## Teaching Guide

Identifying Data					2023/24
Subject (*)	Statistical Analysis of Dependent Data		Code	614G02022	
Study programme	Grao en Ciencia e Enxeñaría de Datos				
Descriptors					
Cycle	Period	Year	Type	Credits	
Graduate	1st four-month period	Third	Obligatory	6	
Language	Spanish				
Teaching method	Face-to-face				
Prerequisites					
Department	Matemáticas				
Coordinador	Aneiros Perez, German	E-mail	german.aneiros@udc.es		
Lecturers	Aneiros Perez, German Fernández Casal, Rubén	E-mail	german.aneiros@udc.es ruben.fcasal@udc.es		
Web					
General description	Modelización de datos que presentan dependencia temporal e/ou espacial. Utilizaranse modelos estatísticos paramétricos. A partir de devanditos modelos, obteranse predicións de valores en novos instantes temporais ou posicións espaciais. Farase uso do software R.				

## Study programme competences

Code	Study programme competences
A17	CE17 - Capacidade para a construción, validación e aplicación dun modelo estocástico dun sistema real a partir dos datos observados e a análise crítica dos resultados obtidos.
A19	CE19 - Capacidade para comprender, expor, formular e resolver problemas susceptibles de ser abordados a través de modelos estatísticos para datos que presentan dependencia.
A20	CE20 - Coñecemento das ferramentas informáticas no campo da análise dos datos e modelización estatística, e capacidade para seleccionar as máis adecuadas para a resolución de problemas.
B2	CB2 - Que os estudantes saiban aplicar os seus coñecementos ao seu traballo ou vocación dunha forma profesional e posúan as competencias que adoitan demostrarse por medio da elaboración e defensa de argumentos e a resolución de problemas dentro da súa área de estudo
B3	CB3 - Que os estudantes teñan a capacidade de reunir e interpretar datos relevantes (normalmente dentro da súa área de estudo) para emitir xuízos que inclúan unha reflexión sobre temas relevantes de índole social, científica ou ética
B4	CB4 - Que os estudantes poidan transmitir información, ideas, problemas e solucións a un público tanto especializado como non especializado
B7	CG2 - Elaborar adecuadamente e con certa orixinalidade composicións escritas ou argumentos motivados, redactar plans, proxectos de traballo, artigos científicos e formular hipóteses razoables.
B8	CG3 - Ser capaz de manter e estender formulacións teóricas fundadas para permitir a introdución e explotación de tecnoloxías novas e avanzadas no campo.
B9	CG4 - Capacidade para abordar con éxito todas as etapas dun proxecto de datos: exploración previa dos datos, preprocesado, análise, visualización e comunicación de resultados.
B10	CG5 - Ser capaz de traballar en equipo, especialmente de carácter multidisciplinar, e ser hábiles na xestión do tempo, persoas e toma de decisións.
C1	CT1 - Utilizar as ferramentas básicas das tecnoloxías da información e as comunicacións (TIC) necesarias para o exercicio da súa profesión e para a aprendizaxe ao longo da súa vida.
C4	CT4 - Valorar a importancia que ten a investigación, a innovación e o desenvolvemento tecnolóxico no avance socioeconómico e cultural da sociedade.

## Learning outcomes

Learning outcomes	Study programme competences



Coñecer e comprender conceptos xerais relativos a procesos estocásticos.	A17 A19	B4 B8	
Identificar e analizar modelos estatísticos susceptibles de xerar a un conxunto de datos dependentes.	A17 A19	B2 B4 B9 B10	C4
Coñecer e aplicar técnicas de estimación dos parámetros presentes nos modelos estatísticos con datos dependentes.	A17 A19	B2 B4 B9 B10	C4
Entender a importancia de levar a cabo unha diagnose dun modelo construído con datos que presenten dependencia temporal e/ou espacial.	A17 A19	B4	C4
Coñecer e aplicar os fundamentos da predición en series de tempo.	A17 A19	B2 B4 B7 B9 B10	C4
Ser capaz de interpretar os modelos propostos e os resultados obtidos ao utilizar técnicas estatísticas para datos dependentes.		B3 B4 B7 B9 B10	C4
Saber manexar con soltura programas informáticos avanzados de análise estatística.	A20		C1

Contents	
Topic	Sub-topic
1. Análise descritiva dunha serie de tempo	1.1 Introducción 1.2 Descomposición dunha serie de tempo
2. Series de tempo e procesos estocásticos	2.1 Introducción 2.2 Procesos estocásticos: concepto e definicións asociadas
3. Metodoloxía Box- Jenkins	3.1 Introducción 3.2 Modelización ARIMA e predición
4. Tópicos adicionais	Introducción a análise de intervención, valores atípicos, regresión con series de tempo, e cluster e clasificación con series de tempo
5. Estatística Espacial	5.1 Tipos de procesos espaciais 5.2 Análise exploratorio de datos espaciais 5.3 Modelado de procesos xeoestatísticos 5.4 Predición Kriging 5.5 Introducción á Estatística espazo-temporal

Planning				
Methodologies / tests	Competencies	Ordinary class hours	Student?s personal work hours	Total hours
Guest lecture / keynote speech	A17 A19 B3 B4 B8 B9 C4	21	42	63
ICT practicals	A20 B2 B4 B7 B9 B10 C1 C4	14	42	56
Seminar	B7 B8 C4	7	14	21



Mixed objective/subjective test	A17 A19 B2 B3 B8 B9 C4	2	0	2
Personalized attention		8	0	8
(*)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	O profesor, coa axuda dos medios audiovisuais pertinentes, expoñerá os contidos teórico-prácticos da materia. O material empregado estará ao dispor dos estudantes.
ICT practicals	Trátase de poñer en práctica a metodoloxía estatística aprendida nas sesións maxistras. Farase uso do software R (gratuíto).  Faránselle chegar ao estudante distintos conxuntos de datos reais para ser analizados ao longo das sesións de prácticas. Á súa vez, recomendaráselle que obteña outros datos pola súa conta para que reforce a súa destreza na análise de datos con dependencia, esta vez nas horas dedicadas ao traballo autónomo.
Seminar	O seu obxectivo é apontoar, a través da realización de exercicios teórico-prácticos, a comprensión dos contidos expostos nas sesións maxistras. Nalgunhas ocasións, poderían dedicarse a completar as prácticas a través de TIC.  Faránselle chegar ao estudante distintos exercicios para que trate de resolvelos (algúns deles ao longo do propio seminario e outros nas horas dedicadas ao traballo autónomo).
Mixed objective/subjective test	Proba tipo test a través da cal se avaliará o grao de aprendizaxe alcanzada polo estudante.

Personalized attention	
Methodologies	Description
Guest lecture / keynote speech ICT practicals Seminar	Ao longo de calquera das actividades nas que o profesor estea presente (sesión maxistral, prácticas a través de TIC, seminario), o estudante ten a oportunidade de expoñerlle as dúbidas que lle xurdan acerca dos contidos que se están impartindo ou levando á práctica. Se a resolución das devanditas dúbidas implica unha clara perda de continuidade da actividade que se está levando a cabo, ou ben se as dúbidas xórdenlle ao estudante en horas non presenciais, pode facer uso das titorías individualizadas para que lle sexan resoltas por parte do profesor.

Assessment			
Methodologies	Competencies	Description	Qualification
Mixed objective/subjective test	A17 A19 B2 B3 B8 B9 C4	Consistirá nunha proba tipo test de coñecementos teórico-prácticos, que será realizada na data oficial aprobada pola Xunta de Facultade.	60
ICT practicals	A20 B2 B4 B7 B9 B10 C1 C4	Parte do grao de aprendizaxe alcanzada nas prácticas a través de TIC avaliarase de maneira continua. Para iso, realizaranse unha ou dúas probas ao longo do curso (no horario establecido para ditas prácticas) e/ou un ou dous traballos (individuais ou en grupo). A súa resolución requirirá do uso do software R e tamén do coñecemento da metodoloxía estatística explicada nas sesións maxistras.	40

Assessment comments



Primeira oportunidade: Si se denotan a través de P e F as notas (sobre 10) obtidas nas "Prácticas a través de TIC" e na "Proba mixta", respectivamente, a nota final será:

(a)  $0.4 \cdot P + 0.6 \cdot F$ , sempre e cando tanto P como F sexan maiores ou iguais que 3. (b)  $4.5 \cdot (0.4 \cdot P + 0.6 \cdot F) / 7.2$ , sempre e cando P e/ou F sexan menores que 3. Como consecuencia do anterior, tense que para superar a materia será necesario obter un mínimo de 3 puntos tanto nas "Prácticas a través de TIC" como na "Proba mixta". Avisarase da data en que se realizará cada proba práctica cunha antelación mínima de 7 días. Os traballos prácticos terán un prazo de entrega de polo menos 7 días. Segunda oportunidade: Por defecto, a puntuación obtida nas "Prácticas a través de TIC" na primeira oportunidade manterase para esta segunda oportunidade. Pero si o alumno así o desexa, terá a opción de volver a facer a(s) proba(s)/traballo(s) correspondente(s) a a(s) parte(s) de prácticas que considere, e as novas puntuacións obtidas reemplazarán ás anteriores. Todos os alumnos deben realizar a "Proba mixta". A nota final obterase da mesma forma que na primeira oportunidade.

## Sources of information

<b>Basic</b>	<ul style="list-style-type: none"> <li>- Bivand R.S., Pebesma E.J., Gómez-Rubio V. (2008). Applied Spatial Data Analysis with R. Springer</li> <li>- Clark, I. (1979). Practical geostatistics. London: Applied Science Publishers</li> <li>- Cowpertwait P.S.P., Metcalfe A.V. (2009). Introductory Time Series with R. Springer</li> <li>- Cryer J.D., Chan K-S. (2008). Time Series Analysis. With Applications in R. Springer (2ª edición)</li> <li>- Diggle P., Ribeiro P.J. (2007). Model-based Geostatistics. Springer</li> <li>- Hyndman R.J., Athanasopoulos G. (2018). Forecasting: Principles and Practice. O Texts (2ª edición) (accesible online en <a href="https://otexts.com/fpp2/">https://otexts.com/fpp2/</a>)</li> <li>- Montero P., Vilar J.A. (2014). TSclust: An R package for time series clustering. Journal of Statistical Software 62(1)</li> </ul>
<b>Complementary</b>	<ul style="list-style-type: none"> <li>- Brockwell P.J., Davis R.A. (2002). Introduction to Time Series and Forecasting. Springer (2ª edición)</li> <li>- Chilès J.P., Delfiner P. (1999). Geostatistics: modeling spatial uncertainty. Wiley</li> <li>- Liao T.W. (2005). Clustering of time series-a survey. Pattern Recognition 38, 1857-1874</li> <li>- Peña D. (2005). Análisis de Series Temporales. Alianza Editorial</li> <li>- Shumway R.H., Stoffer D.S. (2017). Time Series Analysis and Its Applications. With R Examples. Springer (4ª edición)</li> <li>- Wikle C.K., Zammit-Mangion A., Cressie N. (2019). Spatio-temporal Statistics with R. Chapman and Hall/CRC (accesible online en <a href="https://spacetimewithr.org">https://spacetimewithr.org</a>)</li> </ul>

## Recommendations

### Subjects that it is recommended to have taken before

Regression Models/614G02012

Statistical Modeling of High Dimensional Data/614G02013

Statistical Inference/614G02007

Probability and Basic Statistics/614G02003

### Subjects that are recommended to be taken simultaneously

### Subjects that continue the syllabus

Simulation and Resampling Techniques/614G02036

Statistical Analysis of Complex Data/614G02031

### Other comments



-Usarase linguaxe non sexista, utilizarase

bibliografía de autores/as de ambos sexos, propiciarase a intervención en clase de alumnos e alumnas...

-Traballarse para identificar e modificar

prexuízos e actitudes sexistas, e influirase na contorna para modificalos e fomentar valores de respecto e igualdade.-Tratarase de detectar situacións de discriminación por razón de xénero e proporanse accións e medidas para corrixilas.

**(\*)The teaching guide is the document in which the URV publishes the information about all its courses. It is a public document and cannot be modified. Only in exceptional cases can it be revised by the competent agent or duly revised so that it is in line with current legislation.**