Teaching Guide				
Identifying Data				2022/23
Subject (*)	Simulation and Resampling Techniques		Code	614G02036
Study programme	Grao en Ciencia e Enxeñaría de Datos			
Descriptors				
Cycle	Period	Year	Type	Credits
Graduate	1st four-month period	Fourth	Optional	6
Language	Spanish			
Teaching method	Face-to-face			
Prerequisites				
Department	Matemáticas			
Coordinador	Costa Bouzas, Julian	E-mail	julian.costa@udc.es	
Lecturers	Costa Bouzas, Julian Fernández Casal, Rubén	E-mail	julian.costa@udc.es ruben.fcasal@udc.es	
Web				
General description	En esta asignatura se estudian algunas de las técnicas más importantes de simulación y de remuestreo, con el objetivo de que el alumno sea capaz de identificar situaciones donde resulten de utilidad y aplicarlas correctamente. Asimismo también se provee al alumno de las bases metodológicas para llevar a cabo experimentos de simulación que permitan validar empíricamente el correcto comportamiento de un algoritmo de análisis de datos, así como la comparación entre algoritmos alternativos.			

Study programme competences	
Code	Study programme competences
A3	CE3 - Capacidade para a análise de datos e a comprensión, modelado e resolución de problemas en contextos de aleatoriedade.
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.
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		
Conocer las principales técnicas de simulación estadística	A3 A17 A20	B2 B3 B4 B7 B8 B9 B10	C1 C4
Conocer los principios del método de remuestreo bootstrap	A3 A17 A20	B2 B3 B4 B7 B8 B9 B10	C1 C4
Saber aplicar las técnicas bootstrap en problemas de regresión y con datos dependientes	A3 A17 A20	B2 B3 B4 B7 B8 B9 B10	C1 C4
Ser capaz de aplicar las principales técnicas de simulación para obtener datos simulados, así como aplicar las principales técnicas de remuestreo a conjuntos de datos reales o simulados	A3 A17 A20	B2 B3 B4 B7 B8 B9 B10	C1 C4

Contents	
Topic	Sub-topic
1. Introducción a la simulación	1.1 Conceptos básicos 1.2 Ventajas e inconvenientes de la simulación 1.3 Tipos de números aleatorios 1.4 Números aleatorios en R
2. Generación de números pseudoaleatorios	2.1 Generadores congruenciales y extensiones 2.2 Análisis de la calidad de un generador 2.3 Análisis de resultados de simulación
3. Métodos de simulación de distribuciones	3.1 Simulación de variables aleatorias continuas 3.2 Simulación de variables aleatorias discretas 3.3 Simulación de distribuciones multivariantes
4. Aplicaciones de la simulación	4.1 Estudios de simulación 4.2 Integración y optimización Monte Carlo 4.3 Aplicaciones en inferencia estadística
5. Métodos de remuestreo	5.1 Introducción al remuestreo 5.2 Bootstrap uniforme 5.3 Modificaciones del bootstrap uniforme 5.4 Herramientas disponibles en R

6. Aplicaciones del remuestreo	6.1 Estimación del sesgo y la precisión de un estimador 6.2 Intervalos de confianza bootstrap 6.3 Contrastes de hipótesis bootstrap 6.4 Inferencia en modelos de regresión 6.5 Bootstrap con datos dependientes

Planning				
Methodologies / tests	Competencies	Ordinary class hours	Student?s personal work hours	Total hours
Guest lecture / keynote speech	A17 A20 A3 B2 B3 B4 B7 B8 B9 B10 C1 C4	21	21	42
ICT practicals	A17 A20 A3 B2 B3 B4 B7 B8 B9 B10 C1 C4	21	77	98
Mixed objective/subjective test	A17 A20 A3 B2 B3 B4 B7 B8 B9 C1	4	0	4
Personalized attention		6	0	6
(*)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	Presentación de los aspectos teóricos relevantes de la materia, de modo que el alumnado pueda abordar la realización de las prácticas a través de TIC. Se fomentará en todo momento la participación y el debate
ICT practicals	Aplicación práctica de las metodologías explicadas en las sesiones magistrales. Se hará uso de herramientas de software libre, principalmente R, y también de recursos web
Mixed objective/subjective test	El alumnado deberá demostrar su dominio de los aspectos teóricos de la materia y su capacidad para la resolución de problemas

Personalized attention	
Methodologies	Description
Guest lecture / keynote speech ICT practicals	Tanto en las sesiones magistrales como en las prácticas, será importante atender personalmente a los alumnos ante cualquier duda que pueda surgir. Esta atención servirá al profesor para detectar posibles problemas en la metodología utilizada para impartir la materia y, además, a los alumnos para consolidar los conocimientos y expresar sus inquietudes sobre la materia

Assessment			
Methodologies	Competencies	Description	Qualification
ICT practicals	A17 A20 A3 B2 B3 B4 B7 B8 B9 B10 C1 C4	A lo largo del curso, el alumno demostrará su interés y dominio de la materia, y su capacidad de trabajo en equipo, mediante la realización de trabajos, individuales o en grupo. La calificación obtenida se mantendrá entre las dos oportunidades de la convocatoria de cada curso	40
Mixed objective/subjective test	A17 A20 A3 B2 B3 B4 B7 B8 B9 C1	El examen final consistirá en una prueba escrita teórico-práctica	60

Assessment comments

En la fecha establecida por la Facultad de Informática en su programación anual, el alumno realizará, por escrito, el examen final de la materia (prueba mixta), en el que deberá responder preguntas teóricas, resolver cuestiones teórico-prácticas, y calcular la solución de diversos problemas. Para esta prueba el alumno sólo podrá llevar consigo el material que se autorice de forma expresa. Para aprobar la materia será necesario superar las dos partes (prácticas a través de TIC y prueba mixta). La oportunidad de julio (segunda oportunidad) estará sometida a los mismos criterios que la oportunidad de enero (primera oportunidad). Los alumnos que se presenten a la segunda oportunidad podrán optar entre mantener la nota de prácticas a través de TIC o volver a examinarse también de esta parte. No se considera necesario realizar ninguna adaptación para los alumnos a tiempo parcial o con dispensa académica. La realización fraudulenta de las pruebas o actividades de evaluación implicará directamente la calificación de suspenso (0) en la materia en la oportunidad correspondiente.

Sources of information	
Basic	- Cao, R. (2002). Introducción a la simulación y a la teoría de colas. NetBiblo - Cao R. y Fernández-Casal R. (2021). Técnicas de Remuestreo. Libro online: https://rubenfcasal.github.io/book_remuestreo - Chernick, M.R. y LaBudde, R.A. (2014). An introduction to bootstrap methods with applications to R. John Wiley & Sons - Davison, A.C. y Hinkley, D.V. (1997). Bootstrap Methods and their Application. Cambridge University Press - Fernández-Casal, R. y Cao, R. (2022). Simulación Estadística. Libro online: https://rubenfcasal.github.io/simbook - Jones, O. et al. (2009). Introduction to Scientific Programming and Simulation Using R. CRC - Robert, C.P. y Casella, G. (2010). Introducing Monte Carlo Methods with R. Springer
Complementary	- Efron, B. y Tibshirani, R.J. (1993). An Introduction to the Bootstrap. Chapman and Hall - Fox, J. y Weisberg, S. (2018). An R companion to applied regression. Sage - Gentle, J.E. (2003). Random number generation and Monte Carlo methods. Springer?Verlag - Ripley, B.D. (1987). Stochastic Simulation. John Wiley & Sons - Ross, S.M. (1999). Simulation. Prentice Hall - Shao, J. y Tu, D. (1995). The Jackknife and Bootstrap. Springer-Verlag - Suess, E.A. y Trumbo, B.E. (2010). Introduction to probability simulation and Gibbs sampling with R. Springer

Recommendations
Subjects that it is recommended to have taken before
Regression Models/614G02012 Statistical Inference/614G02007 Probability and Basic Statistics/614G02003
Subjects that are recommended to be taken simultaneously
Subjects that continue the syllabus
Other comments

(*)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.