



Guía Docente				
Datos Identificativos			2024/25	
Asignatura (*)	Probabilidade. estatística e elementos de biomatemática	Código	614522007	
Titulación				
Descritores				
Ciclo	Período	Curso	Tipo	Créditos
Mestrado Oficial	1º cuatrimestre	Primeiro	Obrigatoria	6
Idioma	CastelánGalegoInglés			
Modalidade docente	Presencial			
Prerrequisitos				
Departamento	Matemáticas			
Coordinación	López Cheda, Ana	Correo electrónico	ana.lopez.cheda@udc.es	
Profesorado	López Cheda, Ana	Correo electrónico	ana.lopez.cheda@udc.es	
Web	<a href="https://dm.udc.es/modes/">https://dm.udc.es/modes/</a>			
Descrición xeral	<p>Preténdese que os estudantes adquiran competencias na identificación de situacións nas que a teoría de probabilidade e os métodos da inferencia estatística son ferramentas axeitadas para a análise cuantitativa de bases de datos xerados na área de bioinformática. Para iso, tratarase de que o alumnado complemente o seu coñecemento dos conceptos básicos de probabilidade e inferencia estatística, obteña soltura no manexo do software estatístico R, utilizando un gran número de recursos, e introducíndose na programación nesta contorna. Tamén se pretende que os estudantes se familiaricen cos modelos probabilísticos de procesos estocásticos en tempo discreto e adquiran unha formación básica en técnicas de remostraxe (bootstrap) como ferramenta para a posta en marcha e avaliación de diferentes algoritmos estatísticos.</p>			

Competencias / Resultados do título	
Código	Competencias / Resultados do título

Resultados da aprendizaxe			
Resultados de aprendizaxe	Competencias / Resultados do título		
G2 - Capacidade de aplicación de algoritmos de resolución dos problemas e manexo do software adecuado.	AP5 AP6 AP10	BP1	CP3
G3 - Capacidade de traballo en equipo e de xeito autónomo	AP5 AP6	BP1 BP4 BP5	CP3 CP6 CP8
G4 - Capacidade de formular problemas en termos estatísticos, e de resolvelos utilizando as técnicas axeitadas.	AP5 AP6 AP10	BP1	CP3 CP6
G11 - Adquirir destreza para o desenvolvemento de software	AP5 AP6	BP5	CP3
G14 - Representar un problema real mediante un modelizado estatístico axeitado.	AP5 AP6 AP10	BP1 BP5	
E5 - Coñecer algoritmos de resolución dos problemas e manexar o software axeitado.	AP5 AP6 AP10	BP1 BP5	CP3 CP6 CP8
E12 - Realizar inferencias respecto aos parámetros que aparecen no modelo.	AP5 AP6 AP10	BP1 BP4 BP5	CP3 CP6 CP8



E19 - Tratamento de datos e análise estatística dos resultados obtidos.	AP5 AP6 AP10	BP1 BP4 BP5	CP3
E27 - Obter os coñecementos precisos para unha análise crítica e rigurosa dos resultados.	AP5 AP10	BP1 BP4 BP5	CP6 CP8
E82 - O estudante será capaz de comprender a importancia da Inferencia Estatística como ferramenta de obtención de información sobre a poboación en estudo, a partir do conxunto de datos observados dunha mostra representativa de esta. Para iso deberá recoñecer a diferenza entre estatística paramétrica e non paramétrica.	AP5 AP10	BP1 BP4 BP5	CP6 CP8
E84 - Ser quen de manexar diverso software (en particular R) e interpretar os resultados que proporcionan nos correspondentes estudos prácticos.	AP5 AP6 AP10	BP4 BP5	CP3
E86 - Soltura no manexo da teoría da probabilidade e as variables aleatorias.	AP5 AP10	BP1 BP4 BP5	CP6
Coñecemento dos conceptos básicos de probabilidade e inferencia estatística e a súa aplicación na bioinformática	AP1 AP5		
Coñecemento e aplicación de técnicas estatísticas para a análise cuantitativa de bases de datos xeradas no ámbito da Bioinformática.	AP1 AP2 AP3 AP5 AP6 AP7	BP2 BP4 BP5 BP6	CP3
Obter soltura co software estatístico R, manexando un importante número de recursos e introducindo ó estudiantado na programación neste contorna.	AP3 AP6		CP3
Familiarizarse con modelos probabilísticos de procesos estocásticos en tempo discreto.	AP1	BP2 BP4	
Formación en técnicas de remostraxe (bootstrap) como ferramenta para a aplicación e/ou avaliación de diferentes algoritmos estatísticos.	AP1 AP2 AP3 AP5 AP7	BP2	CP3

Contidos	
Temas	Subtemas
1. Revisión de conceptos básicos de probabilidade e estatística.	a. Probabilidade. Variables aleatorias e distribucións notables discretas e continuas. Distribucións multivariantes. b. Inferencia estatística: estimación, contrastes de hipóteses e intervalos de confianza.
2. Revisión da linguaxe de programación estatística R.	a. Introducción ao R. Primeiros pasos. Funcións internas. Axuda en R. Funcións, bucles, vectores. Funcións estatísticas. Gráficas. Recursividade. R studio. b. Principais distribucións de probabilidade en R. c. Introducción á simulación en R. d. Estatística descritiva en R. e. Contrastes de hipóteses e intervalos de confianza con R.



3. Modelos estatísticos lineais.	<p>a. O modelo de regresión linear simple. Hipóteses básicas. Estimación. Contrastes. Predición. Diagnose do modelo.</p> <p>b. O modelo de regresión linear múltiple. Hipóteses básicas. Estimación. Contrastes. Predición. Diagnose do modelo.</p> <p>c. Modelos básicos do deseño experimental. Análise da Varianza (ANOVA) dunha e dúas vías, sen e con interacción. Hipóteses básicas. Estimación. Contrastes. Diagnose do modelo.</p> <p>d. O problema dos contrastes múltiples. False discovery rate.</p>
4. Introducción aos procesos estocásticos.	<p>a. Paseo aleatorio simple.</p> <p>b. Proceso de Poisson e procesos de renovación. Procesos de nacemento e morte.</p> <p>c. Procesos Markovianos. Cadeas de Markov.</p>
5. Introducción aos métodos de remostraxe.	<p>a. O Bootstrap uniforme. Cálculo da distribución Bootstrap: distribución exacta e distribución aproximada por Monte Carlo. Exemplos. Aplicación do Bootstrap á estimación da precisión e o nesgo dun estimador.</p> <p>b. Modificacións do Bootstrap uniforme. Bootstrap paramétrico, simetrizado e suavizado. Discusión e exemplos.</p> <p>c. Métodos bootstrap para a construción de intervalos de confianza: método percentil, percentil-t, percentil-t simetrizado. Exemplos.</p> <p>Estudos de simulación.</p>
6. Revisión de métodos numéricos de optimización.	<p>a. Conceptos básicos de optimización.</p> <p>b. Optimización de funcións sen restricións.</p>

Planificación				
Metodoloxías / probas	Competencias / Resultados	Horas lectivas (presenciais e virtuais)	Horas traballo autónomo	Horas totais
Presentación oral	A5 A6 A10 B1 B4 B5 C8	24	36	60
Prácticas a través de TIC	A5 A6 A10 B4 B5 C3 C6	18	36	54
Proba de resposta múltiple	A5 B1 B5 C8	1	9	10
Solución de problemas	A5 A6 A10 B1 B4 B5 C3 C6 C8	4	16	20
Atención personalizada		6	0	6

\*Os datos que aparecen na táboa de planificación son de carácter orientativo, considerando a heteroxeneidade do alumnado

Metodoloxías	
Metodoloxías	Descrición
Presentación oral	Presentación con ordenador
Prácticas a través de TIC	Análise estatística de conxuntos de datos usando R.
Proba de resposta múltiple	Proba de resposta múltiple sobre conceptos.
Solución de problemas	Elección das ferramentas estatísticas e estratexias para resolver problemas. Formulación de modelos lineais. Deseño de experimentos. Formulación de plans de remostraxe.

Atención personalizada	
Metodoloxías	Descrición



Prácticas a través de TIC	Asistencia e participación nas clases teóricas. Exame escrito de múltiple opción.
Solución de problemas	Participación en prácticas e seminarios. Suposto práctico a realizar polo alumno.

Avaliación			
Metodoloxías	Competencias / Resultados	Descrición	Cualificación
Presentación oral	A5 A6 A10 B1 B4 B5 C8	Presentación oral do traballo orixinal mencionado no item &quot;Solución de problemas&quot; e seguimento continuado e obxectivo dunha participación activa.	10
Prácticas a través de TIC	A5 A6 A10 B4 B5 C3 C6	Práctica de ordenador usando o software estatístico libre R (exame final).	30
Solución de problemas	A5 A6 A10 B1 B4 B5 C3 C6 C8	Traballo orixinal sobre algún dos temas da materia nun contexto de interese en Bioinformática e seguimento continuado e obxectivo dunha participación activa.	30
Proba de resposta múltiple	A5 B1 B5 C8	Proba de comprensión dos conceptos impartidos (exame final).	30

Observacións avaliación
<p>A avaliación realizarase por medio dunha proba sobre prácticas con R, un traballo individual do/da alumno/a, así como unha proba escrita de conceptos. A calificación da proba de conceptos representará o 30% da calificación global, a proba de práctica en R corresponderá ao 30% mentres que o 40% restante corresponderá ao traballo individual, que ten que ser presentado en público polos estudantes. A cuarta parte da puntuación deste traballo individual (10% da calificación global) corresponde á presentación oral do mesmo.</p> <p>Para superar a materia será necesario obter unha calificación de alomenos 5 sobre 10 no conxunto da materia.</p> <p>Na oportunidade de xullo os estudantes poderán liberarse de facer as probas correspondentes nas que a súa calificación na oportunidade de xaneiro fora de alomenos 4 sobre 10.</p> <p>Os estudantes a tempo parcial serán avaliados polo sistema de avaliación continua. Polo tanto, poderán realizar e presentar o traballo individual durante o cuadrimestre.</p> <p>Na primeira oportunidade (xaneiro-febreiro), se os estudantes que non se teñan presentado a ningunha das probas avaliadas que figuran arriba e que non entregaran o traballo individual obterán a calificación de NON PRESENTADO. En xullo obterán a calificación de NON PRESENTADO os estudantes que non tiveran presentado ao exame final desa data.</p> <p>Se algún estudante quere facer algunha das probas nun idioma oficial específico (galego ou español), debe avisar ó profesorado alomenos 1 semana antes da correspondente proba.</p> <p>A realización fraudulenta das probas ou actividades de avaliación, unha vez comprobada, implicará directamente a calificación de suspenso na convocatoria en que se cometa: o/a estudante será cualificado con ?suspenso? (nota numérica 0) na convocatoria correspondente do curso académico, tanto se a comisión da falta se produce na primeira oportunidade como na segunda. Para isto, procederase a modificar a súa calificación na acta de primeira oportunidade, se fose necesario.</p>

Fontes de información
-----------------------



<b>Bibliografía básica</b>	<ul style="list-style-type: none"><li>- Cao Abad, R., Francisco Fernández, M., Naya Fernández, S., Presedo Quindimil, M.A., Vázquez Brage, M (2001). Introducción a la Estadística y sus Aplicaciones. Pirámide</li><li>- Ewens, W.J. and Grant, G.R. (2005). Statistical Methods in Bioinformatics. Springer</li><li>- Peña Sánchez de Rivera, D. (2000). Estadística: Modelos y Métodos. Alianza Editorial</li><li>- Ross, S.M. (1995). Stochastic Processes. Wiley</li><li>- Efron, B. and Tibshirani, R.J. (1993). An Introduction to the Bootstrap. Chapman and Hall</li><li>- Davison, A.C. and Hinkley, D.V. (1997). Bootstrap Methods and their Application. Cambridge University Press</li></ul>
<b>Bibliografía complementaria</b>	

## Recomendacións

### Materias que se recomenda ter cursado previamente

### Materias que se recomenda cursar simultaneamente

Introdución ás bases de datos/614522002

Xenómica/614522006

Fundamentos de bioinformática/614522008

Introdución á programación/614522001

Fundamentos de intelixencia artificial/614522003

### Materias que continúan o temario

Estruturas de datos e algoritmia para secuencias biolóxicas/614522013

Procesamento avanzado de secuencias biolóxicas/614522020

Intelixencia computacional para datos de alta dimensionalidade/614522024

Traballo fin de mestrado/614522025

Intelixencia computacional para bioinformática/614522012

Métodos estatísticos avanzados en bioinformática/614522009

## Observacións

Segundo se recolle nas distintas normativas de aplicación para a docencia universitaria, deberase incorporar a perspectiva de xénero nesta materia (usarase linguaxe non sexista, utilizarase bibliografía de autores/as de ambos sexos, propiciarse a intervención en clase de alumnos e alumnas, etc.) Traballárase para identificar e modificar prexuízos e actitudes sexistas e influirase na contorna para modificalos e fomentar valores de respecto e igualdade. Deberanse detectar situacións de discriminación por razón de xénero e proporanse accións e medidas para corrixilas.

(\*A Guía docente é o documento onde se visualiza a proposta académica da UDC. Este documento é público e non se pode modificar, salvo casos excepcionais baixo a revisión do órgano competente dacordo coa normativa vixente que establece o proceso de elaboración de guías