COURSE DATA

Data Subject	
Code	42227
Name	Mathematical and computer tools
Cycle	Master's degree
ECTS Credits	18.0
Academic year	2023 - 2024

Study (s)

Degree	Center	Acad. year	Period
2002 - M.D. in Biostatistics	Faculty of Mathematics	1	First term

Subject-matter

Degree	Subject-matter	Character
2002 - M.D. in Biostatistics	1 - Mathematical and computer tools	Obligatory

Coordination

Name	Department
FORTE DELTELL, ANABEL	130 - Statistics and Operational Research

SUMMARY

El módulo Herramientas Matemáticas e Informáticas tiene como objetivo proporcionar al estudiante la formación específica en matemáticas, informática y probabilidad necesaria para el desempeño profesional bioestadístico.

El módulo es autocontenido y no se requieren conocimientos previos.

PREVIOUS KNOWLEDGE

Relationship to other subjects of the same degree

There are no specified enrollment restrictions with other subjects of the curriculum.

Other requirements

OUTCOMES

2002 - M.D. in Biostatistics

- Poseer habilidades en el uso de las TICs para el desempeño profesional.
- Saber programar en lenguajes de uso habitual para el análisis estadístico de datos.
- Ser capaces de diseñar e implementar programas informáticos para la resolución computacional de cualquier tipo de problema matemático y/o estadístico que pueda surgir en la práctica bioestadística.
- Ser capaces de programar, gestionar y consultar bases de datos con fines estadísticos.
- Ser capaces de diseñar y generar informes técnicos y realizar consultas de interés.

LEARNING OUTCOMES

Al finalizar el proceso de enseñanza-aprendizaje el estudiante habrá aprendido a:

1: Recopilar y procesar conjuntos de datos.

2: Conocer y emplear los conceptos y tecnologías de bases de datos en sus tres aspectos fundamentales: diseño, utilización y funcionamiento.

3: Emplear el lenguaje estándar de consulta e interacción con bases de datos (SQL).

4: Entender el funcionamiento interno de los sistemas de gestión de bases de datos.

5: Obtener un análisis estadístico descriptivo completo a partir de una fuente de datos.

6: Caracterizar y modelizar estadísticamente una fuente de datos con el lenguaje R.

7: Conocer y emplear los conceptos matemáticos de función, derivación, integración y ecuación diferencial.

8: Manejar con R los elementos y operaciones comunes del álgebra matricial.

9: Conocer y emplear los conceptos básicos de Cálculo de Probabilidades.

10: Realizar simulación de variables y vectores aleatorios.

DESCRIPTION OF CONTENTS

1. Computación y Programación en R

Tema 1. Introducción y preliminares. Motivos para su uso. Distribución e instalación de R. La ayuda en R. Salvar e iniciar sesiones anteriores. Permanencia de los datos y eliminación de objetos. Edición y ejecución de comandos desde otros programas: scripts. Librerías de R. Introducción a Rcommander.
Tema 2. Manejo de datos. Características de los objetos en R: modos y atributos. Coerción de tipos. Vectores. Generación de secuencias regulares. Variables indexadas (arrays). Matrices. Operaciones

con matrices. Listas. Hojas de datos: funciones attach y dettach. Lectura de ficheros de datos. Importar y exportar datos de otros programas. La librería foreign.
Tema 3. Descripción numérica y gráfica de datos. Tablas de frecuencias. Medidas de localización, dispersión y forma. Descripción gráfica de datos en R. Gráficos para datos discretos, continuos y/o multivariantes. Estimación de densidades. Parámetros gráficos. La función par(). Exportando gráficos. Dispositivos gráficos.
Tema 4. Análisis de datos con R. Inferencia en problemas univariantes. Inferencia en problemas de dos muestras. Análisis de datos categóricos. Análisis de la Varianza. Regresión lineal simple.
Tema 5. Programación de funciones y subrutinas. Órdenes para la ejecución condicional y para la ejecución repetitiva en bucles y ciclos. Funciones: sintaxis y llamada. Nombres de argumentos y valores por defecto. El argumento . Funciones de control y parada. Introducción a las clases y a la creación de librerías.

2. Matemáticas para Bioestadística

Tema 1. Funciones. Funciones elementales. Función exponencial, logarítmica, trigonométrica. Funciones Especiales. Funciones de densidad de probabilidad y funciones de distribución de algunas variables aleatorias. Gráficas de Funciones en R.
Tema 2. Derivación. Comprensión. La función derivada. Cálculo de derivadas en R.
Tema 3. Comportamiento de Funciones. Crecimiento y decrecimiento. Máximos y mínimos. Puntos de Inflexión. Comportamiento en el infinito.
Tema 4: Ecuaciones diferenciales para la dinámica de poblaciones. El modelo de Maltus. La ecuación logística.
Tema 5. Integración. La función primitiva. Cálculo de áreas en R. Relación con la teoría de la probabilidad.
Tema 6. Álgebra Matricial. Notación matricial de un sistema lineal. Valores y vectores propios. Ortogonalidad. Resolución de sistemas lineales. Descomposiciones matriciales: La descomposición LU.

3. Gestión de Bases de Datos

Tema 1: Introducción a las bases de datos relacionales y al álgebra relacional.
Tema 2: El lenguaje SQL: Consulta de datos.
Tema 3: El lenguaje SQL: Creación, modificación y borrado de tablas.
Tema 4: El lenguaje SQL: Inserción, borrado y modificación de datos.
Tema 5: Diseño de bases de datos.
Tema 6: Adquisición de datos desde Internet.
Tema 7: Acceso a bases de datos desde R.

4. Probabilidad y Simulación

Tema 1. Probabilidad: Una medida de la incertidumbre. Modelos de probabilidad. Propiedades de los modelos de probabilidad. Probabilidad condicional e independencia. Dígitos aleatorios y números pseudo-aleatorios. Simulación de sucesos. Selección de muestras aleatorias en poblaciones finitas.
Tema 2. Variables aleatorias y distribuciones. Variables aleatorias. Distribuciones discretas y continuas. Distribuciones conjuntas. Condicionamiento e independencia. Simulación de distribuciones de probabilidad: métodos basados en transformaciones. Métodos de composición. Métodos de Aceptación-rechazo.
Tema 3. Valores esperados. Esperanza. Varianza, covarianza y correlación. Valor esperado condicional. Leyes de los grandes números. Integración Monte Carlo. Métodos de reducción de la varianza.
Tema 4. Distribuciones muestrales. Aproximaciones asintóticas: Teorema Central del Límite. Aproximaciones por Monte Carlo. Teoría de la distribución Normal.
Tema 5. Cadenas de Markov. Cadenas de Markov. Distribuciones estacionarias y Teoremas Límite. Métodos MCMC: Metropolis-Hastings y Gibbs Sampling.

WORKLOAD

ACTIVITY	Hours	% To be attended
Tutorials	108,00	100
Computer classroom practice	72,00	100
Development of group work	40,00	0
Development of individual work	100,00	0
Study and independent work	40,00	0
Preparation of evaluation activities	20,00	0
Preparation of practical classes and problem	60,00	0
Resolution of case studies	100,00	0
TOTAL	**540,00**	

TEACHING METHODOLOGY

La metodología docente consistirá en clases presenciales teóricas y prácticas y en trabajos a desarrollar por el estudiante. Las clases presenciales se dividirán en:

- Clases teóricas, en las que se expondrán los conceptos básicos de cada uno de los puntos del temario.
- Clases prácticas, en las que se desarrollarán ejercicios prácticos de lo expuesto en las clases de teoría con el fin de reforzar su comprensión. Estas clases servirán, además, para generar nuevos puntos de vista y enfoques no analizados en las clases teóricas, así como comprobar el grado de adquisición de los conocimientos teóricos por parte de los alumnos.

Por su parte, el estudiante deberá desarrollar diferentes trabajos y actividades con la ayuda mediante tutorías del profesor, que servirán para comprobar el grado de adquisición de las competencias. Éstos deberán ser eminentemente prácticos, aunque podrán versar sobre aspectos teóricos vistos en el curso. Trabajos posibles son, por ejemplo, un análisis, diseño e implementación de una base de datos ó un estudio de análisis de datos o un estudio de simulación.

EVALUATION

Dependiendo de la asignatura, la evaluación se realizará a partir de un examen para demostrar la consecución de los objetivos generales de la asignatura, y/o de la valoración de las prácticas y trabajos realizados como prueba de la adquisición de las destrezas indicadas.

La evaluación del módulo provendrá de promediar las calificaciones obtenidas en la evaluación de las asignaturas del mismo, ponderadas por sus créditos ECTS. Será imprescindible haber conseguido en todas y cada una de las asignaturas del módulo una calificación superior a 3,5 puntos (sobre 10).

El sistema de evaluación para cada una de las asignaturas que conforman el módulo estará basado en tres posibles tipos de actividades evaluables no excluyentes:

- Prácticas/ejercicios/tests
- Proyecto(s) de trabajo(s), generalmente basado(s) en el estudio de un(os) caso(s).
- Examen final

Las prácticas/tests se presentarán al estudiante en las diferentes sesiones de la asignatura, así como los proyectos de trabajo (si procede), que habrán de entregarse resueltos en forma de informe técnico. La exposición de trabajos por parte de los estudiantes y las tutorías programadas permitirán valorar los resultados de aprendizaje obtenidos por los estudiantes. Esta valoración se completará con los informes entregados y los exámenes realizados.

La calificación máxima en cada asignatura será de 10 y la mínima de 0.

El módulo podrá ser reconocido a aquellos estudiantes que hayan superado un conjunto adecuado de asignaturas de matemáticas, bases de datos, análisis de datos y probabilidad de nivel universitario previo análisis del programa o programas cursados o de la actividad demostrable.

REFERENCES

Basic

- Michael J. Crawley (2007). The R book. John Wiley and Sons, Ltd.

- Milton, J.S. (1994) Estadística para Biología y Ciencias de la Salud (Segunda Edición). Mc-Graw-Hill.
- Rius, F. y Barón, F.J. (2005) Bioestadística. Thomson.
- Forrest Houlette (2003). Fundamentos de SQL. McGraw-Hill Interamericana.
- James R. Groff, Paul N. Weinberg (1990). Aplique SQL. McGraw-Hill.
- Roger Jennings (2002) Microsoft Access 2002: edición especial Pearson Educación.

Additional

- Evans, M.J. y Rosenthal, J.S. (2005) Probabilidad y Estadística. Reverté.
- Robert, C.P. y Casella, G. (2005). Monte Carlo Statistical Methods. Springer.
- Shorack, G.R. (2000). Probability for Statisticians. Springer.
- Miguel Ángel Martín Tardío (2001). Manual avanzado de Microsoft Access 2002 Office XP. Anaya Multimedia.
- M. Samuels y J. Witmer (2002). Statistics for the Life Sciences. Prentice Hall.