Universitat Autònoma de Barcelona

2023/2024

Estadística y Análisis de Datos

Código: 44079
Créditos ECTS: 9

Titulación	Tipo	Curso	Semestre
4313861 Física de Altas Energias, Astrofísica y Cosmología	OB	0	1

Contacto

Nombre: Francisco Javier Rico Castro

Correo electrónico: franciscojavier.rico@uab.cat

Idiomas de los grupos

Puede consutarlo a través de este enlace. Para consultar el idioma necesitará introducir el CÓDIGO de la asignatura. Tenga en cuenta que la información es provisional hasta el 30 de noviembre del 2023.

Equipo docente

Abelardo Moralejo Olaizola

Carles Sánchez Alonso

Jorge Carretero Palacios

Pau Tallada Crespí

Martin Borstad Eriksen

Francesc d'Assis Torradeflot Curero

Prerrequisitos

Para el "Bootcamp" de Python (parte 2), es necesario el uso de un ordenador portátil personal con una instalación de Python 3.

Se recomienda instalar Python 3 con el intalador de Anaconda. De esta manera, su distribución de Python contendrá todos los paquetes asociados necesarios para este curso.

Siga estos pasos:

1. Descargue el instalador de Anaconda para Python 3 aquí https://www.anaconda.com/download/

2. Siga las instrucciones de instalación; tanto la GUI como las versiones de terminal funcionan bien. Si se le solicita, seleccione la opción para agregar el nuevo directorio de anaconda a su ruta.

El uso de GNU/Linux es altamente recomendado.

Objetivos y contextualización

En este curso aprenderemos cómo extraer conocimiento científico de datos experimentales, un proceso que se basa en métodos estadísticos. Aprenderemos los conceptos básicos de Probabilidad y Estadística (en sus marcos Frequentista y Bayesiano). Además, estudiaremos y practicaremos varios métodos estadísticos particulares y técnicas de análisis de datos que normalmente se utilizan en los campos de Física de Alta Energía, Astrofísica y Cosmología. Con ese fin, aprenderemos y practicaremos el uso de herramientas modernas de software de análisis y estadísticas.

Competencias

- Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio.
- Trabajar en grupo, asumir responsabilidades compartidas e interaccionar profesionalmente y de manera constructiva con otras personas con un respeto absoluto a sus derechos.
- Usar las matemáticas para describir el mundo físico, seleccionar las ecuaciones apropiadas, construir modelos adecuados, interpretar resultados matemáticos y comparar críticamente con experimentación y observación.
- Usar software adecuado, lenguajes de programación y paquetes informáticos en la investigación de problemas relacionados con la física de altas energías, la astrofísica o la cosmología.

Resultados de aprendizaje

1. Aplicar las técnicas de análisis de datos a problemas tanto del ámbito de la física de partículas, la astrofísica y la cosmología, como de ámbitos cercanos pero distintos.
2. Aprender el funcionamiento de programas informáticos de análisis estadístico.
3. Trabajar en pequeños grupos para resolver problemas de análisis de datos
4. Utilizar las técnicas Monte Carlo para modelar problemas reales de Física.

Contenido

Parte 1: Conceptos básicos sobre probabilidad, estadística y técnicas de Monte Carlo.

Parte 2: Python para estadísticas y análisis de datos

Parte 3: Estimación de parámetros, test de hipótesis y "Unfolding".

Parte 4: Estadística Bayesiana

Metodología

- Conferencias teóricas que incluyen ejemplos prácticos en los campos de Física de Alta Energía, Astrofísica y Cosmología.
- Ejercicios y tareas para ser resueltos por las alumnas y alumnos individualmente o en pequeños grupos.
- Discusión de problemas durante las clases y tutorías.
- Sesiones prácticas sobre herramientas de software para estadísticas y análisis de datos (en lenguaje de programación Python)
- Explicación y discusión de código / algoritmos de muestra en lenguajes de programación Python durante clases y tutoriales

Nota: se reservarán 15 minutos de una clase dentro del calendario establecido por el centro o por la titulación para que el alumnado rellene las encuestas de evaluación de la actuación del profesorado y de evaluación de la asignatura o módulo.

Actividades

Título	Horas	ECTS	Resultados de aprendizaje
Tipo: Dirigidas			
Clases	56	2,24	1, 2, 4
Estudio de la teoría y ejemplos prácticos	64	2,56	1, 2, 3, 4
Tipo: Autónomas			
Discusión, grupos de trabajo, resolución de problemas.	60	2,4	1, 2, 3, 4

Evaluación

La evaluación tendrá en cuenta:

- Asistencia y participación activa a las clases.
- Resolución de ejercicios específicos a lo largo del curso.
- Resolución de un examen final.

Para aquellas y aquellos estudiantes que no aprueben el curso después del procedimiento de evaluación regular, habrá una ronda de evaluación de recuperación consistente en ejercicios específicos para las diferentes partes del curso y de un examen final de síntesis. No habrá una nota de umbral para ser elegible para la ronda de evaluación de recuperación, a parte del requisito general de haber sido evaluada/evaluado al menos de un 66% del total de actividades de calificación en la primera ronda.

Esta asignatura/módulo no prevee el sistema de evaluación única.

Actividades de evaluación continuada

Título	Peso	Horas	ECTS	Resultados de aprendizaje
Asistencia y participación activa en las clases.	5%	0	0	1, 2, 4
Resolución de ejercicios de clase	45%	40	1,6	1, 2, 3, 4
Resolución de un examen final de síntesis	50%	5	0,2	1, 2, 4

Bibliografía

- G. Bohm and G. Zech; "Introduction to Statistics and Data Analysis for Physicists", 3rd Edition, 2017, Verlag Deutsches Elektronen-Synchrotron (available on-line https://s3.cern.ch/inspire-prod-files-d/da9d786a06bf64d703e5c6665929ca01)
- F. James; "Statistical Methods in Experimental Physics", 2nd Edition, 2006, World Scientific
- G. Cowan; "Statistical Data Analysis", 1998, Oxford University Press

- A. Gelman, J. B. Carlin, H. S. Stern, et al. "Bayesian Data Analysis", 3rd Edition, 2013, CRC Press

Software

Introduciremos y haremos uso del lenguaje de programación Python (consulte la sección "Prerequisitos" para obtener información detallada sobre la instalación)

En particular, estudiaremos y utilizaremos las siguientes librerías Python: NumPy, pandas, matplotlib, scipy y scikit learn