Universidad de Navarra

Análisis de Datos (Ing.Gr.)

Guía docente 2025-26

PRESENTACIÓN

Breve descripción:

En esta asignatura se explican las técnicas matemáticas para analizar todo tipo de datos utilizados en ingeniería. Se hará especial hincapié en los métodos estadísticos.

Titulación (Módulo/Materia):

- Ingeniería en Tecnologías Industriales (Bloque Especializado de Tecnologías Industriales/Matemáticas y Ciencias de Computación)
- Ingeniería Mecánica (Bloque Especializado Mecánica/Matemáticas)
- Ingeniería Eléctrica (Bloque Especializado de electricidad/Matemáticas y Ciencias de la Computación)
- Ingeniería en Electrónica Industrial (Bloque Especializado de Electrónica Industrial/Matemáticas y Ciencias de la Computación)
  - Ingeniería en Sistemas de telecomunicación (Bloque Común a la Rama de Telecomunicación/Fundamentos de Sistemas de Transmisión)
  - Ingeniería en Organización Industrial (Métodos Cuantitativos/ Estadística Industrial)
  - Ingeniería en Diseño industrial y Desarrollo de productos (Conocimientos Científicos Aplicados/Conocimientos Científicos Aplicados)
  - Ingeniería Biomédica (Fundamentos de Biología/ Bioestadística)
- Ingeniería en Inteligencia Artificial (Inteligencia Artificial / Ciencia de Datos)

Detalles:

- **ECTS**: 4 ECTS
- **Curso, semestre**: 2.º curso, 1.º semestre
- **Carácter**: Obligatorio
- **Idioma**: Castellano

Profesores de la asignatura:

- Rubio Díaz-Cordovés, Ángel / Profesor Catedrático
- Ferrer-Bonsoms Hernández, Juan / Ayudante Doctor

RESULTADOS DE APRENDIZAJE (Competencias)

INGENIERÍA EN TECNOLOGÍAS INDUSTRIALES

CG6 - Proporcionar las bases científico-tecnológicas necesarias para el aprendizaje autónomo, o para cursar estudios de postgrado que le permitan profundizar y/o especializarse en diferentes campos de la ingeniería.

CG7 - Capacitar profesionales para gestionar, evaluar y mejorar sistemas de información basados en tecnologías de la información y las telecomunicaciones.

CG11 - Habilitar al egresado de destrezas técnicas y de una sensibilización que le permita impulsar, organizar y llevar a cabo mejoras e innovaciones tanto en procesos, bienes y servicios.

CE3 - Conocimientos básicos sobre el uso y programación de los ordenadores, sistemas operativos, bases de datos y programas informáticos con aplicación en ingeniería.

CE1 - Capacidad para la resolución de los problemas matemáticos que puedan plantearse en la ingeniería. Aptitud para aplicar los conocimientos sobre: álgebra lineal; geometría; geometría diferencial; cálculo diferencial e integral; ecuaciones diferenciales y en derivadas parciales; métodos numéricos; algorítmica numérica; estadística y optimización.

INGENIERÍA MECÁNICA

CG3 - Conocimiento en materias básicas y tecnológicas, que les capacite para el aprendizaje de nuevos métodos y teorías, y les dote de versatilidad para adaptarse a nuevas situaciones.

CE1 - Capacidad para la resolución de los problemas matemáticos que puedan plantearse en la ingeniería. Aptitud para aplicar los conocimientos sobre: álgebra lineal; geometría; geometría diferencial; cálculo diferencial e integral; ecuaciones diferenciales y en derivadas parciales; métodos numéricos; algorítmica numérica; estadística y optimización.

CE3 - Conocimientos básicos sobre el uso y programación de los ordenadores, sistemas operativos, bases de datos y programas informáticos con aplicación en ingeniería.

INGENIERÍA ELÉCTRICA

CG3 - Conocimiento en materias básicas y tecnológicas, que les capacite para el aprendizaje de nuevos métodos y teorías, y les dote de versatilidad para adaptarse a nuevas situaciones.

CE3 - Conocimientos básicos sobre el uso y programación de los ordenadores, sistemas operativos, bases de datos y programas informáticos con aplicación en ingeniería.

CE1- Capacidad para la resolución de los problemas matemáticos que puedan plantearse en la ingeniería. Aptitud para aplicar los conocimientos sobre: álgebra lineal; geometría; geometría diferencial; cálculo diferencial e integral; ecuaciones diferenciales y en derivadas parciales; métodos numéricos; algorítmica numérica; estadística y optimización.

INGENIERÍA EN ELECTRÓNICA INDUSTRIAL

CG3 - Conocimiento en materias básicas y tecnológicas, que les capacite para el aprendizaje de nuevos métodos y teorías, y les dote de versatilidad para adaptarse a nuevas situaciones.

CE25 - Conocimiento y capacidad para el modelado y simulación de sistemas.

CE1 - Capacidad para la resolución de los problemas matemáticos que puedan plantearse en la ingeniería. Aptitud para aplicar los conocimientos sobre: álgebra lineal; geometría; geometría diferencial; cálculo diferencial e integral; ecuaciones diferenciales y en derivadas parciales; métodos numéricos; algorítmica numérica; estadística y optimización.

CE3 - Conocimientos básicos sobre el uso y programación de los ordenadores, sistemas operativos, bases de datos y programas informáticos con aplicación en ingeniería.

INGENIERÍA EN SISTEMAS DE TELECOMUNICACIÓN

R23 - Conocimiento en materias básicas y tecnológicas, que les capacite para el aprendizaje de nuevos métodos y teorías, y les dote de versatilidad para adaptarse a nuevas situaciones.

R2 - Capacidad para la resolución de los problemas matemáticos que puedan plantearse en la ingeniería. Aptitud para aplicar los conocimientos sobre: álgebra lineal; geometría; geometría diferencial; cálculo diferencial e integral; ecuaciones diferenciales y en derivadas parciales; métodos numéricos; algorítmica numérica; estadística y optimización.

R3 - Conocimientos básicos sobre el uso y programación de los ordenadores, sistemas operativos, bases de datos y programas informáticos con aplicación en ingeniería.

R7 - Capacidad para aprender de manera autónoma nuevos conocimientos y técnicas adecuados para la concepción, el desarrollo o la explotación de sistemas y servicios de telecomunicación.

R8 - Capacidad de utilizar aplicaciones de comunicación e informáticas (ofimáticas, bases de datos, cálculo avanzado, gestión de proyectos, visualización, etc.) para apoyar el desarrollo y explotación de redes, servicios y aplicaciones de telecomunicación y electrónica.

INGENIERÍA EN ORGANIZACIÓN INDUSTRIAL

CG7 - Promover las capacidades y competencias dirigidas hacia la resolución de problemas, la iniciativa, la toma de decisiones, la creatividad, el análisis y el razonamiento crítico.

CE21 - Conocimientos y manejo adecuados de técnicas estadísticas que permitan su integración en la toma de decisiones estratégicas de una manera organizada.

INGENIERÍA EN DISEÑO INDUSTRIAL Y DESARROLLO DE PRODUCTOS

CE1 - Capacidad para la resolución de los problemas matemáticos que puedan plantearse en la ingeniería. Aptitud para aplicar los conocimientos sobre: álgebra lineal; geometría; geometría diferencial; cálculo diferencial e integral; métodos numéricos; algorítmica numérica; estadística y optimización.

CE3 - Conocimientos básicos sobre el uso y programación de los ordenadores, sistemas operativos, bases de datos y programas informáticos con aplicación en ingeniería.

INGENIERÍA EN INGENIERÍA BIOMÉDICA

CG1 - La formación debe proporcionar al egresado una base científica sólida que permita abordar con rigor los retos profesionales del sector biomédico.

CE3 - Conocer las ventajas y desventajas de los diferentes métodos estadísticos y las diversas hipótesis estudiadas.

INGENIERÍA EN INTELIGENCIA ARTIFICIAL

R9 - Aplicar las técnicas de análisis de datos que ayudan a la toma de decisiones.

R19 - Extraer información a partir de datos heterogéneos en distintas aplicaciones de IA.

PROGRAMA

1. Introducción y repaso:

Universidad de Navarra

1. Funciones de densidad y distribución habituales. Continuas y discretas.
2. Función de densidad Chi-2, t-Student, F-Fisher
3. Muestra de una distribución.

1b. Lenguaje de programación R

1. Introducción

2. Tipos de datos

3. Lectura de datos de ficheros

4. Aplicación a estadística: tipos de distribuciones.

5. Control del flujo de programas. Creación de funciones.

2. Inferencia Estadística: test de diferencia de medias, análisis de la varianza tipo I

1. Intervalos de confianza para la media y la varianza,...

2. Inferencia estadística: hipótesis nula y p-value.

3. T-test a una cola y a dos colas. T-test pareados.

4. ANOVA type I.

3. Regresión lineal simple

1. Resolución por mínimos cuadrados.
2. Distribuciones estadísticas de los parámetros estimados de la regresión.
3. Inferencia en regresión lineal. Distribución estadística de las estimaciones de la regresión.
4. Comparación con el test de diferencia de medias y el análisis de la varianza.

4. Regresión lineal múltiple (Modelos lineales)

1. Matriz de diseño y matriz de contrastes.
2. Influencia de cada factor.
3. Varianza explicada y significancia de la regresión.

Universidad de Navarra

English version:

1. Introduction and review:

1. Density functions and usual distributions. Continuous and discrete.
2. Chi-2, t-Student, and F-Fisher distributions
3. Sample of a distribution.

1b. Programming in R

1. Introduction
2. Type of data
3. Read data from files
4. Application to statistics: type of distributions.
5. Program flow control. Creation of functions.

2. Statistical inference: test of difference of means, analysis of variance type I.

1. Confidence intervals for mean and variance
2. Statistical inference: null hypothesis and p-value.
3. One-tailed and two-tailed t-tests. Paired t-tests.
4. ANOVA type I.

3. Simple linear regression

1. Least squares resolution.
2. Statistical distributions of the estimated regression parameters.
3. Inference in linear regression. Statistical distribution of regression estimates.
4. Comparison with the test of difference of means and analysis of variance.

4. Multiple linear regression (Linear models)

1. Design matrix and matrix of contrasts.
2. Influence of each factor.
3. Explained variance and significance of the regression.

Universidad de Navarra

ACTIVIDADES FORMATIVAS

Durante la asignatura habrá diferentes actividades docentes:

1. Clases magistrales en el aula.
2. Resolución de problemas en el aula.
3. Clases prácticas en aula de ordenadores.
4. Evaluación continua mediante test teóricos y prácticos.

Estudio personal:

1. El alumno debe estudiar las bases teóricas de la asignatura.
2. El alumno debe aprender a programar con fluidez en R.
3. El alumno debe, individualmente, resolver los problemas propuestos
4. Trabajo optativo por grupos: cada grupo deberá analizar un data set diferente y aplicar los métodos estadísticos explicados en la asignatura.

- En el área interna, en la sección de "Contenidos" está disponible la programación de todas clases de pizarra y de ordenadores. En este calendario también está señalado la fecha de los tests correspondientes a la evaluación continua.

EVALUACIÓN

CONVOCATORIA ORDINARIA

Para calcular la nota final de la asignatura se tendrá en cuenta:

- **Examen parcial sobre R y test de hipótesis (30%)**
- **El examen final (50%).**
- **Tests teóricos (10%). (Media de todos menos el peor).**
- **Tests prácticos (10%). (Media de todos menos el peor).**

Trabajo:

- **Opcional**. Si no se presenta, las cuentas son las anteriores.
- Si se presenta, **el trabajo vale un 20%** y **las cuentas anteriores 85%**. Es decir, podría sacarse más de 10 si se hace trabajo aunque la evaluación final será como mucho de 10.
- No se otorgará la matrícula de honor si no se ha hecho trabajo.
- Los trabajos se harán por parejas (sin excepciones).
- Se indicará una fecha límite para indicar la pareja de alumnos y el título (aproximado del trabajo).
- El trabajo consisitirá en analizar un set de datos a elegir por los alumnos:
  - El alumno deberá entregar un archivo zip que contenga el RMarkdown del análisis de los datos, el HTML del mismo compilado y un video de duración 5-7 minutos explicando procedimiento y resultados.

En el examen parcial **sólo se podrá consultar un archivo con un conjunto de fórmulas** que se indicarán en su momento. Se podrá utilizar la ayuda de R.

Universidad de Navarra

El examen final tendrá preguntas teóricas y de resolución práctica por ordenador. Para la parte práctica **se podrán consultar los apuntes de ADI. No se podrá consultar Internet. No se podrá usar el móvil, ni portatil ni tablet.** En ADI no estarán visibles los exámenes de años anteriores. Como es lógico no se podrá preguntar a otra persona.

En caso de que por un **motivo justificado** (por ejemplo examenes o viajes oficiales, enfermedad con informe médico, enfermedad grave de un familiar, actividades formativas de la Universidad o sus Colegios Mayores, etc.) no se pueda hacer un test, el valor de ese test no se considerará en la asignatura y se disminuirá el porcentaje de los test no presentados con justificación un 2% en la nota final. Ese procentaje se añadirá al examen final. **En ningún caso ser repetirán los test fuera del horario establecido.**

Ejemplo de notas:

- Nota Examen parcial: 7, Nota Examen final: 6, No presenta trabajo,
- Tests prácticos: 4,5,6, NP, 5., Tests teóricos: 7,6,5,5,10

Las cuentas **si el no presentado está justificado** (como se ha mencionado antes) serían:

- Nota prácticos: (5+6+7)/3 = 6, Nota teóricos: (7+6+5+10)/4 = 7
- Nota total: 7 * 0.3 + 6 *.52 + 6*.08 + 7*.10 = 6.4

Si no estuviera justificado, la nota de ese examen sería un cero. Las nuevas cuentas serían:

- Nota prácticos: (4+5+6+7)/4 = 5.5, Nota teóricos: (7+6+5+10)/4 = 7
- Nota total: 7 * 0.3 + 6 *.5 + 5.5*.10 + 7*.10 = 6.35

El resumen, es que no presentarse a alguno de los test no supone una gran penalización en la nota final (en este ejemplo 5 centésimas de punto).

CONVOCATORIA EXTRAORDINARIA

Para calcular la nota final de la asignatura se tendrá en cuenta:

- **El examen final (100%).**

Para la convocatoria extraordinaria **no se tendrá en cuenta las notas de los test teóricos y prácticos ni el trabajo de la ordinaria**.

Si un alumno se quiere presentar al trabajo en la convocatoria extraordinaria, el trabajo:

- Deberá ser diferente al realizado en la convocatoria ordinaria y con un data set diferente a cualquier otro trabajo de otro grupo en la convocatoria ordinaria.
- Deberá ser individual.
- Al igual que en la convocatoria ordinaria, **el trabajo es opcional.**
  - Si un alumno se presenta la trabajo, para la nota final de la asignatura se tendrá en cuenta: **Examen final (85%).**
  - **Trabajo (15%).** (a **entregar antes de la fecha del examen de la convocatoria extraordinaria** y **defenderlo presencialmente** a fecha concordada con los profesores de la asignatura). No será necesario hacer el vídeo.

Es decir, en la convocatoria extraordinaria, no puede sacar una nota superior a 10 aunque se presente el trabajo.

HORARIOS DE ATENCIÓN

Dr Angel Rubio (arubio@unav.es), **Dr Juan Ferrer-Bonsoms** (jafhernandez@unav.es),

- Despacho M018. Edificio Miramón. Planta 0.
- Para resolución de dudas , se deberá concertar una reunión con este link

https://calendar.app.google/xzPK5tJA8k6wvoFK8

- Si ningún horario es adecuado, contactar con por email con alguno de los profesores.

BIBLIOGRAFÍA

La bibliografía es muy amplia. En la página web de la biblioteca se puede acceder a diversos libros que son gratuitos de forma electrónica para alumnos y profesores de la Universidad de Navarra.

De entre estos, los más interesantes son:

- Biostatistics with R: An Introduction to Statistics Through Biological Data, por Babak Shahbaba. Localizalo en la biblioteca (Formato electrónico)
- Introduction to Data Analysis and Graphical Presentation in Biostatistic with R: Statistic in the Large, by Thomas W. McFarland. Localízalo en la biblioteca (Formato electrónico)

Existen otros libros electrónicos gratuitos. Una lista se encuentra en http://stats.stackexchange.com/questions/170/free-statistical-textbooks.