Máster en Técnicas Estadísticas

DATOS IDENTIFICATIVOS DE LA MATERIA

Código de la materia: 614493111
Nombre de la materia: Métodos No Paramétricos
Número de créditos ECTS: 5
Curso académico: 2022/2023
Profesorado: José Antonio Vilar Fernández (coordinador y docente, 5 créditos)

OBJETIVOS DE LA MATERIA

Esta materia constituye la primera toma de contacto en el máster con técnicas no paramétricas de inferencia y modelización. Los objetivos a alcanzar como resultado del aprendizaje son:

• Familiarizarse con las técnicas no paramétricas básicas de estimación de la función de distribución de probabilidad, la función de densidad de probabilidad y la regresión.

• Conocer y saber aplicar las principales pruebas no paramétricas de bondad de ajuste y de asociación.

• Conocer con rigor las ventajas y limitaciones del enfoque no paramétrico en el análisis de datos.

• Desarrollar autonomía para aplicar herramientas no paramétricas en el análisis de datos, en situaciones complejas y/o multidisciplinares.

• Saber presentar el análisis de datos mediante técnicas no paramétricas a un público tanto especializado como no.

CONTENIDOS DE LA MATERIA

Tema 1. Estimación no paramétrica de la función de distribución.
La distribución empírica. Propiedades. Estimación de momentos y cuantiles.

Tema 2. Contrastes no paramétricos clásicos para una muestra.
Contrastes de bondad de ajuste a un modelo paramétrico: Prueba de Kolmogorov-Smirnov. Análisis de normalidad: Gráfico Q-Q, test de Lilliefors, test de Shapiro-Wilk, transformaciones para conseguir normalidad. Contrastes de localización: test de los signos, test de Wilcoxon de los rangos signados.

Tema 3. Tests para dos y más muestras.
Comparación de dos muestras: Test de Kolmogorv-Smirnov para dos muestras, test de Mann-Whitney-Wilcoxon. Generalizaciones a más de dos muestras: Test de Kruskal-Wallis, test de Friedman.

Tema 4. Tests basados en tablas de contingencia.
Análisis de tablas de contingencia. Tests χ^2 de bondad de ajuste, de homogeneidad y de independencia en tablas de contingencia.

Tema 5. Métodos de suavización: estimación no paramétrica de la función de densidad.
El histograma. Estimación tipo núcleo de la densidad. Medidas de error en la estimación de la función de densidad. Selección del parámetro de suavizado en la estimación tipo núcleo: validación cruzada y plug-in. Estimación de la densidad multivariante.

Tema 6. Estimación no paramétrica de la función de regresión.
Regresión tipo núcleo. Regresión polinómica local. Suavización por vecinos más próximos. Selección del parámetro de suavizado: validación cruzada y plug-in. Algoritmo loess. Breve introducción a la regresión por splines.

BIBLIOGRAFÍA BÁSICA Y COMPLEMENTARIA

Básica

Conover W.J. (1999) *Practical Nonparametric Statistics* (3rd edition). Willey Series in Probability and Statistics, 350.

Corder G.W., Foreman D.I. (2014) *Nonparametric Statistics: A Step-by-Step Approach* (2nd edition). John Willey & Sons.

Fan J., Gijbels I. (1996) *Local polynomial modelling and its applications*. Monographs on Statistics and Applied Probability 66. Chapman & Hall/CRC.

Gibbons J.D, Chakraborti S. (2010) *Nonparametric Statistical Inference* (5th edition). Statistics: Textbooks and Monographs. Chapman & Hall/CRC.

Hollander M., Wolfe D.A., Chicken E. (2014) *Nonparametric Statistical Methods* (3rd edition). Wiley Series in Probability and Statistics, Wiley.

Silverman, B. W. (1986) *Density Estimation for Statistics and Data Analysis*. Monographs on Statistics and Applied Probability 26. Chapman & Hall/CRC.

Wand M.P., Jones M.C. (1995) *Kernel smoothing*. Monographs on Statistics and Applied Probability 60. Chapman & Hall/CRC.

Complementaria

Bowman A.W., Azzalini A. (1997) *Applied Smoothing Techniques for Data Analysis*. Oxford: Oxford University Press.

McKean J.W., Kloke J. (2014) *Nonparametric Statistical Methods using R*. The R Series. Chapman and Hall/CRC.

Simonoff J.S. (1996) *Smoothing methods in statistics*. Springer Series in Statistics. New York: Springer.

Smeeton N.C, Sprent P. (2007) *Applied Nonparametric Statistical Methods* (4th edition) Chapman & Hall/CRC Texts in Statistical Science. Chapman & Hall/CRC.

Wasserman L. (2006). *All of Nonparametric Statistics*. Springer Texts in Statistics. New York: Springer.

COMPETENCIAS

En esta materia se trabajarán las competencias básicas, generales y transversales recogidas en la memoria del título. Se indican a continuación cuáles son las competencias específicas, que se potenciarán en esta materia:

E1 - Conocer, identificar, modelar, estudiar y resolver problemas complejos de estadística e investigación operativa, en un contexto científico, tecnológico o profesional, surgidos en aplicaciones reales.
E2 - Desarrollar autonomía para la resolución práctica de problemas complejos surgidos en aplicaciones reales y para la interpretación de los resultados de cara a la ayuda en la toma de decisiones.
E3 - Adquirir conocimientos avanzados de los fundamentos teóricos subyacentes a las distintas metodologías de la estadística y la investigación operativa, que permitan su desarrollo profesional especializado.
E4 - Adquirir las destrezas necesarias en el manejo teórico-práctico de la teoría de la probabilidad y las variables aleatorias que permitan su desarrollo profesional en el ámbito científico/académico, tecnológico o profesional especializado y multidisciplinar.
E5 - Profundizar en los conocimientos en los fundamentos teórico-prácticos especializados del modelado y estudio de distintos tipos de relaciones de dependencia entre variables estadísticas.
E6 - Adquirir conocimientos teórico-prácticos avanzados de distintas técnicas matemáticas, orientadas específicamente a la ayuda en la toma de decisiones, y desarrollar capacidad de reflexión para evaluar y decidir entre distintas perspectivas en contextos complejos.
E8 - Adquirir conocimientos teórico-prácticos avanzados de las técnicas destinadas a la realización de inferencias y contrastes relativos a variables y parámetros de un modelo estadístico, y saber aplicarlos con autonomía suficiente un contexto científico, tecnológico o profesional.

METODOLOGÍA DOCENTE

La enseñanza constará de clases expositivas e interactivas, así como de la tutorización del aprendizaje y de las tareas encomendadas al alumnado.

En las clases expositivas e interactivas se introducirán los fundamentos teóricos de las técnicas y se enfatizará su utilidad con ejemplos aplicados. Se propondrán y resolverán ejemplos con ayuda de scripts con código de software R, por lo que es necesario que el alumnado disponga de un ordenador en el desarrollo de estas sesiones.

Se propondrán actividades para el alumnado, que consistirán en la resolución de cuestiones, ejercicios y ejemplos relacionados con técnicas de inferencia y modelización no paramétrica.

Se proporcionarán apuntes de la materia, presentaciones y diverso material orientativo del aprendizaje del software. Los apuntes y otros instrumentos didácticos estarán disponibles a través de alguna herramienta de acceso por vía web (campus virtual, Teams, base de datos del máster).

CRITERIOS Y MÉTODOS DE EVALUACIÓN

Evaluación continua (30%): la evaluación continua se realizará en base a la resolución por

parte del alumnado, vía presencial o no presencial, de cuestiones y/o problemas propuestas por el docente, pudiendo requerirse el empleo de software y la redacción de informes correctamente estructurados. La calificación obtenida se conservará entre las oportunidades (ordinaria y extraordinaria) dentro de la convocatoria de cada curso. Con las distintas actividades que se propondrán a lo largo del curso, se valorará el nivel de adquisición de las competencias básicas CB6-CB10 y generales CG1-CG5. También se evaluará el nivel alcanzado en las competencias específicas E2, E3, E5 y E8. Así mismo se tendrá en cuenta en la evaluación el nivel alcanzado en las competencias transversales CT1-CT5.

Examen final (70%): el examen final constará de varias cuestiones teórico-prácticas sobre los contenidos de la materia, dentro de las que se podrá incluir la interpretación de resultados obtenidos con el lenguaje estadístico utilizado en la docencia interactiva (R). En el examen, se evaluarán las competencias específicas: E1, E2, E3, E4, E5, E6 y E8.

Presentación a la evaluación: se considera que un/a alumno/a concurre a una convocatoria cuando participa en actividades que le permiten obtener al menos un 50% de la evaluación final.

El peso de la evaluación continua en la oportunidad extraordinaria de recuperación (en el mes de julio) será el mismo que en la evaluación ordinaria. En la segunda oportunidad de evaluación (recuperación), se realizará un examen y la nota final será el máximo de tres cantidades: la nota de la evaluación ordinaria, la nota del nuevo examen y la media ponderada del nuevo examen y la evaluación continua.

TIEMPO DE ESTUDIO Y DE TRABAJO PERSONAL QUE DEBE DEDICAR UN ESTUDIANTE PARA SUPERAR LA MATERIA

Cada crédito ECTS se traduce en 7 horas de clase de tipo presencial. Se estima que el alumnado necesitará, por cada hora presencial, una hora adicional para la compresión global de los contenidos y otra hora y media para completar las actividades asociadas a ejercicios y a otras tareas propuestas. En total resultarán 24.5 horas por crédito ECTS.

RECOMENDACIONES PARA EL ESTUDIO DE LA MATERIA

Es conveniente que el alumnado posea conocimientos básicos de cálculo de probabilidades e inferencia paramétrica. También es recomendable disponer de unas habilidades medias en el manejo de ordenadores, y en concreto de software estadístico. Para un mejor aprendizaje de la materia, conviene tener presente el sentido práctico de los métodos que se están conociendo. Se aconseja participar activamente en el proceso de aprendizaje de la materia, incluyendo la asistencia y participación en las sesiones teóricas y prácticas.

RECURSOS PARA EL APRENDIZAJE

Bibliografía, software libre y material de apoyo proporcionado vía la plataforma web del Máster en Técnicas Estadísticas.

OBSERVACIONES

El desarrollo de los contenidos de la materia se realizará teniendo en cuenta que las competencias a adquirir por el alumnado deben cumplir con el nivel MECES3. En este

sentido, conviene enfatizar: (i) los contenidos de carácter inferencial de materias previas (estimación y contrastes) son radicalmente diferentes al considerar únicamente técnicas paramétricas, y (ii) los métodos no paramétricos abordados en esta materia son herramientas avanzadas de análisis de datos, de tal modo que, o bien no se incluyen en los diseños curriculares de la mayoría de titulaciones de grado o, si se imparten, lo hacen con carácter introductorio y sin el rigor y el carácter exhaustivo con que se propone impartir en esta materia. Así por ejemplo se establecerán las hipótesis estructurales que permiten la aplicación de las técnicas, sus limitaciones según los modelos de probabilidad y de regresión a examinar, y diferentes criterios de evaluación de su comportamiento.

Para los casos de realización fraudulenta de ejercicios o pruebas, será de aplicación lo recogido en las respectivas normativas de las universidades participantes en el Máster en Técnicas Estadísticas.

Esta guía y los criterios y metodologías en ella descritos están sujetos a las modificaciones que se deriven de normativas y directrices de las universidades participantes en el Máster en Técnicas Estadísticas.