



Estimación no paramétrica de funciones

(Fecha última actualización:04/10/2024)

Máster en Estadística Aplicada. CURSO ACADÉMICO 2024-2025	
Título	Estimación no paramétrica de densidades en R
Tipo	INVESTIGACIÓN <input checked="" type="checkbox"/> ORIENTACIÓN PRÁCTICA <input type="checkbox"/>
Número de alumnos	1
Profesor(es)/ email	María Dolores Martínez Miranda (mmiranda@ugr.es) María Luz Gámiz Pérez (mgamiz@ugr.es)
Descripción	<p>La estimación no paramétrica de densidades es una herramienta clave en Machine Learning dentro de lo que se denomina Aprendizaje No Supervisado. Se trata de comprender el proceso de generación de los datos que generó la muestra que tenemos. También juegan un papel central en algoritmos de clasificación.</p> <p>Desde el punto de vista de la Estadística, la estimación de la función de densidad constituye un problema “canónico”. Si tenemos la función de densidad lo tenemos todo, ya que proporcionan una descripción completa de la población.</p> <p>En Inferencia Estadística clásica la función de densidad habitualmente se estima usando modelos paramétricos, idealmente por máxima-verosimilitud. La inferencia no paramétrica se presenta como una alternativa más flexible donde se otorga a los datos un papel central. Bajo este último enfoque, y compartiendo terreno con el Machine Learning, los denominados estimadores kernel de la densidad (a veces traducidos como estimadores tipo núcleo) ofrecen una solución eficiente al problema y constituyen una de las herramientas más utilizadas hoy en día en la práctica.</p> <p>El entorno de análisis y programación estadística R proporciona diversas funciones para el cálculo de los estimadores kernel de la densidad. Algunas funciones elementales forman parte del sistema base como la función “density”. Además, existen varios paquetes específicos disponibles en CRAN que implementan los últimos avances en este campo.</p> <p>Esta propuesta de TFM supondrá para el estudiante en una aproximación computacional práctica al problema de estimación no paramétrica de la función de densidad. Una vez que haya adquirido las nociones teóricas básicas el resto del trabajo se centrará en aplicaciones prácticas en R usando datos reales y simulaciones.</p>
Objetivos particulares	<ul style="list-style-type: none"> • Formular y comprender el problema de estimación de la función de densidad desde el punto de vista no paramétrico. • Describir los estimadores kernel de la función de densidad más relevantes. • Comprender el papel que juega el parámetro de suavizado y describir métodos para su selección en la práctica. • Revisar y documentar las herramientas disponibles en R para la computación de los estimadores descritos. • Realizar aplicaciones con datos reales y simulados.
Prerrequisitos y recomendaciones	Se requiere destreza en el uso del entorno de programación y análisis estadístico R, incluyendo habilidad para estudiar paquetes específicos disponibles en CRAN y realizar tareas de programación de funciones. Aunque el enfoque del TFM es computacional se requiere una base adecuada de Inferencia Estadística.
Plan de trabajo	<ol style="list-style-type: none"> 1. Revisión de los aspectos teóricos. 2. Implementación computacional: estudio de herramientas disponibles R y creación de



	<ol style="list-style-type: none"> 3. funciones originales en su caso. 4. Aplicaciones con datos reales y datos simulados. 5. Reflexión sobre limitaciones y extensiones de la metodología considerada. 6. Elaboración de las conclusiones del trabajo.
<p>Competencias generales y específicas</p>	<p>Competencias generales:</p> <p>CG1 - Los titulados han de saber 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.</p> <p>CG7 - Los titulados han de realizar una contribución a través de una investigación original que amplíe las fronteras del conocimiento desarrollando un corpus sustancial, del que parte merezca la publicación referenciada a nivel nacional o internacional.</p> <p>CG8 - Los titulados deben ser críticos en el análisis, evaluación y síntesis de ideas nuevas y complejas.</p> <p>Competencias específicas:</p> <p>CE1 - Conocer métodos para el Análisis de Datos</p> <p>CE5 - Adquirir conocimientos avanzados en Inferencia Estadística</p> <p>CE17 - Adquirir capacidades de elaboración y construcción de modelos y su validación</p> <p>CE18 - Ser capaz de realizar un análisis de datos</p> <p>CE20 - Ser capaz de realizar una correcta representación gráfica de datos</p> <p>CE21 - Conocer, identificar y seleccionar fuentes estadísticas</p> <p>CE22 - Ser capaz de interpretar resultados a partir de modelos estadísticos</p> <p>CE24 - Ser capaz de extraer conclusiones y redactar informes</p> <p>CE25 - Ser capaz de identificar relaciones o asociaciones</p> <p>CE26 - Saber utilizar con destreza entornos de programación y análisis estadístico</p>
<p>Bibliografía</p>	<ol style="list-style-type: none"> 1. Cattaneo, M. D., Jansson, M., & Ma, X. (2022). lpdensity: Local Polynomial Density Estimation and Inference. Journal of Statistical Software, 101(2), 1–25. https://doi.org/10.18637/jss.v101.i02 2. Duong, T. (2018). ks: Kernel Smoothing. R package version 1.11.7. 3. García Portugués, E. (2022). Notes for Nonparametric Statistics. Version 6.5.9. ISBN 978-84-09-29537-1. https://bookdown.org/egarpor/NP-UC3M/ 4. Hastie, T., Tibshirani, R. y Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer Series in Statistics. Springer, New York. 5. Silverman, B. (1986). Density Estimation for Statistics and Data Analysis. Chapman and Hall, London. 6. Wand, M. (2021). KernSmooth: Functions for Kernel Smoothing Supporting Wand & Jones (1995). R package version 2.23-20. https://CRAN.R-project.org/package=KernSmooth 7. Wand, M. P. y Jones, M. C. (1995). Kernel Smoothing. Chapman and Hall, London.