



Estimación no paramétrica de funciones

(Fecha última actualización 30 septiembre 2022)

| Máster en Estadística Aplicada. CURSO ACADÉMICO 2022-2023 | |
|---|---|
| Título | Estimación no paramétrica de la función de densidad |
| Tipo | INVESTIGACIÓN <input checked="" type="checkbox"/> ORIENTACIÓN PRÁCTICA <input type="checkbox"/> |
| Profesor(es)/ email | María Dolores Martínez Miranda (mmiranda@ugr.es) María Luz Gámiz Pérez (mgamiz@ugr.es) |
| Descripción | <p>La inferencia no paramétrica constituye una herramienta flexible que permite descubrir y describir estructuras subyacentes en los datos. De ahí que sus principios sean claves en disciplinas en auge como Data Science o Machine Learning. Si bien en sus inicios pudo asociarse con la pura exploración de datos, hoy en día la inferencia no paramétrica ha alcanzado la madurez suficiente como para dar respuestas integrales a problemas complejos en todos los ámbitos de la ciencia dentro del contexto formal de la Estadística Matemática.</p> <p>Esta propuesta se dirige a un problema “canónico” de la Estadística. Se trata de comprender el proceso de generación de los datos que generó la muestra que tenemos. Para ello el foco se pone en la estimación de la función de densidad, función clave ya que proporciona una descripción completa de la población. En el contexto del Machine Learning la estimación de la función de densidad constituye uno de los problemas fundamentales del Aprendizaje No Supervisado con aplicaciones en problemas de clasificación.</p> <p>En Inferencia Estadística clásica la función de densidad 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, 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> |
| Objetivos particulares | <ul style="list-style-type: none"> • Formular 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. • Estudiar medidas de error apropiadas para los estimadores descritos y comprender aspectos clave en la inferencia no paramétrica de funciones y el aprendizaje automático como son el balance entre sesgo y varianza, y la complejidad del modelo. • Comprender el papel que juega el parámetro de suavizado y describir métodos para su selección en la práctica. • Ilustrar los métodos estudiados con datos reales y/o simulados usando paquetes específicos disponibles en CRAN. • Tomar conciencia de las ventajas e inconvenientes de los métodos estudiados. |
| Prerrequisitos y recomendaciones | Se requiere tener una base sólida en Estadística Matemática como la adquirida en un grado de Estadística o Matemáticas. También se requieren conocimientos básicos del entorno de programación y análisis estadístico R, concretamente el uso de paquetes especializados en CRAN y elementos básicos de programación. |
| Plan de trabajo | <ol style="list-style-type: none"> 1. Revisión de los aspectos teóricos. 2. Aplicaciones con datos. 3. Reflexión sobre limitaciones y extensiones de la metodología considerada. 4. Elaboración de las conclusiones del trabajo. |
| Competencias generales y específicas | Competencias generales: |



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.
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.
CG8 - Los titulados deben ser críticos en el análisis, evaluación y síntesis de ideas nuevas y complejas.

Competencias específicas:

CE1 - Conocer métodos para el Análisis de Datos
CE5 - Adquirir conocimientos avanzados en Inferencia Estadística
CE17 - Adquirir capacidades de elaboración y construcción de modelos y su validación
CE18 - Ser capaz de realizar un análisis de datos
CE20 - Ser capaz de realizar una correcta representación gráfica de datos
CE21 - Conocer, identificar y seleccionar fuentes estadísticas
CE22 - Ser capaz de interpretar resultados a partir de modelos estadísticos
CE24 - Ser capaz de extraer conclusiones y redactar informes
CE25 - Ser capaz de identificar relaciones o asociaciones
CE26 - Saber utilizar con destreza entornos de programación y análisis estadístico

Bibliografía

1. Cattaneo, M. D., Jansson, M., & Ma, X. (2020). Simple Local Polynomial Density Estimators. *Journal of the American Statistical Association. Theory and Methods*, 115 (531), 1449-1455.
2. García Portugués, E. (2022). Notes for Nonparametric Statistics. Version 6.5.9. ISBN 978-84-09-29537-1. <https://bookdown.org/egarpor/NP-UC3M/>
3. 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.
4. Silverman, B. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London.
5. Wand, M. P. y Jones, M. C. (1995). *Kernel Smoothing*. Chapman and Hall, London.