

Programa de máster
“Matemáticas y aplicaciones”
Departamento de Matemáticas (UAM)
Curso 2010-2011

Curso avanzado de Estadística

Profesor: Antonio Cuevas

ALCANCE Y OBJETIVOS DEL CURSO

El curso está centrado en algunos temas de “**estadística infinito-dimensional**”, utilizando esta expresión para designar aquellos problemas de estadística en los que los datos y/o los “parámetros” de interés son elementos de un espacio de dimensión infinita. El libro de Grenander (1981) es una referencia pionera en estos temas. En esa obra se presentan por primera vez, bajo el título *Abstract Inference*, los problemas de estimación infinito-dimensional en un marco unificado, con una gran amplitud y claridad de ideas. Históricamente, el precedente más claro de este tipo de problemas es quizá la **teoría de estimación no paramétrica**, cuyo origen se remonta a los años cincuenta del siglo XX y que ha alcanzado su apogeo en los años noventa. Esta teoría se ocupa de la estimación de funciones (típicamente la función de densidad o la de regresión) y, más recientemente, de la estimación de conjuntos. La información muestral consiste en una muestra aleatoria de observaciones en la recta real o en el espacio euclídeo. En este caso, por tanto, es el espacio paramétrico (es decir, el conjunto en el cual se encuentra el “parámetro” desconocido que se desea estimar) el que resulta ser infinito-dimensional.

A partir de los años noventa se ha desarrollado también con fuerza la **estadística con datos funcionales**, que se ocupa de aquellos problemas en los que son los elementos de la muestra son funciones y, por tanto, “viven” usualmente en un espacio de dimensión infinita. Esta teoría se ha visto impulsada por motivaciones estrictamente prácticas, relacionadas con la demanda de los usuarios de la estadística en diferentes ciencias experimentales (incluyendo diversas áreas de la economía y otras ciencias sociales) en las que se dispone de una creciente abundancia de datos “en tiempo continuo” que pueden considerarse como funciones.

El propósito de este curso es presentar una perspectiva actualizada de la teoría de estimación no paramétrica (estimación de la densidad, de la regresión y estimación de conjuntos) y de la estadística con datos funcionales. Se pondrá especial énfasis en aquellos aspectos que actualmente presentan una elaboración matemática más completa. Se prestará también atención a los aspectos computacionales, utilizando para ello el conocido software **R**.

PROGRAMA

1. Estimación funcional no paramétrica y sus aplicaciones

- 1.1 Estimación no paramétrica de las funciones de densidad y de regresión. Algunas clases importantes de estimadores: conceptos básicos, ejemplos, propiedades asintóticas.
- 1.2 El problema de selección de los parámetros de suavizado..
- 1.3 Estimación de conjuntos. Estimación del soporte y de conjuntos de nivel. Algunos resultados básicos. Aplicaciones.
- 1.4 Aspectos computacionales: el software **R**.

2. Una introducción a la metodología estadística con datos funcionales

- 2.1 Planteamiento general. Ejemplos. Datos longitudinales y datos funcionales.
- 2.2 Análisis exploratorio de datos funcionales. Medidas de profundidad.
- 2.3 Regresión y análisis de la varianza con datos funcionales. Metodología funcional de componentes principales.
- 2.4 El problema de clasificación supervisada (o discriminación). Planteamiento y resultados básicos. El caso funcional: sus diferencias con el problema análogo en dimensión finita.
- 2.5 The problema de clasificación no supervisada (o análisis de conglomerados) con datos funcionales. .
- 2.6 Aspectos computacionales: análisis de datos funcionales con **R**.

Bibliografía

1. Baíllo, A., Cuevas, A. y Fraiman, R. (2010). Classification methods with functional data. En *Oxford handbook on Statistics and FDA*, F. Ferraty y Y. Romain, eds. Oxford University Press (en prensa).
2. Cuevas, A. (2009). Set estimation: another bridge between statistics and geometry. *BEIO*, 25, 71-85.
3. Cuevas, A. y Fraiman, R. (2009). Set estimation. En *New Perspectives on Stochastic Geometry*, W.S. Kendall y I. Molchanov, eds., pp. 366-389. Oxford University Press.
4. Devroye, L., Györfi, L. y Lugosi, G. (1996). *A Probabilistic Theory of Pattern Recognition*. Springer-Verlag.
5. Ferraty, F. y Vieu, P. (2006). *Nonparametric Functional Data Analysis. Theory and Practice*. Springer.
6. Grenander, U. (1981). *Abstract Inference*. Wiley.
7. Györfi, L., Kohler, M., Krzyżak, A. y Walk, H. (2002). *A Distribution-Free Theory of Nonparametric Regression*. Springer-Verlag.

8. Hastie, T., Tibshirani, R. y Friedman, J. (2001). *The Elements of Statistical Learning*. Springer.
9. Maindonald, J. y Braun, J. (2003). *Data Analysis and Graphics Using R*. Cambridge University Press.
10. Ramsay, J.O. y Silverman, B.W. (2005). *Functional Data Analysis (2^a ed.)*. Springer.
11. Ramsay, J.O. y Silverman, B.W. (2002). *Applied functional data analysis methods and case studies*. Springer.
12. Scott, D.W. (1992). *Multivariate density estimation: theory, practice and visualization*. Wiley.
13. Simonoff, J.S. (1996). *Smoothing Methods in Statistics*. Springer.
14. Wand, M.P. y Jones, M.C. (1995). *Kernel Smoothing*. Chapman & Hall.