



Asignatura: Curso Avanzado de Estadística
Código: 30075
Centro: Facultad de Ciencias
Titulación: Máster en Matemáticas y Aplicaciones
Nivel: Máster M2
Tipo: Optativa
Nº de créditos: 8 ECTS

ASIGNATURA / COURSE TITLE

Curso Avanzado de Estadística / [Advanced Course in Statistics](#)

1.1. Código / Course number

30075

1.2. Materia / Content area

Estadística / [Statistics](#)

1.3. Tipo / Course type

Formación optativa / [Elective course](#)

1.4. Nivel / Course level

Máster M2 / [Master M2](#)

1.5. Curso / Year

2013-2014

1.6. Semestre / Semester

2º / [2nd \(Spring semester\)](#)

1.7. Número de créditos / Credit allotment

8 créditos ECTS / [8 ECTS credits](#)

1.8. Requisitos previos / Prerequisites

Conocimientos de probabilidad a nivel de *Probabilidad I*.
Conocimientos de estadística a nivel de *Estadística I* (ver Plan de Estudios de Grado en Matemáticas).

[Background on Probability Theory at the level of the course *Probability I*.
Statistical training at the level of the course *Statistics I* \(see the corresponding Programmes in the webpage of the Mathematics Department\).](#)



Asignatura: Curso Avanzado de Estadística
Código: 30075
Centro: Facultad de Ciencias
Titulación: Máster en Matemáticas y Aplicaciones
Nivel: Máster M2
Tipo: Optativa
Nº de créditos: 8 ECTS

1.9. Requisitos mínimos de asistencia a las sesiones presenciales / **Minimum attendance requirement**

60% de las clases / 60% of the classes

1.10. Datos del equipo docente / **Faculty data**

Docente(s) / **Lecturer(s)** Amparo Baíllo
Departamento de Matemáticas / **Department of Mathematics**
Facultad / **Faculty** Ciencias
Despacho - Módulo / **Office - Module** 203 - 8
Teléfono / **Phone**: +34 91 497 7698
Correo electrónico/**Email**: amparo.baillo@uam.es
Página web/**Website**: <http://www.uam.es/amparo.baillo>
Horario de atención al alumnado/**Office hours**: Se fija a petición individual del alumno / **To be fixed individually with each student**

1.11. Objetivos del curso / **Course objectives**

El curso está centrado en algunos temas de Estadística Multivariante. El objetivo es conocer algunos aspectos y técnicas de la teoría estadística que permite analizar datos de dimensión finita mayor que 1 (vectores). El análisis estadístico multivariante es una herramienta de investigación muy útil, tanto desde el punto de vista teórico (es un punto de partida para estudiar datos funcionales, se emplea en análisis de series temporales y en modelos de matemática financiera) como desde el punto de vista práctico (la mayoría de los datos reales disponibles en la actualidad son como mínimo multivariados).

El propósito de este curso es presentar algunos temas esenciales de la Estadística Multivariante que permitan a los estudiantes comprender y trabajar en temas de investigación actual como, por ejemplo, el análisis de datos de muy alta dimensión o dimensión infinita. Se pondrá especial énfasis en aquellos aspectos que presentan una elaboración matemática más interesante. También se describirán problemas de investigación de actualidad en cada uno de los temas del programa. Las clases teóricas se completarán con la implementación de muchas técnicas en programas de Matlab o R.

This course focuses on some topics of Multivariate Statistics. The aim is to introduce some aspects and techniques of this statistical theory developed to handle data of finite dimension larger than 1 (vectors). Multivariate statistical analysis is a very useful research tool, both from the theoretical viewpoint (a starting point for Functional Data Analysis; useful in Time Series



Asignatura: Curso Avanzado de Estadística
Código: 30075
Centro: Facultad de Ciencias
Titulación: Máster en Matemáticas y Aplicaciones
Nivel: Máster M2
Tipo: Optativa
Nº de créditos: 8 ECTS

Analysis and Mathematical Finance) and from the practical viewpoint (the majority of real data available nowadays are at least multivariate).

The objective of this course is to introduce the students to some essential aspects of Multivariate Statistics, which will allow them to understand and work on current research topics such as, for instance, the analysis of very high- or infinite-dimensional data. There will be a special emphasis on mathematically more interesting aspects. Current research topics will be mentioned all along the course. A fundamental complement to the theoretical lectures will be the implementation of many multivariate techniques in Matlab or R.

1.12. Contenidos del programa / Course contents

1. Descripción de datos multivariantes / Description of multivariate data

- 1.1. Descripción numérica: matriz de datos, vector de medias, matriz de covarianzas, medidas globales de variabilidad. Dependencia lineal.
- 1.2. Numerical description: data matrix, mean vector, covariance matrix, global measures of variability. Linear dependence.
- 1.3. Análisis gráfico y transformaciones.
- 1.4. Graphical analysis and transformations.

2. Distribuciones multivariantes / Multivariate distributions

- 2.1. Vectores aleatorios: Media, matriz de covarianzas, transformaciones lineales, medidas de dependencia.
- 2.2. Random vectors: mean, covariance matrix, linear transformations, dependence measures.
- 2.3. La distribución normal multidimensional. Distribuciones elípticas. Modelos de mixturas.
- 2.4. Multidimensional normal distribution. Elliptical distributions. Mixture models.

3. Reducción de la dimensión / Dimension reduction

- 3.1. Componentes principales.
- 3.2. Principal components.
- 3.3. Análisis factorial. Aplicación a finanzas.
- 3.4. Factor analysis. Applications fo finance.
- 3.5. Regresión PLS (*Partial Least Squares*).
- 3.6. PLS regression.

4. Estimación no paramétrica multidimensional / Multivariate nonparametric estimation

- 4.1. Estimación *kernel* de densidades.



Asignatura: Curso Avanzado de Estadística
Código: 30075
Centro: Facultad de Ciencias
Titulación: Máster en Matemáticas y Aplicaciones
Nivel: Máster M2
Tipo: Optativa
Nº de créditos: 8 ECTS

- 4.2. [Kernel density estimation.](#)
- 4.3. Regresión no paramétrica: estimadores *kernel* y polinómico local.
- 4.4. [Nonparametric regression: kernel and local linear estimators.](#)

5. Análisis de conglomerados / **Cluster analysis**

- 5.1. Medidas de distancia y similaridad.
- 5.2. [Distance and similarity measures.](#)
- 5.3. El algoritmo de *k*-medias.
- 5.4. [k-means algorithm.](#)

6. Clasificación o análisis discriminante / **Classification or discriminant analysis**

- 6.1. Introducción al problema matemático. Regla Bayes. Metodología *plug-in*.
- 6.2. [Introduction to the mathematical problem. Bayes rule. Plug-in methodology.](#)
- 6.3. La regla lineal de Fisher.
- 6.4. [Fisher's linear rule.](#)
- 6.5. La regla de *k* vecinos más próximos.
- 6.6. [k-nearest neighbours rule.](#)
- 6.7. Metodología basada en medidas de profundidad.
- 6.8. [Depth-based methodologies.](#)

1.13. Referencias de consulta / **Course bibliography**

- [1] Barker, M. y Rayens, W. (2003). Partial least squares for discrimination. *Journal of Chemometrics*, 17, 166-173.
- [2] Cuadras, C. M. (2010). *Nuevos Métodos de Análisis Multivariante*. CMC Editions.
- [3] Cudeck, R. (2007). *Factor Analysis at 100 Historical Developments and Future Directions*. Lawrence Erlbaum Associates.
- [4] Devroye, L., Györfi, L. y Lugosi, G. (1996). *A Probabilistic Theory of Pattern Recognition*. Springer, New York.
- [5] Fan, J. (1992). Design-adaptive nonparametric regression. *Journal of the American Statistical Association*, 87, 998-1004.
- [6] Fujikoshi, Y. (2010). *Multivariate Statistics: High-Dimensional and Large-Sample Approximations*. Wiley.
- [7] Ghosh, A. K. y Chaudhuri, P. (2005). On maximum depth and related classifiers. *Scandinavian Journal of Statistics*, 32, 327-350.
- [8] Haenlein, M. y Kaplan, A. M. (2004). A beginner's guide to partial least squares analysis. *Understanding Statistics*, 3, 283-297.
- [9] Hall, P., Park, B. U. y Samworth, R. J. (2008). Choice of neighbor order in nearest-neighbor classification. *The Annals of Statistics*, 36, 2135-2152.
- [10] Härdle, W. (2007). *Applied Multivariate Statistical Analysis*. Springer.



Asignatura: Curso Avanzado de Estadística
Código: 30075
Centro: Facultad de Ciencias
Titulación: Máster en Matemáticas y Aplicaciones
Nivel: Máster M2
Tipo: Optativa
Nº de créditos: 8 ECTS

- [11] Helland I. S. (1990). PLS regression and statistical models. *Scandinavian Journal of Statistics*, 17, 97-114.
- [12] Izenman, A. J. (2008). *Modern Multivariate Statistical Techniques. Regression, Classification and Manifold Learning*. Springer.
- [13] Johnson, R. A. y Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis*. Prentice Hall.
- [14] de Jong, S. (1993). PLS fits closer than PCR. *Journal of Chemometrics*, 7, 551-557.
- [15] McLachlan, G. y Peel, D. (2000). *Finite Mixture Models*. Wiley.
- [16] Peña, D. (2002). *Análisis de datos multivariantes*. McGraw-Hill.
- [17] Ruppert, D. y Wand, M. P. (1994). Multivariate locally weighted least squares regression. *The Annals of Statistics*, 22, 1346-1370.
- [18] Schneeweiss, H. y Mathes, H. (1995). Factor analysis and principal components. *Journal of Multivariate Analysis*, 55, 105-124.
- [19] Scott, D.W. (1992). *Multivariate Density Estimation: Theory, Practice and Visualization*. Wiley.
- [20] Serfling, R. y Zuo, Y. (2000). General notions of statistical depth function. *The Annals of Statistics*, 28, 461-482.
- [21] Simonoff, J.S. (1996). *Smoothing Methods in Statistics*. Springer.
- [22] Stone, C. J. (1977). Consistent nonparametric regression. *The Annals of Statistics*, 5, 595-645.
- [23] Wand, M.P. y Jones, M.C. (1995). *Kernel Smoothing*. Chapman & Hall.

2. Métodos docentes / Teaching methodology

Clases presenciales y discusión personal a través de tutorías o lecturas dirigidas. Presentación de trabajos.
Classroom sessions and personal discussion via tutorship activities or guided readings. Written essays based on the classroom materials.

3. Tiempo de trabajo del estudiante / Student workload

		Nº de horas	
Presencial	Clases teóricas	42h(21%)	68h (33%)
	Clases prácticas	4 h (2%)	
	Tutorías	10 h (5%)	
	Seminarios y trabajos	10 h (5%)	



Asignatura: Curso Avanzado de Estadística
Código: 30075
Centro: Facultad de Ciencias
Titulación: Máster en Matemáticas y Aplicaciones
Nivel: Máster M2
Tipo: Optativa
Nº de créditos: 8 ECTS

		Nº de horas	
	Examen final / proyecto	2h (1%)	
No presencial	Elaboración de problemas	80h(40%)	132h (66%)
	Estudio semanal	46h(23%)	
	Preparación del examen	6h(3%)	
Carga total de horas de trabajo: 25 horas x 8 ECTS		200 h	

4. Métodos de evaluación y porcentaje en la calificación final / Evaluation procedures and weight of components in the final grade

Examen y/o elaboración y presentación oral de un trabajo: 50% de la nota final.

Entrega de ejercicios, presentaciones orales y participación en clase: 50% de la nota final.

Final exam (elaboration and oral presentation of a research or review work): 50% of final mark.

Assignments and active participation in lectures: 50% of final mark.

EVALUACIÓN EXTRAORDINARIA / Make up exam:

Examen ante tribunal de Máster / Examination by a committee.

5. Cronograma* / Course calendar

Semana Week	Contenido Contents	Horas presenciales Contact hours	Horas no presenciales Independent study time
1	Apartado 1	4	6
2	Apartado 2.1	4	7
3	Apartado 2.2	4	9
4	Apartados 3.1 y 3.2	5	10
5	Apartado 3.3	4	9
6	Apartado 1.5	5	10



Asignatura: Curso Avanzado de Estadística
Código: 30075
Centro: Facultad de Ciencias
Titulación: Máster en Matemáticas y Aplicaciones
Nivel: Máster M2
Tipo: Optativa
Nº de créditos: 8 ECTS

Seman a Week	Contenido Contents	Horas presenciales Contact hours	Horas no presenciales Independent study time
7	Apartado 4.1	5	10
8	Apartado 4.2	4	7
9	Apartado 5.1	4	9
10	Apartado 5.2	5	10
11	Apartados 6.1 y 6.2	4	6
12	Apartado 6.3	5	10
13	Apartado 6.4	5	11
14	Exposición de trabajos por los alumnos / presentation of students' works	4	9
15	Exposición de trabajos por los alumnos / presentation of students' works	4	9

*Este cronograma tiene carácter orientativo / This course calendar should be considered just as an approximate guidance.