

Universidad
Autónoma de Madrid



UNIVERSIDAD AUTÓNOMA DE MADRID

GUÍA DOCENTE



Asignatura: INTRODUCCIÓN AL ANÁLISIS BAYESIANO DE DATOS

Código: 31215

Tipo: Seminario

Nivel: Posgrado

Centro: Facultad de Psicología

Titulación: Máster en Metodología de las Ciencias Sociales y de la Salud

Curso académico: 2017-2018



1. ASIGNATURA / SUBJECT

1.1. Nombre / Name

INTRODUCCIÓN AL ANÁLISIS BAYESIANO DE DATOS

1.2. Código / Code

1.3. Tipo / Type

SEMINARIO

1.4. Nivel / Level

POSGRADO

1.5. Curso / Year

POSGRADO

1.6. Semestre / Semester

SEGUNDO

1.7. Número de créditos / Credits Number or ECTS

2



1.8. Idioma en el que se imparte / **Subject Language**

Español

1.9. Requisitos Previos / **Previous Aptitudes or Previous Requirements**

Es requisito imprescindible simultanear, o haber cursado previamente, la asignatura Fundamentos de Estadística Teórica. También es necesario cursar Técnicas de Simulación.

Son necesarios conocimientos previos de matemáticas a nivel de bachillerato: funciones, derivadas e integrales.

Conocimientos de probabilidad y estadística al nivel de las asignaturas de grado en ciencias sociales: distribuciones, probabilidad condicionada y valor esperado de una variable aleatoria. Estimación por máxima verosimilitud. Contraste de hipótesis.

En el seminario se realizan lenguaje R. Es útil conocer la programación informática a nivel elemental.

1.10. ¿ Aula ? / **Classroom?**

1.11. ¿ Es obligatoria la asistencia? / **Is it obligatory to attend the courses?**

No

1.12. Objetivos y Competencias a Desarrollar / **Objectives and Competitions to Develop**

El propósito del seminario es proporcionar una introducción al análisis bayesiano aplicado mediante técnicas analíticas y de simulación. En concreto, los objetivos detallados son:



1. Conocimiento de los conceptos básicos de inferencia estadística desde el enfoque clásico y el bayesiano.
2. Análisis matemático de la estimación bayesiana.
3. Estimación mediante simulación de Monte Carlo con cadenas de Markov (MCMC).
4. Realizar análisis de datos en casos reales utilizando los lenguajes de programación R y Stan.

Las competencias específicas que se trabajan en el seminario son

CB10 - Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida auto-dirigido o autónomo.

CG2 - Desarrollar el razonamiento crítico y la capacidad para realizar análisis y síntesis de la información disponible.

CE4 - Analizar datos identificando diferencias y relaciones. Esto implica conocer las diferentes herramientas de análisis así como su utilidad y aplicabilidad en cada contexto.

1.13. Contenidos del Programa / Program Contents

1. Introducción a la probabilidad. Función de probabilidad y densidad, valor esperado y varianza de una variable aleatoria. Probabilidad condicionada. Teorema de la probabilidad total. Distribución normal, Bernoulli y Poisson.
2. Teorema de Bayes. Distribución a-priori y a-posteriori.
3. Inferencia bayesiana. Estimador modal a posteriori y esperado a posteriori. Varianza a posteriori.
4. Simulación MCMC, conceptos básicos de la estimación mediante simulación.
5. Programación de la estimación bayesiana en Stan.



6. Estimación de la media de una distribución normal con varianza conocida y desconocida.
7. Análisis de datos dicotómicos: modelo beta-binomial.
8. Modelo lineal generalizado. Función de enlace, variable dependiente e independiente, heterocedasticidad, modelos jerárquicos y de efectos aleatorios.



1.14. Referencias de Consulta Básicas / **Basic References of Consultation, Recommended Bibliography**

El material básico para cursar el seminario son los apuntes proporcionados por el profesor. También será necesario manejar los apuntes de la asignatura Fundamentos de Estadística Teórica.

De manera complementaria, puede ampliarse información en el libro:

Revuelta, J. y Ponsoda, V. (2005). *Fundamentos de estadística*, segunda edición. Madrid. UNED Ediciones.

Los siguientes libros constituyen una referencia esencial sobre estadística bayesiana. Pueden utilizarse para ampliar conocimientos sobre estadística Bayesiana partiendo de la base adquirida en el seminario:

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A. y Rubin, D. B. (2013). *Bayesian Data Analysis*, Third Edition. Boca Raton, FL. Chapman & Hall/CRC Texts in Statistical Science.

Kruschke, J. K. (2015). *Doing Bayesian Data Analysis, : A Tutorial with R, JAGS, and Stan*. Second Edition. Academic Press / Elsevier.

Libros de texto sobre probabilidad y estadística, con mayor énfasis en la estadística clásica que en la bayesiana:

Peña, D. (2008). *Fundamentos de estadística*, segunda edición. Madrid. Alianza editorial.

Ruiz-Maya, L. (2004). *Fundamentos de inferencia estadística*, tercera edición. Madrid. AC.



2. Datos del Profesor / **Professor Data**

2.1. Nombre y Apellidos / **Name and Surname**

Javier Revuelta

2.2. Centro / **Department or Centre**

Facultad de Psicología
Departamento de Psicología Social y Metodología
Universidad Autónoma de Madrid

2.3. Despacho / **Office**

Despacho número 5 (edificio anexo de Psicología)

2.4. Correo Electrónico / **E -Mail address**

javier.revuelta@uam.es

2.5. Teléfono / **Telephone Number**

914973242

2.6. Web del profesor / **Professor Website**

Los materiales estarán disponibles en la plataforma moodle de la UAM (<https://moodle.uam.es/>).

2.7. ¿Horario de tutoría? / **Professor's Consultations Timetable**

A convenir con el alumno.



3. Métodos Docentes / Educational Methods

El seminario comprende tres tipos de actividades:

1. Estudio teórico. Estudio de los conceptos teóricos.
2. Resolución de ejercicios de lápiz y papel utilizando cálculo diferencial e integral.
3. Ejercicios de informática. Programación de la estimación bayesiana en el lenguaje Stan. También se utilizará el lenguaje R para preparar los datos y analizar los resultados.

4. Tiempo estimado de Trabajo del Estudiante / Estimated work task student time

El tiempo estimado de trabajo por parte del alumno es de 50 horas a lo largo del cuatrimestre. La distribución de esas horas en función del tipo de actividad es la que se resume en la siguiente tabla:

Tipo de actividad	Horas
Estudio teórico	15
Ejercicios de lápiz y papel	15
Ejercicios de informática	20
	50



5. Métodos de Evaluación y Porcentaje en la Calificación Final / **Evaluation Methods and Percentage in the Final Qualification**

La calificación del seminario se realizará del siguiente modo en cada una de las convocatorias:

- Convocatoria ordinaria. La convocatoria ordinaria se basa en un sistema de evaluación continua.

Se realizarán ejercicios prácticos obligatorios a lo largo del seminario, aproximadamente uno por semana o cada dos semanas. Cada ejercicio se puntuará de 0 a 10 y la media resultante será la calificación obtenida en la convocatoria ordinaria.

- Convocatoria extraordinaria. Los alumnos que no hayan superado el seminario en la convocatoria ordinaria pueden presentarse a una prueba única de recuperación durante el mes de junio puntuada de 0 a 10. Las calificaciones obtenidas en los ejercicios correspondientes a la evaluación continua no se tendrán en cuenta para la convocatoria extraordinaria.



6. Cronograma de Actividades / Activities Cronogram

El seminario se imparte entre los meses de febrero y abril.

El siguiente cronograma está organizado en un calendario de siete semanas. El cronograma describe la secuencia en que se presentan los conceptos fundamentales y se realizan las prácticas a lo largo del curso.

Semana	Actividad Teórica	Actividad Práctica
1	Introducción a la probabilidad Variable aleatoria Valor esperado y varianza	Ejercicios de lápiz y papel
2	Probabilidad condicionada y teorema de Bayes Distribuciones discretas y continuas.	Ejercicios de lápiz y papel
3	Estimación máximo verosímil y bayesiana Errores típicos	Ejercicios de lápiz y papel
4	Inferencia bayesiana mediante MCMC	Ejercicios de lápiz y papel Prácticas con R y Stan
5	Estimación de la distribución normal	Prácticas con R y Stan
6	Análisis de datos dicotómicos	Prácticas con R y Stan
7	Modelos de regresión	Prácticas con R y Stan