



MODELOS ESTADÍSTICOS AVANZADOS (47105)
MODELOS ESTADÍSTICOS (40029) del Grado en Matemáticas (394)

Grado en ESTADÍSTICA (549) curso 2016/2017

Curso 4º **Primer Cuatrimestre**

Carácter: **Optativa** Créditos: **6 ECTS**

Profesor: **José A. Menéndez**

e-mail: josan@eio.uva.es; web personal: <http://www.eio.uva.es/~josan/>

Tfno: +34 983 42 30 00 ext. 4169 /// Fax: +34 983 42 30 13

Departamento de Estadística e Investigación Operativa, oficina A230.

Facultad de Ciencias:

Campus Miguel Delibes. Paseo de Belén, nº 7. 47011 Valladolid. España.

Universidad de Valladolid

La asignatura MODELOS ESTADÍSTICOS (40029) del Grado en Matemáticas se imparte conjuntamente con la asignatura MODELOS ESTADÍSTICOS AVANZADOS (47105) del Grado en Estadística, con la que está “amaestrada”.

OBJETIVOS

Este curso está orientado a cualquier estudiante que tenga una formación básica en regresión lineal, y proporcionará una introducción de diferentes modelos estadísticos, desde una perspectiva aplicada. El objetivo prioritario es el desarrollo de la capacidad de ajustar estos modelos en diferentes contextos, para lo cual se expondrán, además de la teoría básica, diversos ejemplos con datos reales en los que pueda ser útil el ajuste de los modelos presentados. El estudiante aprenderá a ajustar modelos no lineales normales, modelos lineales generalizados, ecuaciones estimadoras generalizadas y modelos mixtos, con aplicaciones variadas en el estudio de datos longitudinales o de medidas repetidas con observaciones correladas.

Por último, un objetivo, que se verá cumplido a la par que los anteriores, es el manejo de R y SAS para el ajuste de los modelos en las aplicaciones.

PROGRAMA

Introducción

1. Modelos de Regresión Lineal y No Lineal para una respuesta Normal

Modelos de regresión lineal y no lineal.

Estimación de parámetros.

Inferencia estadística de parámetros y predicción.

Introducción al ajuste de modelos con R y SAS.

2. El Modelo Lineal Generalizado (MLG)

La familia exponencial de distribuciones con parámetro de dispersión.
Elementos del MLG.
Estimación de parámetros: ecuaciones estimadoras y algoritmo IRWLS.
Inferencias sobre los parámetros del modelo.
Descripción de los MLG's más habituales: normal, binomial, multinomial, poisson, binomial negativa, gamma, normal inversa.
Estimación del parámetro de dispersión.
Selección de un MLG. La deviance y el AIC.
Validación del modelo. Residuos. Influencia.

3. Modelos de Regresión Logística

El modelo de regresión logística.
Interpretación y uso. Modelos alternativos.
Predicción. Estimación de la "dosis efectiva".
Reglas de clasificación. Curvas ROC.
Sobredispersión. Quasi-verosimilitud binomial.
Modelo mixto logístico-normal: verosimilitud marginal exacta.
Regresión logística condicional. Muestras apareadas.
Regresión logística multinomial nominal y ordinal.

4. Modelos de Regresión de Poisson

El modelo de regresión de poisson.
Inferencias sobre parámetros. Predicción.
Modelos log-lineales para el análisis de tablas de contingencia.
Sobredispersión. Quasi-verosimilitud de poisson.
Regresión binomial negativa.

5. Otros Modelos Lineales Generalizados

El modelo gamma. Estimación de la dispersión.
El modelo normal inverso.

6. Ecuaciones Estimadoras Generalizadas

Datos longitudinales.
Modelos para la correlación.
Ecuaciones estimadoras generalizadas para la quasi-verosimilitud. Solución iterativa mediante IRWLS.
Estimación directa y robusta de la matriz de covarianzas.

7. Modelos Lineales Mixtos Generalizados

Introducción del modelo lineal mixto.
El modelo lineal mixto generalizado.
Inferencias sobre los coeficientes del modelo y sobre las componentes de la varianza.
Ajuste del modelo.

BIBLIOGRAFÍA

Básica:

- Agresti, A. (2015). *Foundations of Linear and Generalized Linear Models*. Wiley.
- James, G., Witten, D., Hastie, T.J., Tibshirani, R.J. (2013). *An Introduction to Statistical Learning: With Applications in R* (corrected 6th printing 2015). Springer. <http://www-bcf.usc.edu/~garth/ISL/ISLR%20First%20Printing.pdf>
- Myers, R.H., Montgomery, D.C., Vining, G.G. and Robinson, T.J. (2010) *Generalized Linear Models with Applications in Engineering and the Sciences*, second edition. Wiley.

- Madsen H. and Thyregod, P. (2011). *Introduction to General and Generalized Linear Models*. CRC Press.
- Pawitan Yudi. (2004). *In All Likelihood: Statistical Modelling and Inference Using Likelihood*. Oxford Science Publications.

Complementaria:

- Agresti, A. (2013). *Categorical Data Analysis*, (3rd edition). Wiley.
- Collett D. (2003). *Modelling Binary Data*, (2nd edition). Chapman & Hall/CRC.
- Davison, A. C. (2003). *Statistical Models*. Cambridge University Press.
- Dobson, A.J. and Barnett, A.G. (2008). *An Introduction to Generalized Linear Models*, third edition. Chapman & Hall/CRC.
- Fahrmeir, L. and Tutz, G. (2001). *Multivariate Statistical Modelling Based on Generalized Linear Models*, second edition. Springer.
- Faraway, J.F. (2006). *Extending the Linear Model with R, Generalized Linear, Mixed Effects and Nonparametric Regression Models*. Chapman & Hall/CRC.
- Gelman, A. and Hill, J. (2007). *Data Analysis using Regression and Multilevel/Hierarchical Models*. Cambridge.
- Harrell, Frank (2015). *Regression Modeling Strategies* (2nd edition). Springer.
- Hastie, T.J. and Tibshirani, R.J. (1990). *Generalized Additive Models*. Chapman & Hall.
- Hastie, T.J., Tibshirani, R.J and Friedman, J. (2008). *The Elements of Statistical Learning, Data Mining, Inference, and Prediction*, (2nd edition). Springer.
- Hilbe, Joseph M. (2011). *Negative Binomial Regression*, (2nd edition). Cambridge.
- Hilbe, Joseph M. (2009). *Logistic Regression Models*. Chapman & Hall/CRC.
- Hardin, J.W. and Hilbe, J.M. (2003). *Generalized Estimating Equations*. Chapman & Hall/CRC.
- Jong, P. and Hellen G.Z. (2008). *Generalized Linear Models for Insurance Data*. Cambridge.
- Lee, Y., Nelder, J.A. and Pawitan, Y. (2006). *Generalized Linear Models with Random Effects*. Chapman & Hall/CRC.
- Maindonald, J. and Braun, W.J. (2010). *Data Analysis and Graphics Using R*, third edition. Cambridge.
- Tutz, Gerhard (2012). *Regression for Categorical Data*. Cambridge University Press.
- Weiss, R.E. (2005). *Modeling Longitudinal Data*. Springer.
- R Development Core Team (2010). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.

Se dispone de la bibliografía en la biblioteca de la facultad. En la página web del profesor (<http://www.eio.uva.es/~josan/>) se puede encontrar el resto del material necesario para seguir la asignatura, una colección de ejercicios, junto a la resolución de algunas aplicaciones cuyos resultados serán objeto de análisis a lo largo del curso.

METODOLOGÍA

Clases:

- El profesor presentará diversos problemas reales de distintos ámbitos de aplicación en los que se precisa la utilización de los métodos que el estudiante aprenderá a manejar en la asignatura.
- La teoría básica necesaria será expuesta en clase por el profesor de la asignatura y se ilustrará su aplicación mediante ejemplos. Esto hace difícil diferenciar claramente entre clases de teoría y clases prácticas. No obstante, podemos estimar que la "teoría" ocupará un 25% del tiempo total dedicado a las clases.

- Los estudiantes realizarán prácticas de ordenador en el Laboratorio de Estadística para familiarizarse con la aplicación de R y SAS a la resolución de problemas.
- Los estudiantes tendrán a su disposición los resultados de los análisis de diversos casos reales, cuya interpretación ocupará buena parte del tiempo dedicado a las clases. Esta actividad se realizará en el aula y en la mayoría de las ocasiones consistirá en prácticas tuteladas.

Trabajos:

- Se realizarán dos trabajos propuestos por el profesor en los plazos que se indicarán oportunamente. El informe de cada trabajo deberá ir firmado por su autor o autores. En el caso de que haya varios autores, cada uno de ellos deberá presentar en uno o dos folios un resumen personal explicando el trabajo realizado y sus aportaciones principales al mismo. El informe será revisado y valorado por el profesor, tanto en contenidos como en presentación, pudiendo ser requeridas de los alumnos cuantas explicaciones se consideren oportunas. Cada estudiante tendrá acceso a su informe, debidamente revisado y valorado.

Exámenes Parciales:

- Se realizarán dos exámenes parciales de una hora de duración.

Examen Final: 12 de Enero de 2017 (recuperación el día 31 de Enero)

Tutorías:

- Las **tutorías individualizadas** podrán ser atendidas los lunes, martes y jueves de 16:30 a 18:30, dentro del período lectivo, en el Departamento de Estadística. Fuera del horario anterior podrá consultarse al profesor previa cita con el mismo.

El profesor solicitará de los estudiantes al finalizar el curso la realización de una **encuesta específica** de la asignatura, con objeto de conocer su opinión sobre distintos aspectos de la misma.

EVALUACIÓN

La evaluación se llevará a cabo de la siguiente forma:

Denotemos por **T1** y **T2** las notas en cada uno de los dos trabajos, por **P1** y **P2** las notas en cada uno de los exámenes parciales, y por **EF** la nota en el examen final de la convocatoria ordinaria. Las notas se darán en una escala de 0 a 10.

Consideremos las ponderaciones: $T=0.10*T1+0.15*T2$; $P=0.10*P1+0.15*P2$; $F=0.5*EF$;

Si $T \geq 1$, $P \geq 1$ y $EF \geq 3$, entonces la calificación final será $C = \max(T+P+F, EF)$.

En otro caso, la calificación final será $C = EF$.

Caso de suspender en la convocatoria ordinaria, será posible la recuperación mediante el examen extraordinario. En este caso, la calificación final de la asignatura será la del examen extraordinario.

Cronograma de Actividades

Horario: Miércoles y Viernes de 13 a 14 h. Jueves de 10 a 12 h.

Laboratorio: 308.

Fechas de entrega de los trabajos: **T1, T2.**

Exámenes Parciales: **P1, P2.**

2016 SEPTIEMBRE

semana	Lunes	Martes	Miércoles	Jueves	Viernes
1	5	6	7	8	9
2	12	13	14	15	16
3	19	20	21	22	23
4	26	27	28	29	30

2016 OCTUBRE

semana	Lunes	Martes	Miércoles	Jueves	Viernes
5	3	4	5	6	7
6	10	11	12	13	14
7	17	18	19	20	21
8	24	25	26	27	28 T1

2016 NOVIEMBRE

semana	Lunes	Martes	Miércoles	Jueves	Viernes
9	31	1	2	3	4 P1
10	7	8	9	10	11 S Alberto
11	14	15	16	17	18
12	21	22	23	24	25
13	28	29	30	1	2 P2

2016 DICIEMBRE

semana	Lunes	Martes	Miércoles	Jueves	Viernes
14	5	6	7	8	9
15	12	13	14	15	16 T2
16	19	20	21	22	23
	26	27	28	29	30

2017 ENERO

	Lunes	Martes	Miércoles	Jueves	Viernes
	2	3	4	5	6
17	9	10	11	12 MEA-exfin	13
18	16	17	18	19	20
	23	24	25	26	27
	30	31 MEA-extra	1	2	3

Las fechas propuestas en el cronograma pueden verse alteradas por la necesidad de realizar algún cambio no previsto en el momento de su elaboración.