

MODELOS LINEALES

Curso 2010/2011

(Código:2115210-)

1.PRESENTACIÓN

Se proporcionan los fundamentos teóricos que justifican el uso de algunas técnicas estadísticas clásicas. Concretamente, se inicia al alumno en el ajuste de modelos útiles para la descripción y predicción de relaciones estocásticas. La metodología más elemental es la basada en los Modelos Lineales, que dan nombre a esta asignatura. Será preciso adquirir agilidad en el manejo de la distribución de algunos tipos de vectores aleatorios, así como de formas lineales y cuadráticas, para profundizar en el estudio de procedimientos de inferencia –estimación y contraste de hipótesis– que informen sobre las posibles fuentes de variación de una respuesta cuantitativa.

2.CONTEXTUALIZACIÓN

Se trata de una materia básica dentro del área de *Estadística*, en la especialidad de *Estadística e Investigación Operativa*. Supone el primer contacto del alumno con los *modelos estadísticos* como medio de representación de la realidad observable y un punto de partida para el estudio de materias más avanzadas como la “Inferencia Estadística Robusta” o el “Análisis Estadístico Multivariante”.

3.REQUISITOS PREVIOS RECOMENDABLES

El desarrollo de la *teoría de los modelos lineales* que se adopta en esta asignatura se apoya fuertemente en la representación matricial -compacta- de los mismos; por ello, es necesario tener algunos conocimientos de *álgebra matricial: operaciones con matrices, particionamiento de matrices, autovalores y autovectores de matrices definidas y semidefinidas positivas*. Igualmente, se precisa estar familiarizado con la teoría elemental de *vectores aleatorios*, al nivel que proporciona un curso introductorio sobre *probabilidad*, y con los conceptos básicos de los *métodos de inferencia estadística* -referidos a *muestras de una y dos poblaciones*.

4.RESULTADOS DE APRENDIZAJE

En líneas generales, se persiguen dos fines: por un lado, poner en contacto al alumno con el mundo de los *modelos estadísticos*, que protagonizará el resto de su formación en dicha materia, y por otro -el principal, dar a conocer la teoría que soporta a los métodos más elementales y, posiblemente, más utilizados en el análisis de datos procedentes de la observación o experimentación. Concretamente, cuando finalice esta asignatura, el alumno debería poseer los/as siguientes:

- Conocimientos:

1. Definición, estructura y propiedades de la distribución de probabilidad *normal multivariante*.
2. Distribución de probabilidad de las *formas cuadráticas* sobre vectores aleatorios.
3. Metodología básica para la *elaboración y el contraste de teorías* cuando se propone un *modelo de regresión lineal* como medio de representación de la asociación presente en los datos.
4. Análogamente para los *modelos de análisis de la varianza*.
5. Visión unificada de ambos enfoques y previsión de desarrollos futuros.

- Destrezas y habilidades:

1. Aplicar la teoría para justificar razonadamente los procedimientos que se utilizan en el análisis de datos mediante modelos lineales.
2. Utilizar la base teórica para ejecutar rigurosamente la selección del modelo, su ajuste y la interpretación de los resultados obtenidos.



3. Aportar posibles alternativas cuando los modelos lineales no proporcionen una solución adecuada.

■ Competencias:

1. Habilitación para hacer un uso científico de las numerosas facilidades computacionales y gráficas disponibles en la actualidad.
2. Predisposición para el estudio de generalizaciones y otras metodologías más modernas, que se alejan, en mayor o menor medida, de la rigidez impuesta por la estructura paramétrica que aquí se considera.

5.CONTENIDOS DE LA ASIGNATURA

Programa

0. Introducción al Entorno Computacional "R"
1. Distribución Normal Multivariante
2. Distribución de Formas Cuadráticas
3. Modelo de Regresión Lineal
4. Estimación y Contraste de Hipótesis
5. Validación y Diagnóstico
6. Regresión con Variables Predictoras Aleatorias
7. Modelos de Análisis de la Varianza
8. Análisis de la Varianza para Uno y Dos Tratamientos

Comentarios al programa

Como ya se ha mencionado en otro punto de esta guía, se supone que el alumno tiene conocimientos de teoría de la *Probabilidad*, hasta el punto de manejar con soltura conceptos básicos relativos a vectores aleatorios y su distribución de probabilidad, así como de *Álgebra Matricial*, sobre todo en lo relacionado con las matrices simétricas definidas positivas. Por otro lado, como ocurre en todas las asignaturas que tratan con la construcción de modelos, tanto la experimentación con los resultados teóricos aportados como la resolución de problemas con datos reales, pueden acelerar la comprensión de la materia. Para facilitar este aspecto práctico, el curso se inicia con una breve introducción a "R" (0), entorno informático orientado al análisis gráfico y numérico de datos, que puede descargarse de forma gratuita por la Red.

Con este bagaje comienza el desarrollo de la materia objeto de esta asignatura. El primer paso es el estudio de la distribución Normal multivariante (1), que desempeña un papel destacado tanto en la teoría de los modelos lineales como en la inferencia estadística multivariante. A continuación se hace un recorrido breve, seleccionando sólo los resultados necesarios para desarrollos posteriores, por el terreno de las formas cuadráticas (2) definidas sobre vectores aleatorios; en particular se estudian las propiedades asociadas al caso en el que el vector involucrado tiene una distribución de probabilidad *Normal*. En este punto se definen las versiones *no centrales* de algunas distribuciones conocidas: *Chi-cuadrado*, *F* y *t*.

Ahora ya se puede abordar el estudio de uno de los protagonistas de este curso: el modelo de regresión lineal (3), en el que la linealidad se utiliza para representar el efecto de un conjunto de variables cuantitativas -explicativas o predictoras, x 's, sobre la media de una variable respuesta unidimensional, y , con el mismo carácter que las anteriores. Los primeros apartados de interés, como siempre que se trabaja con modelos, son la estimación y el contraste de hipótesis (4) sobre los parámetros de la estructura probabilística impuesta. La validación del modelo y el diagnóstico (5) son otros aspectos importantes para poder hacer un uso riguroso de la técnica. El bloque de temas dedicados a la regresión lineal finaliza con una revisión de los puntos clave del problema cuando se supone que el carácter aleatorio también afecta a las



variables explicativas/predictoras y no sólo a la respuesta; lo denominamos regresión con x's aleatorias (6).

La última fase del curso se dedica a una variante del planteamiento previo: el objetivo sigue siendo una respuesta cuantitativa, pero ahora se trata de explicarla mediante variaciones no cuantificables; esta situación aparece muy ligada a los diseños experimentales en los que, con frecuencia notable, las variables explicativas son de tipo cualitativo. En la literatura especializada, los modelos que se adaptan a estas especificaciones se conocen como modelos de análisis de la varianza (7) y se identifican con las siglas ANOVA. En esta asignatura se analizan los métodos para uno y dos tratamientos aditivos (8) como muestra de la enorme casuística asociada con el problema.

6.EQUIPO DOCENTE

DATOS NO DISPONIBLES POR OBSOLESCENCIA

7.METODOLOGÍA

El alumno debe estudiar los temas que componen el programa a través del texto base propuesto para este fin. El equipo docente dirigirá y apoyará el estudio por medio de internet.

Para el aprendizaje de la materia contenida en el programa de esta asignatura, el alumno deberá combinar adecuadamente el estudio de la *teoría* con la *experimentación* y el *análisis de casos prácticos*. El plan que se expone a continuación tiene carácter orientativo y deberá reajustarse en función de las necesidades y hábitos personales. Una recomendación de carácter general es la de no perder nunca de vista el objetivo que hay tras los desarrollos matemáticos; esto se podría conseguir con un acercamiento a base de lecturas sucesivas, cada vez más detalladas.

Plan de Trabajo

Tema del Programa	Capítulo/s del Texto Base	Horas Teoría	Horas Práctica	Dedicación Total
0	-	2	2	4
1	4	10	3	13
2	5	8	2'5	10'5
3	6	4	1'5	5'5
4	7 y 8	36	14	50
5	9	6	2	8
6	10	12	3	15
7	11	16	4	20
8	12 y 13	28	8	36

Tutorización y utilización de recursos del curso virtual: 25'5 horas.

8.BIBLIOGRAFÍA BÁSICA

ISBN(13): 9780471754985
Título: LINEAR MODELS IN STATISTICS (2ª ed.)
Autor/es: Alvin C. Rencher & G. Bruce Schaalje ;
Editorial: WILEY- INTERSCIENCE

Buscarlo en librería virtual UNED

Buscarlo en bibliotecas UNED



Buscarlo en la Biblioteca de Educación

Buscarlo en Catálogo del Patrimonio Bibliográfico

Comentarios y anexos:

- Aunque se tomará como referencia esta segunda edición, la asignatura se puede estudiar completamente con la primera edición del texto.

9.BIBLIOGRAFÍA COMPLEMENTARIA

Comentarios y anexos:

General

- Arnold, S.F. (1981) *The Theory of Linear Models and Multivariate Analysis*, Wiley.
- Carmona, F. (2005) *Modelos Lineales*, Publicacions i Edicions UB.
- Christensen, R. (1987) *Plane Answers to Complex Questions. The Theory of Linear Models*, Springer-Verlag.
- Graybill, F.A. (1996) *Theory and Application of the Linear Model*, Wadsworth.
- Montgomery, D.C., Peck, E.A. and Vining, G.G. (2002) *Introducción al Análisis de Regresión Lineal*, CECSA.
- Rao, C.R. and Toutenburg, H. (1999) *Linear Models: Least Squares and Alternatives*, Springer-Verlag.
- Ravishanker, N. and Dey, D.K. (2002) *A First Course in Linear Model Theory*, Chapman & Hall.
- Sengupta, D. (2003) *Linear Models. An Integrated Approach*, World Scientific.

Preliminares

- Graybill, F.A. (1997) *Introduction to Matrices with Applications in Statistics*, Wadsworth.
- Harville, D. (1997) *Matrix Algebra from a Statistician's Perspective*, Springer-Verlag.
- Krzanowski, W. (1998) *An Introduction to Statistical Modelling*, Edward Arnold.

Aspectos computacionales y Aplicaciones

- Faraway, J.J. (2005) *Linear Models with R*, Chapman & Hall.
- Fox, J. (2002). *An R and S-Plus Companion to Applied Regression*, Sage Pub.
- Venables, W.N. & Ripley B.D. (2002) *Modern Applied Statistics with S, 4th edition*, Springer-Verlag.
- Weisberg, S. (2005) *Applied Linear Regression, 3rd ed.*, Wiley.

Generalizaciones y otras Extensiones

- Dobson, A.J. (2002) *An Introduction to Generalized Linear Models, 2nd edition*, Chapman & Hall.
- Faraway, J.J. (2006) *Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric*



Regression Model, Chapman & Hall.

McCullagh, P. and Nelder, J.A. (1989) *Generalized Linear Models*, 2nd edition, Chapman & Hall.

10. RECURSOS DE APOYO AL ESTUDIO

Internet es una fuente de materiales multimedia que pueden ser útiles para el aprendizaje de esta materia. Dado el carácter dinámico de esta información, se comunicará al alumno al comienzo del curso.

11. TUTORIZACIÓN Y SEGUIMIENTO

Esta fase del proceso se ejecutará con los medios actualmente disponibles en esta Universidad. Por un lado, se utilizará una plataforma de virtualización desde la cual se responderá a las cuestiones planteadas por los alumnos y, por otro, se habilitará un horario para consultas telefónicas y presenciales, que se comunicará a los alumnos oportunamente.

12. EVALUACIÓN DE LOS APRENDIZAJES

Se realizará una Prueba de Evaluación Presencial que tendrá lugar en los Centros Asociados. Dicha prueba consistirá en la resolución de una serie de ejercicios mediante los que se valorará el grado de comprensión de los conceptos y de sus relaciones. La duración prevista es de dos horas.

13. COLABORADORES DOCENTES

Véase equipo docente.

