
Profesor: Carlos Felipe Valencia
cf.valencia@uniandes.edu.co
Of. ML 718

Horario de Clase: Martes y Jueves de 5:00 a 6:20 pm
Salón: B-401

A. DESCRIPCIÓN GENERAL Y OBJETIVOS DEL CURSO

En este curso se presentan los conceptos fundamentales de los modelos estadísticos lineales, tanto teóricos como aplicados. Los modelos lineales, son la base de la mayoría de modelos estadísticos usados actualmente y contienen las características esenciales para comprender los mecanismos de estimación e inferencia estadística. Son usados extensivamente en diversas ramas como ingeniería, economía, ciencias naturales, ciencias sociales y administración de negocios.

La idea fundamental de los modelos desarrollados en el curso es construir una relación funcional lineal entre variables explicativas y una variable de interés que se quiere entender o predecir. Los mecanismos de estimación permiten comprobar estadísticamente las relaciones más importantes, así como desarrollar pruebas que permiten comprobar o proponer teorías que representen la realidad de los fenómenos que se registran con los datos.

El curso pretende que el estudiante adquiera todas las habilidades técnicas para la correcta aplicación de los modelos, pero también que desarrolle un pensamiento crítico que le permita evaluar las ventajas y desventaja de los mismos.

Los objetivos primarios de curso son:

- Desarrollar la capacidad de formular modelos estadísticos apropiados para describir fenómenos aleatorios
- Comprender los conceptos fundamentales de los modelos estadísticos lineales, incluyendo sus supuestos y limitaciones
- Aprender a utilizar herramientas computacionales que permitan la correcta aplicación de los métodos vistos
- Desarrollar habilidades para el análisis, comprensión y comunicación de resultados de los modelos.

B. CONTENIDO

PARTE A. MODELO DE REGRESIÓN LINEAL

I. Introducción y conceptos generales de inferencia estadística y uso de los modelos lineales

- Conceptos generales de inferencia estadística y formulación del modelo lineal
- Relaciones funcionales lineales en modelos estadísticos y causalidad
- Estimación en el modelo de regresión lineal múltiple
- Notación matricial e interpretación vectorial del modelo

II. Inferencia, pruebas de hipótesis y predicciones

- Inferencia de parámetros por intervalos de confianza y pruebas de hipótesis
- Predicción de nuevas observaciones
- Análisis de Varianza y descomposición en formas cuadráticas
- Pruebas conjuntas e inferencia con restricciones

III. Selección de modelos

- Descomposición sesgo-varianza
- Criterios de selección de modelos
- Selección de variables vs. Pruebas de hipótesis
- Selección con regularización

IV. Diagnóstico y validación de modelos

- Multicolinealidad
- Varianza no-constante
- *Outliers* y valores influyentes

V. Formas generales

- Variables categóricas y diseño de experimentos
- Regresión polinomial
- Modelo lineal general y mínimos cuadrados ponderados

PARTE B. MODELO LINEAL GENERALIZADO

I. Introducción conceptos fundamentales

- Conceptos básicos del modelo generalizado en familias exponenciales
- Pruebas asintóticas e inferencia en GLM
- Sobre-dispersión y validación de los modelos generalizados

II. Modelo de regresión logística

- Planteamiento, inferencia y validación

III. Modelo de regresión de Poisson

- Planteamiento, inferencia y validación

PARTE C. MODELO LINEAL EN DATOS PANEL

I. Introducción conceptos fundamentales de la regresión en datos panel

II. Estimación con efectos fijos

III. Modelo de componentes del error (efectos aleatorios)

C. BIBLIOGRAFÍA

- **Kutner, M.; Nachtsheim, C.; Neter, J.; Li, W.** (2005) Applied Linear Statistical Models. Fifth edition. McGraw Hill.
- **Seber, G.; Lee, A.** (2003) Linear Regression Analysis. Second edition. Wiley.
- **Montgomery, D.; Peck, E.; Vinning, G.** (2012) Introduction to Linear regression Analysis. Fifth edition. Wiley.
- **Myers, R.; Montgomery, D.; Vinning, G.; Robinson, T.** (2010) Generalized linear models: with applications in engineering and sciences. Second edition, Wiley.
- **McCullagh, P.; Nelder, A.** (1989) Generalized Linear Models. Second edition. Chapman & Hall/CRC.
- **Mátyás, L.; Sevestre, P.** (1996) The Econometrics of Panel Data. Second edition. Springer.

D. EVALUACIÓN

- **Examen I** (25%)
- **Examen II** (25%)
- **Talleres y evaluaciones en clase** (15%)
- **Proyecto Final** (10%)
- **Tareas** (25%)

E. ASPECTOS GENERALES

- **Software y uso de herramientas computacionales.** Durante el curso se hará énfasis en el aspecto práctico de los modelos y herramientas vistas en clase, por lo cual, es indispensable el manejo de software estadístico que permita solucionar problemas basados en datos reales. El paquete estadístico que se utilizará mayormente será **R**, el cual se puede obtener de manera gratuita desde <http://cran.r-project.org>. Otros softwares que pueden ser utilizados en menor medida son SAS y módulos de Phyton. El estudiante es libre de utilizar el software de su preferencia que le permita solucionar los problemas prácticos.
- **Metodología.** El contenido de los temas será presentado por el profesor durante las clases magistrales, además de casos de aplicación y uso de software presentados por el asistente del curso. Se espera que el estudiante realice un trabajo individual de manera semanal para afianzar los conceptos y obtener las habilidades requeridas. Solución de dudas o asesorías extras por parte del profesor pueden ser programadas por correo electrónico o durante la clase.
- **Conducta y Código de Honor.** Se espera que todos los estudiantes tengan un comportamiento respetuoso frente a sus compañeros de clase y a sus profesores. Así mismo, se espera completa honestidad en las evaluaciones, de forma tal que estas reflejen el trabajo individual y el conocimiento que se tiene sobre los temas del curso. Cualquier violación al reglamento general de estudiantes será notificada y tratada mediante el conducto regular.
- **Prerrequisitos.** Conocimientos básicos de probabilidad, estadística y álgebra lineal.