



XS-2130: MODELOS DE REGRESIÓN APLICADOS – GRUPO 01
PROGRAMA
II SEMESTRE 2021

Docente: Ricardo Alvarado Barrantes

Correo electrónico: estad.ucr@gmail.com

Oficina: 17 Estadística.

Teléfono: 84021263

Horario de clases: L y J: 9:00-10:10

Horas de consulta: M: 11 am a 1 pm

Enlace Zoom:

<https://udecr.zoom.us/j/82154992116>

1. Descripción

Los modelos de regresión son un tema central en la práctica de la estadística y forman la base de un amplio rango de métodos estadísticos. En este curso se presentarán las bases de los modelos de regresión múltiple y los procedimientos desarrollados para asegurar una aplicación correcta de los mismos.

Además de los conocimientos teóricos se brindará al estudiante la posibilidad de aplicar los métodos mediante el uso del lenguaje de programación R.

- **Requisitos:** XS-2110 Métodos Estadísticos, XS-2210 Estadística Computacional I y MA-0313.
- **Correquisitos:** XS-2330 Estadística Computacional II.
- **Horas:** 5 horas semanales (2 de teoría y 3 de práctica).
- **Créditos:** 4.

2. Objetivo General

Proveer a los estudiantes el conocimiento conceptual y los instrumentos estadísticos para la aplicación de los modelos de regresión lineal.

3. Objetivos Específicos

Al finalizar el curso el estudiante tendrá criterio y conocimiento básico para:

- Reconocer las situaciones donde se pueden aplicar las técnicas de regresión lineal.
- Utilizar un modelo con fines de predicción (un valor estimado individual, un valor estimado promedio).
- Usar diferentes técnicas de selección de variables para formar un modelo apropiado a partir de un conjunto de predictores disponibles.
- Llevar a cabo los diagnósticos del modelo de regresión ajustado.
- Aplicar las medidas apropiadas en los casos de no conformidad con los supuestos del modelo de regresión.





- Analizar casos con respuestas de dos categorías y conteos.

4. Contenidos

| | |
|---|---|
| I. El modelo de regresión lineal. | |
| 1.1 | El modelo y los supuestos. |
| 1.2 | Estimación de los coeficientes (notación matricial). |
| 1.3 | Coefficientes estandarizados. |
| 1.4 | Valores ajustados y residuales. |
| 1.5 | Inferencias. El cuadrado de medio de error. Varianzas de los estimadores. Intervalos de confianza para los coeficientes. |
| 1.6 | Inferencias para la respuesta media y valores individuales. |
| 1.7 | Coefficiente de determinación. |
| 1.8 | Comparación de modelos anidados con la distribución F. Uso de la distribución t. |
| 1.9 | Predictores categóricos. Modelos con interacciones. |
| II. Selección de predictores. | |
| 2.1 | Procedimiento con todos los posibles subconjuntos: R^2 ajustado, suma de cuadrados de predicción (PRESS) y P^2 . |
| 2.2 | Selección por pasos: criterios de información (Akaike y Bayes). |
| III. Diagnósticos. | |
| 3.1 | Verificación de los supuestos: a. Normalidad: QQ plot. Pruebas de Kolmogorov-Smirnov, Jarque Bera, Shapiro. b. Homoscedasticidad: prueba de Breush Pagan. c. No multicolinealidad: factor de inflación de la variancia (VIF). d. Linealidad entre predictor y respuesta: gráficos de regresión parcial. |
| 3.2 | Medidas para detectar valores extremos: residuales estudentizados, "leverage". |
| 3.3 | Medidas para detectar casos de influencia: DFFITs, distancia de Cook, DFBETAs. |
| IV. Construcción del modelo. | |
| 4.1 | Transformaciones de Box-Cox. |
| 4.2 | Mínimos cuadrados ponderados. |
| 4.3 | Regresión poynomial. |
| 4.4 | Evaluación de la precisión en casos no estándar: bootstrap. |
| 4.5 | Validación del modelo. |
| V. Modelos lineales generalizados. | |
| 5.1 | Regresión logística. |





5. Metodología

- El curso seguirá la modalidad virtual usando la plataforma Zoom para las lecciones, las cuales serán grabadas.
- Presentaciones teóricas: se impartirán lecciones sincrónicas por parte del docente donde se explicarán los conceptos y sus aplicaciones. Las lecciones serán grabadas y estarán disponibles para que los estudiantes las puedan descargar.
- Prácticas: se realizarán laboratorios estructurados con ejercicios sobre los contenidos desarrollados en las clases teóricas. Estos laboratorios también se grabarán y la solución también estará disponible. Durante las sesiones de laboratorio se utilizará el lenguaje de programación R para realizar ejercicios de la materia vista en clase.
- Juegos: Los estudiantes tendrán la oportunidad de realizar varios juegos didácticos para el aprendizaje y/o evaluación de algunos temas.
- Tareas: se asignarán ejercicios de práctica que incluirán aplicaciones con datos para ser analizados con R, así como interpretaciones de los resultados
- Trabajo creativo: Los estudiantes realizarán un trabajo creativo donde se ilustren los conocimientos aprendidos durante el curso.
- Trabajo con datos reales:
 - Los estudiantes realizarán un trabajo de análisis de datos reales utilizando lo aprendido en el curso. El trabajo será en grupos de 3 (máximo).
 - Las aplicaciones podrán ser tomadas de investigaciones en áreas como la economía, medicina, psicología, biología, población, etc.
 - Los estudiantes presentarán un anteproyecto donde expondrán los objetivos del estudio, el cual será criticado por el profesor y devuelto para que se realicen las mejoras sugeridas.

6. Evaluación

| Exámenes | 70 % |
|-------------------------------------|-------------|
| Primer examen parcial | 25% |
| Segundo examen parcial | 25% |
| Tercer examen parcial | 25% |
| Trabajos | 25 % |
| Trabajo de análisis de datos reales | 15% |
| Trabajo creativo | 10% |





7. Cronograma

| | CAPITULO | L | J | Actividad |
|-----------|-----------------------------------|----|---|---|
| AGOSTO | I. Modelo de regresión lineal | 16 | | Presentación |
| | | 19 | | El modelo |
| | | 23 | | Lab.0 - DESCRIPTIVO |
| | | 26 | | Estimación de coeficientes |
| SETIEMBRE | I. Modelo de regresión lineal | 30 | | Coefficientes estandarizados |
| | | 2 | | Lab. No.1 - ESTIMACION DE COEFICIENTES Y COEFICIENTES ESTANDARIZADOS |
| | | 6 | | Valores ajustados y residuales / Inferencias sobre los betas - SIMULACION |
| | | 9 | | Lab. No.2 - INTERVALOS DE CONFIANZA PARA LOS BETAS |
| | | 13 | | FERIADO |
| | | 16 | | Lab. No.3 - INTERVALOS PARA LA MEDIA Y RESPUESTA INDIVIDUAL |
| | | 20 | | Coefficiente de determinación |
| | | 23 | | Comparación de modelos anidados |
| | | 27 | | Lab. No.4 - COMPARACION DE MODELOS ANIDADOS |
| | | 30 | | EXAMEN I |
| OCTUBRE | II. Selección de Predictores | 4 | | Lab. No.5 - PREDICTORES CATEGORICOS |
| | | 7 | | Lab. No.5 - PREDICTORES CATEGORICOS (continuación) |
| | | 11 | | RECESO DE VIRTUALIDAD - Selección de variables |
| | | 14 | | TRABAJO EN PROYECTO CREATIVO |
| | | 18 | | Lab. No.6 - SELECCION DE VARIABLES |
| OCTUBRE | III. Diagnósticos | 21 | | Los supuestos del modelo / Lab. No.7 - DIAGNOSTICOS I |
| | | 25 | | Lab. No.7 - DIAGNOSTICOS I (continuación) |
| | | 28 | | Lab. No.8 - VALORES EXTREMOS Y CASOS DE INFLUENCIA |
| NOVIEMBRE | IV. Construcción del modelo | 1 | | Lab. No.9 - TRANSFORMACIONES |
| | | 4 | | Lab. No.10 - MINIMOS CUADRADOS PONDERADOS |
| | | 8 | | Lab. No.11 - REGRESION POLINOMIAL |
| | | 11 | | EXAMEN ii |
| | | 15 | | Lab. No. 12 - BOOSTRAP Y VALIDACION |
| NOVIEMBRE | V. Modelos lineales generalizados | 18 | | Regresión logística |
| | | 22 | | Lab. No.13 - REGRESION LOGISTICA |
| | | 25 | | Presentación de trabajos |
| DIC | | 29 | | FERIADO |
| | | 2 | | EXAMEN III |
| | | 6 | | |
| | | 9 | | |
| | | 13 | | AMPLIACION |



8. Referencias bibliográficas

Chatterjee, S. & A.S. Hadi (1988). *Sensitivity Análisis in Linear Regression*. New York, N.Y.: Wiley.

BIBLIOTECA LUIS DEMETRIO TINOCO 519.536 Ch495s c.1

Faraway, J. (2002). *Practical Regression and Anova using R*. Documento en internet: <http://www.biostat.jhsph.edu/~iruczins/teaching/jf/faraway.html>

Faraway, J. (2004). *Linear Models with R*. Chapman & Hall/CRC.

BIBLIOTECA LUIS DEMETRIO TINOCO 519.536 F219L

Faraway, J. (2005). *Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models*. Chapman & Hall/CRC.

BIBLIOTECA LUIS DEMETRIO TINOCO 519.536 F219e

Fox, J. (2011). *An R ans S-Plus Companion to Applied Regression*. 2a edición. Thousand Oaks, Cal.: Sage Publications, Inc.

BIBLIOTECA LUIS DEMETRIO TINOCO 519.536.028.551.33 F792a2

Gujarati, D. (1990). *Econometría básica*. McGraw-Hill.

BIBLIOTECA LUIS DEMETRIO TINOCO 330.015.195 G969e2

Kutner, M, Nachtsheim, C, Neter, J, Li, William (2005). *Applied Linear Statistical Models*. 5a. ed. WCB, McGraw-Hill. (libro de texto)

BIBLIOTECA LUIS DEMETRIO TINOCO 519.536 K97a5