



XS-2130: MODELOS DE REGRESIÓN APLICADOS – GRUPO 01

PROGRAMA

II SEMESTRE 2022

Docente: Ricardo Alvarado Barrantes

Correo electrónico: estad.ucr@gmail.com

Oficina: 17 Estadística.

Teléfono: 84021263

Horario de clases:

L: 9:00-11:50 (lab 108 CE)

J: 9:00-10:50 (aula 211 AU)

Horas de consulta:

M: 1:00 a 3:00 pm por Zoom

<https://udecr.zoom.us/j/85015834757>

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. No autocorrelación: prueba de Durbin-Watson. b. Normalidad: QQ plot. Pruebas de Kolmogorov-Smirnov, Jarque Bera, Shapiro. c. Homocedasticidad: prueba de Breush-Pagan. d. No multicolinealidad: factor de inflación de la variancia (VIF). e. 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 polinomial.
4.4	Regresión robusta.
4.5	Componentes principales.
4.6	Evaluación de la precisión en casos no estándar: bootstrap.
V. Modelos lineales generalizados.	
5.1	Regresión logística.



5. Metodología

- Presentaciones teóricas: se impartirán lecciones sincrónicas por parte del docente donde se explicarán los conceptos y sus aplicaciones.
- Laboratorios: se realizarán laboratorios estructurados con ejercicios sobre los contenidos desarrollados en las clases teóricas. Durante las sesiones de laboratorio se utilizará el lenguaje de programación R para realizar ejercicios de la materia vista en clase.
- Ejercicios: se asignarán ejercicios para asegurar la “puesta en práctica” de los conceptos estudiados. Los ejercicios incluirán aplicaciones con datos para ser analizados con R, así como interpretaciones de los resultados
- Trabajos creativos: Los estudiantes realizarán trabajos creativos 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	65 %
Primer examen parcial	20%
Segundo examen parcial	20%
Tercer examen parcial	25%
Trabajos	35 %
Trabajo de análisis de datos reales	20%
Trabajos creativos	10%
Revisión bibliográfica	5%



7. Cronograma

	CAPITULO	L	J	Actividad	
AGOSTO		15		FERIADO	
			18	Presentación	
	I. Modelo de regresión lineal	22			
		29			
SEPTIEMBRE		1			
		5			
		8			
		12			
		15			
		19		FERIADO	
		22			
		26		EXAMEN I	
		29			
OCTUBRE		3		ANTEPROYECTO	
		6			
	II. Selección de predictores	10			
		13			
	III. Diagnósticos	17			
		20			
		24		EXAMEN II	
		27			
		31			
NOVIEMBRE	IV. Construcción del modelo	3		REVISION BIBLIOGRAFICA	
		7			
		10			
		14			
		17			
		21			
	V. Regresión logística	21			
		24			
		28		ARTICULO	
DIC			1		
		5		EXAMEN III	
			8		



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