



CARTA AL ESTUDIANTE
Modelos de Regresión Aplicados
XS2130
II Semestre 2024

	GRUPO 01	GRUPO 02
Docente:	Ricardo Alvarado Barrantes	Alejandra Arias Salazar
Correo:	estad.ucr@gmail.com	alejandra.ariassalazar@ucr.ac.cr
Teléfono:	84021263	
Clases:	L: 10:00-11:50 am (241 CE) J: 10:00-11:50 am (Lab 107 AU)	L: 10:00-11:50 am (Lab 015 CE) J: 10:00-11:50 am (240 CE)
Consulta:	M: 1:00-3:00 pm	L-J: 9:00-10:00 am
Zoom:	https://udecr.zoom.us/j/86862167699	https://udecr.zoom.us/j/86911965318?pwd=fowehlmmguZZfl3ReAhqRgEqppWyt.1
Materiales:	https://6f33fa7f78ea46e2aaca-my.sharepoint.com/:f:/g/personal/ricardo_alvarado_ucr_ac_cr/EgwdeeTVVRtlrE5tX4L2Px4B8kdZJL4ICxXf2_e6jYc4Yg?e=UEnYtZ	

Características del curso

Horas:	2 teoría y 2 práctica
Créditos:	4
Requisitos:	XS0122 Modelos Probabilísticos I (Equiv. XS2310), MA0252 Álgebra Lineal (Equiv. MA1004), XS0129 Programación para Estadística I (Equiv. XS2210), XS3150 Diseño de Experimentos (Equiv. XS2110).
Correquisitos:	ninguno
Ciclo:	IV
Clasificación:	propio





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 presentan las bases de los modelos de regresión lineal y los procedimientos necesarios para asegurar su correcta aplicación, con el fin de seleccionar el modelo óptimo. Se estudian las técnicas para determinar si un conjunto de datos es apto para la aplicación de un modelo lineal y se proponen alternativas ante situaciones particulares. Además de adquirir conocimientos teóricos, cada estudiante debe aplicar las técnicas, utilizando lenguajes de programación estadística, y también debe realizar un trabajo de investigación que se lleva a cabo con una situación real, al lado de una persona profesional experta del área de aplicación.

Objetivo general

Aplicar los fundamentos y características del modelo de regresión lineal múltiple para ajustarlo a situaciones que se presenten en el contexto empírico.

Objetivos específicos

Al finalizar el curso el/la estudiante estará en la capacidad de:

1. Reconocer las situaciones en las que se puede estimar un modelo de regresión lineal para contribuir con los objetivos de una investigación.
2. Evaluar las técnicas de selección de variables más apropiadas para estimar un modelo lineal parsimonioso con el conjunto de predictores disponibles.
3. Seleccionar las técnicas más apropiadas para evaluar el cumplimiento de los supuestos de un modelo de regresión lineal.
4. Aplicar las técnicas que permitan estimar una ecuación que esté acorde con los supuestos del modelo de regresión gaussiano, para generar interpretaciones correctas y predicciones apropiadas.



Habilidades y conocimientos (perfil de salida)

Habilidades	Conocimientos
HM01 - Manejar entidades matemáticas asociadas a la estadística	CM02 - Conocimientos intermedios en distribuciones estadísticas y de probabilidad
HM02 - Emplear lenguaje matemático para expresar propiedades estadísticas	CM04 - Conocimientos avanzados de aspectos teórico-matemáticos que dan fundamento al uso de las técnicas de análisis estadístico
HE01 - Identificar y aplicar modelos estadísticos apropiados según el problema de investigación	CE01 - Conocimientos avanzados de técnicas clásicas y modernas de análisis de datos univariados y multivariados para comprender los fenómenos en diferentes áreas del conocimiento CE02 - Aplicación de modelos estadísticos a problemas de diversas áreas del conocimiento
HE05 - Simular diversos fenómenos mediante modelos estadísticos usando escenarios asociados a condiciones experimentales u observacionales	CE09 - Conocimientos básicos en desarrollo de procesos de simulación de escenarios estadísticos
HT02 - Crear, interpretar y modificar programas de código escrito en lenguajes de programación (ej: R, Python y SQL)	CT08 - Conocimientos intermedios en lenguajes de programación para análisis estadístico (ej: R, Python)
HI01 - Identificar y aplicar metodologías y diseños de investigación adecuados	CI01 - Conocimientos avanzados en los diferentes tipos de diseños de estudios experimentales, cuasiexperimentales, por muestreo y otros observacionales CI02 - Conocimientos intermedios de metodologías de investigación
HI05 - Contextualizar el problema de investigación y los resultados al campo de aplicación HI06 - Aprender de forma autónoma elementos del campo de aplicación HI07 - Comprender artículos científicos tanto de estadística como de disciplinas sustantivas a las cuales aplica la estadística	CI07 - Conocimiento de estrategias de aprendizaje autodidacta CI08 - Conocimientos intermedios de técnicas de lectura (en español e inglés)
HI09 - Identificar fuentes bibliográficas confiables	CI12 - Conocimientos intermedios de citación CI13 - Conocimientos básicos de búsqueda de fuentes bibliográficas (SIBDI)
HC01 - Expresar y transmitir conocimientos técnicos mediante lenguaje adaptado al público meta HC02 - Comunicarse con profesionales de otros campos para entender sus necesidades	CC01 - Conocimientos básicos en técnicas para la comunicación oral efectiva



HC03 - Comunicar conceptos técnicos en la escritura formal de un documento académico	CC02 - Conocimientos avanzados de escritura matemática CC03 - Conocimientos avanzados en los aspectos teóricos de las técnicas y modelos estadísticos
HC04 - Comunicar asertivamente mediante la manipulación creativa de datos pero manteniendo su integridad	CC04 - Conocimientos avanzados en construcción de cuadros y gráficos
HC05 - Comunicar adecuadamente de forma escrita diferentes elementos del quehacer estadístico	CC06 - Conocimientos intermedios de técnicas de escritura de artículos e informes técnicos CC07 - Conocimientos avanzados de redacción y ortografía

Contenidos

1. Modelo de regresión lineal:
 - a) Tipos de estudios. Usos del análisis de regresión.
 - b) El modelo y los supuestos.
 - c) Valores ajustados y residuales.
 - d) Estimación de los coeficientes: mínimos cuadrados ordinarios y máxima verosimilitud (concepto de verosimilitud).
 - e) Interpretación de coeficientes: coeficientes estandarizados.
 - f) Inferencias sobre los coeficientes: cuadrado de medio residual, distribución de los estimadores (concepto del error estándar), intervalos de confianza para los coeficientes.
 - g) Intervalos de confianza para la respuesta media y predicción de nuevas observaciones.
 - h) Coeficiente de determinación R^2 .
 - i) Comparación de modelos anidados: distribuciones F y t.
 - j) Predictores cualitativos: modelos con interacciones.
 - k) Variables confusoras, mediadoras y de control.
2. Selección de predictores:
 - a) Procedimiento con todos los posibles subconjuntos: R^2 ajustado, suma de cuadrados de predicción (PRESS) y P^2 .
 - b) Selección por pasos: concepto de devianza, criterios de información (Akaike y Bayes) y criterio de la probabilidad asociada.
 - c) Gráficos de regresión parcial y su relación con la correlación entre predictores.





3. Diagnósticos:

a) Verificación de los supuestos:

- Normalidad: análisis gráfico y pruebas formales.
- Homocedasticidad: gráfico de residuales versus ajustados; prueba de Breusch-Pagan.
- No multicolinealidad: factor de inflación de la varianza.
- Linealidad entre predictor y respuesta promedio: gráficos de dispersión y gráficos de residuales parciales.

b) Medidas para detectar valores extremos.

c) Medidas para detectar casos de influencia.

4. Construcción del modelo:

a) Transformaciones en la variable dependiente (por ejemplo, Box-Cox) y en las predictoras (por ejemplo, regresión polinomial).

b) Mínimos cuadrados ponderados y regresión robusta.

c) Otras técnicas: Bootstrap, LASSO, regresión local.



Metodología

El curso es teórico-práctico y exige el uso frecuente de la computadora. Se espera que el estudiante aprenda los fundamentos teóricos de los modelos de regresión y que aplique las técnicas a archivos de datos utilizando lenguajes de programación estadística. Se propone una combinación de actividades, tales como:

1. Presentaciones teóricas: lecciones por parte del docente donde se explican los conceptos y sus aplicaciones.
2. Laboratorios: sesiones estructuradas con ejercicios sobre los contenidos desarrollados en las clases teóricas con solución disponible. Durante las sesiones de laboratorio se utilizan lenguajes de programación (por ejemplo, R).
3. Prácticas: ejercicios fuera de clase que incluyen aplicaciones con datos para ser analizados, así como interpretaciones de los resultados.
4. Revisión bibliográfica: los estudiantes deben seleccionar un artículo donde se aplique un modelo de regresión, exponerlo y entregar un reporte sobre el artículo que incluya una crítica sobre el mismo.
5. Desarrollo de investigación: incluye contactar a una persona experta en algún campo de interés, plantear objetivos y diseño, ejecución, escritura de un artículo científico y presentación oral de resultados de una manera comprensible para el público meta.
6. Estudios de simulación: en este trabajo se plantea un objetivo metodológico de la estadística relacionado con supuestos u otros temas que puedan estudiarse mediante el planteamiento de escenarios simulados.

Evaluación

1. 3 exámenes parciales (65% del total: 20% parcial I y II, 25% III Parcial).
2. Trabajo de investigación (20%).
3. Trabajo de simulación (10%).
4. Presentación de revisión bibliográfica (5%).



Cronograma

SEMANA / FECHA	CONTENIDO	DETALLE
1: 12-15.08*	Modelo de regresión lineal	Sesión sobre elaboración de artículos
2: 19-22.08		
3: 26-29.08* (jornadas)		
4: 02-05.09		
5: 09-12.09		
6: 16-19.09		
7: 23-26.09	Selección de predictores	Primer examen (sábado 28/9)
8: 30.09-03.10		Entrega de anteproyecto
9: 07-10.10	Diagnósticos	Segundo examen (sábado 26/10)
10: 14-17.10		
11: 21-24.10		
12: 28-31.10	Construcción del modelo	Revisión bibliográfica
13: 04-07.11		Entrega de trabajo de simulación
14: 11-14.11		Entrega trabajos finales
15: 18-21.11		
16: 25-28.11		
17: 02.12		Tercer examen (lunes 2/12)
18: 12.12		Ampliación



Bibliografía

- Fox, J., & Weisberg, S. (2018). *An R companion to applied regression*. Sage publications.
SIBDI: 519.536.028.551.33 F792a2
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). Statistical learning. In *An introduction to statistical learning*. Springer, New York, NY.
- Kutner, M, Nachtsheim, C, Neter, J, Li, William (2005). *Applied Linear Statistical Models*. 5a. ed. WCB, McGraw.
SIBDI: 519.536 K97a5
- Madhyastha, N. M., Ravi, S., & Praveena, A. S. (2020). *A First Course in Linear Models and Design of Experiments*. Singapore: Springer.
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2021). *Introduction to linear regression analysis*. John Wiley & Sons
- Rincón, L. (2009). *Curso básico de modelos lineales*. Universidad de Santo Tomas.

Reglamentación

- La reglamentación sobre sus deberes y derechos como estudiante se encuentra en el **Reglamento de Régimen Académico Estudiantil**
https://www.cu.ucr.ac.cr/normativ/regimen_academico_estudiantil.pdf
- La reglamentación y sanciones ante fraudes en las evaluaciones o comportamientos anómalos por parte de los y las estudiantes, la pueden encontrar en **Reglamento de Orden y Disciplina de los Estudiantes de la Universidad de Costa Rica**
https://www.cu.ucr.ac.cr/normativ/orden_y_disciplina.pdf
- Con el fin de garantizar un espacio libre de violencia y sexismo en el desarrollo de este curso, les recomiendo que revisen **el Reglamento de la Universidad de Costa Rica contra el Hostigamiento Sexual**
https://www.cu.ucr.ac.cr/normativ/hostigamiento_sexual.pdf