

UNIDAD: IZTAPALAPA		DIVISIÓN CIENCIAS BÁSICAS E INGENIERÍA	
NIVEL: LICENCIATURA		EN MATEMÁTICAS	
CLAVE: 213112	UNIDAD DE ENSEÑANZA - APRENDIZAJE: REGRESIÓN		TRIM: X
HORAS TEORÍA: 3	SERIACIÓN 2131149		CRÉDITOS: 9
HORAS PRÁCTICA: 3			OPT/OBL: OPT.

OBJETIVOS

GENERALES:

Al finalizar el curso el alumno será capaz de:

- Reconocer si un problema se puede resolver usando regresión, identificando las variables tanto respuesta como explicativas.
- Proponer el modelo de regresión que conviene usar, hacer el análisis e interpretar los resultados.
- Sugerir la forma de realizar la muestra y obtener los datos, hacer el análisis y utilizar los resultados para obtener el mejor modelo, interpretando los resultados de acuerdo con el problema.
- Expresar en forma oral y escrita los procedimientos y algoritmos utilizados así como sus conclusiones.

ESPECÍFICOS

Al finalizar el curso el alumno será capaz de:

- Usar correctamente los conceptos básicos de regresión lineal múltiple, curvilínea y logística.
- Usar la notación matricial para expresar el modelo de regresión y obtener los estimadores de mínimos cuadrados en regresión lineal múltiple.
- Identificar las sumas de cuadrados: total, del error y de regresión como los numeradores de las estimaciones de las varianzas de la variable respuesta y de los residuales de la regresión.
- Mostrar que la estimación de la varianza del error es insesgada y que la de regresión sólo es insesgada cuando todos los coeficientes de regresión de las variables explicativas son iguales a cero.
- Utilizar de manera adecuada los estadísticos de las pruebas parciales (t) y de regresión (F) y conocer las ecuaciones para obtener la significación muestral o valor p.
- Identificar los puntos atípicos usando diagramas de dispersión y la gráfica de probabilidad normal.
- Revisar los residuales estudentizados y los estadísticos diagonal de matriz de proyección y D de Cook para identificar puntos atípicos de influencia y decidir cómo tratarlos en la regresión.
- Identificar los casos en que ocurre multicolinealidad .
- Decidir con base en los resultados de las pruebas de hipótesis parciales y los valores de R cuadrada sobre la eliminación de variables explicativas para proponer el modelo más adecuado para el problema y los datos.
- Obtener e interpretar los intervalos de confianza para la media de la variable respuesta en valores específicos de las variables explicativas, y hacer lo mismo para las predicciones.
- identificar la forma de la respuesta y proponer un modelo curvilíneo adecuado a un problema específico, usando conocimientos propios del problema y diagramas de dispersión.

NOMBRE DEL PLAN LICENCIATURA EN MATEMÁTICAS		2/4
CLAVE 2131112	UNIDAD DE DE ENSEÑANZA-APRENDIZAJE REGRESIÓN	

- Obtener el óptimo (máximo o mínimo) en una superficie de respuesta cuadrática en dos variables.
- Reconocer los problemas donde se tiene una distribución normal multivariada, y obtener e interpretar la correlación de Pearson entre parejas de variables.
- Usar los residuales de las regresiones de Y y W sobre X para obtener las correlación parcial de Y y W dada X. Generalizar al caso de la correlación parcial dadas X_1, X_2, \dots, X_k usando regresiones sobre X_1, X_2, \dots, X_k .
- Usar la función de verosimilitud de la muestra Bernoulli y la Logit como función de enlace con una regresión lineal múltiple, para obtener el modelo de regresión logística Bernoulli.
- Usar resultados de la familia exponencial para conocer que los estimadores de los parámetros de interés son suficientes y permiten construir sus intervalos de confianza.
- Utilizar la devianza para probar hipótesis sobre los coeficientes de regresión.
- Extender la forma de plantear el modelo logístico para una Bernoulli al caso de una distribución multinomial con k clases.
- Obtener las estimaciones y realizar pruebas de hipótesis usando la devianza, con el modelo logístico para una multinomial.

CONTENIDO SINTÉTICO

1. Regresión lineal múltiple. (4 semanas)

- 1.1. Introducción: Generalizar el modelo de regresión lineal simple a un modelo de regresión lineal múltiple (RLM) con k variables explicativas (independientes) y una variable respuesta (dependiente).
- 1.2. Estimación de los coeficientes de regresión: Utilizar la notación matricial para expresar el modelo de RLM y encontrar los estimadores por el método de mínimos cuadrados.
- 1.3. Estimación de las varianzas: Estimación de la varianza de la variable respuesta, del error y de regresión para construir la tabla de análisis de varianza. Varianzas de los estimadores de los coeficientes de regresión y sus estimaciones.
- 1.4. Pruebas de hipótesis: Hipótesis del modelo de regresión y parciales y sus respectivas pruebas. Estadísticos de prueba y significación muestral.
- 1.5. Puntos atípicos y puntos de influencia, introducción al diagnóstico de la regresión. Multicolinealidad.
- 1.6. Selección del mejor modelo: Utilizando los criterios de significación de coeficientes y de R cuadrada.
- 1.7. Intervalo de confianza para la media e intervalo de predicción para un valor de la variable respuesta, cuando se suma un valor determinado para cada una de las variables explicativas.

2. Regresión curvilínea. (2 semanas)

- 2.1. Identificación de respuestas curvilíneas. Regresión polinomial (cuadrática, cúbica), centrada en la media ($(X - \bar{X})^2$ o $(X - \bar{X})^3$) y polinomios ortogonales. Otras transformaciones: exponencial, logarítmica, raíz cuadrada, etc. Transformación logarítmica de la variable respuesta para resolver la regresión exponencial.
- 2.2. Pruebas de hipótesis: del modelo de regresión y parciales.
- 2.3. Selección del mejor modelo.
- 2.4. La superficie de respuesta cuadrática en dos variables y la obtención del óptimo.

3. La normal multivariada y la correlación entre parejas de variables. (1 semana)

- 3.1. Distribuciones marginales y condicionales.
- 3.2. La correlación parcial.
- 3.3. Estimación de la correlación entre parejas de variables y la correlación parcial utilizando en esta última los residuales.

NOMBRE DEL PLAN LICENCIATURA EN MATEMÁTICAS		3/4
CLAVE 2131112	UNIDAD DE DE ENSEÑANZA-APRENDIZAJE REGRESIÓN	

4. Regresión logística. (4 semanas)

- 4.1. Regresión logística cuando la variable respuesta tiene distribución Bernoulli: Planteamiento del problema y del modelo.
- 4.2. Estimación de parámetros: Función de verosimilitud, funciones de enlace y estadísticos suficientes para estimar parámetros y construir sus intervalos.
- 4.3. Pruebas de hipótesis: Devianza y análisis de devianza para probar las hipótesis del modelo y parciales.
- 4.4. Regresión logística cuando la variable respuesta es multinomial.

MODALIDADES DE CONDUCCIÓN DEL PROCESO DE ENSEÑANZA-APRENDIZAJE

Se usará el paquete estadístico R para el manejo de datos y los análisis.

En la exposición de la teoría se formalizan los conceptos y se presentan ejemplos tomados de varias disciplinas aplicando los conceptos.

En la sesión de práctica los alumnos resuelven ejercicios dirigidos por el profesor. Se puede desarrollar en el salón de clases, o en un laboratorio de cómputo con la ayuda de un paquete computacional. Se buscará que el alumno elabore un acervo personal de métodos y estrategias para la solución de problemas, por ejemplo: leer el problema varias veces, definir variables e identificar los parámetros, identificar los datos y las preguntas a determinar con la solución, y validar e interpretar las soluciones.

Las sesiones de práctica serán organizadas con base en la resolución de problemas que incluyan problemas específicos de regresión.

Se utilizará, en la medida de lo posible, material de apoyo basado en las Tecnologías de la información y la comunicación.

El profesor promoverá que durante el transcurso de las horas teóricas y prácticas los alumnos expresen sus ideas y las expongan ante sus compañeros de manera que desarrollen su capacidad de comunicación oral.

El profesor fomentará que los alumnos realicen trabajos escritos en los que desarrollen su capacidad para comunicar sus ideas en forma escrita.

El profesor impulsará la elaboración de carteles o presentaciones en las que los alumnos comuniquen los conceptos aprendidos.

El profesor tomará especial cuidado en que los alumnos identifiquen y comprendan los argumentos correctos y erróneos tanto en sus participaciones en las clases como a través de sus trabajos escritos.

MODALIDADES DE EVALUACIÓN

GLOBAL

El profesor llevará a cabo al menos dos evaluaciones periódicas y, en su caso, una terminal. En la integración de la calificación se incorporarán aspectos como el desempeño en la solución de listas de ejercicios, la participación en clase y talleres, y la elaboración y presentación de proyectos. Los factores de ponderación serán a juicio del profesor.

En el proceso de evaluación el alumno deberá mostrar su capacidad de comprender y aplicar los conceptos desarrollados en el curso.

RECUPERACIÓN

A juicio del profesor, consistirá en una evaluación que incluya todos los contenidos teóricos y prácticos de la UEA o solo aquellos que no fueron cumplidos durante el trimestre.

NOMBRE DEL PLAN LICENCIATURA EN MATEMÁTICAS		4/4
CLAVE 2131112	UNIDAD DE DE ENSEÑANZA-APRENDIZAJE REGRESIÓN	

BIBLIOGRAFÍA NECESARIA O RECOMENDABLE

1. Dobson, A.J., *An introduction lo generalized linear models*, Chapman and Hall/CRC, 2002.
2. Draper, N.D., Smith, H., *Applied regression analysis*, John Wiley and Sons Inc., 1966.
3. Faraday, J. J., *Linear models with R*, Chapman and Hall/CRC, 2005.
4. Gujarati, D., *Econometría*, McGraw Hill, 2004.
5. Martínez, G. A., Castillo, M. A., *Teoría de la regresión con aplicaciones agronómicas*, 1987.
6. Montgomery, D.C., Peck, E. A., Vinning, C. G., *Introducción al análisis de regresiónlineal*, Ed. Patria Cultural, 2002.
7. Rao, C. R., Tontenburg, H., *Linear models, least squares and alternatives*, Springer, 1995.
8. Steel, R. G. D., Torrie, J. H., *Bioestadística: Principios y procedimientos*, 2^a Edición. McGraw-Hill /Interamericana de México S.A de C.V. 1997.