

UNIVERSIDAD AUTÓNOMA CHAPINGO
DIVISIÓN DE CIENCIAS FORESTALES
LICENCIATURA EN ESTADÍSTICA

MODELOS CATEGÓRICOS I

I. DATOS GENERALES

Unidad Académica:	División de Ciencias Forestales
Programa Educativo:	Licenciatura en Estadística
Asignatura:	MODELOS CATEGÓRICOS I
Carácter:	Obligatorio
Tipo:	Teórico-Práctico
Prerrequisitos:	Probabilidad II, Inferencia Estadística
Nombre del profesor:	Antonio Villanueva Morales
Ciclo Escolar:	2010-2011
Año:	5°
Semestre:	Segundo
Horas teoría/semana:	3 h
Horas práctica/semana:	1.5 h
Duración del semestre:	16 semanas
Horas totales del curso:	72 h
Créditos:	4
Cupo:	Limitado a 30 estudiantes

II. INTRODUCCIÓN

Una variable categórica tiene una escala de medición que consiste de un conjunto de categorías. Por ejemplo, la filosofía política de una persona es a menudo medida como liberal, moderada o conservadora. Diagnósticos para cáncer de pecho basados en una mamografía usan las categorías normal, benigno, probablemente benigno, sospechoso y maligno. El desarrollo de métodos para variables categóricas fue estimulado por investigaciones en las ciencias sociales y biomédicas. Las escalas categóricas se han propagado a las ciencias sociales para medir actitudes y opiniones. Las escalas categóricas en las ciencias biomédicas miden resultados tales como si un tratamiento es o no exitoso. Aunque los datos categóricos son comunes en las ciencias sociales y biomédicas, de ningún modo están restringidos a estas áreas. Los datos categóricos ocurren frecuentemente en las ciencias

del comportamiento (p. ej., tipo de enfermedad mental, con las categorías esquizofrenia, depresión, neurosis), epidemiología y salud pública (p. ej., método anticonceptivo en el último coito, con categorías ninguno, condon, píldora, DIU, otro), genética (tipo de alelo heredado por un descendiente), educación (p. ej., respuestas de un estudiante a una pregunta de examen, con las categorías correcta e incorrecta) y mercadotecnia (p. ej., preferencia del consumidor entre marcas líderes de un producto, con las categorías marca A, marca B y marca C).

En el curso Modelos Categóricos I se abordan métodos estadísticos estándar para el análisis de datos categóricos. El curso es teórico-práctico con clases tipo conferencia. Tanto la parte teórica como la parte práctica del curso serán presentadas en clase. La parte teórica abarca la presentación de los modelos teóricos que generan los datos categóricos y la parte práctica comprende la presentación de la metodología estadística y su ilustración a través del análisis de conjuntos de datos reales. Los recursos didácticos incluirán libros, pizarrón y el análisis de datos se ilustrará mediante el software estadístico SAS. El curso será evaluado con exámenes y tareas extraclase en las que se incluyen derivaciones teóricas y problemas que requieren aplicación de la metodología estadística presentada en clase en el análisis de conjuntos de datos.

El curso Modelos Categóricos I es de formación básica. A nivel horizontal se complementa en su contenido teórico y práctico con los cursos Análisis de Regresión, Diseños Experimentales I y Programación con SAS. A nivel vertical el curso tiene relación con los cursos Probabilidad I, Probabilidad II, e Introducción a la Inferencia Estadística ya que hace uso de los conceptos probabilísticos y de inferencia que son el fundamento de estos cursos y de la disciplina de la Estadística en general.

III. PRESENTACIÓN

En este curso se abordan los modelos distribucionales para respuestas categóricas y los métodos estadísticos tradicionales para el análisis de tablas de contingencia de dos vías (Unidades 1-3). También, se hace una introducción a los modelos lineales generalizados (Unidad 4).

IV. OBJETIVOS

Presentar los modelos distribucionales y métodos estadísticos tradicionales para el análisis de datos categóricos a fin de proveer al estudiante con las herramientas estadísticas apropiadas y el software computacional requerido para analizar este tipo de datos.

V. CONTENIDO

Unidad 1. Distribuciones e inferencia para datos categóricos (18 h)

Objetivo: Presentar los modelos probabilísticos y los métodos de inferencia

para datos categóricos a fin de proporcionar al alumno con el conocimiento de estos modelos y las habilidades técnicas para realizar inferencia estadística sobre sus parámetros.

- 1.1 Datos sobre respuestas categóricas
- 1.2 Distribuciones para datos categóricos
- 1.3 Inferencia estadística para datos categóricos
- 1.4 Inferencia estadística para parámetros binomiales
- 1.5 Inferencia estadística para parámetros multinomiales

Unidad 2 Descripción de tablas de contingencia (18 h)

Objetivo: Presentar tablas que despliegan relaciones entre variables categóricas y definir parámetros que describen esta asociación a fin de que el alumno conozca la estructura probabilística de una tabla de contingencia y adquiera las habilidades técnicas para describir asociación entre tales variables.

- 2.1 Estructura probabilística para tablas de contingencia
- 2.2 Comparación de dos proporciones
- 2.3 Asociación parcial en tablas 2x2 estratificadas
- 2.4 Extensiones para tablas IxJ

Unidad 3. Inferencia para tablas de contingencia (18 h)

Objetivo: Presentar los métodos inferenciales para tablas de contingencia a fin de proporcionar al alumno con el conocimiento y habilidades para implementarlos.

- 3.1 Intervalos de confianza para parámetros de asociación
- 3.2 Prueba de independencia para tablas de contingencia con dos criterios de clasificación
- 3.3 Pruebas de independencia chi-cuadrada
- 3.4 Tablas de dos criterios de clasificación con clasificaciones ordenadas
- 3.5 Pruebas de independencia para muestras pequeñas
- 3.6 Intervalos de confianza para muestras pequeñas en tablas 2x2
- 3.7 Extensiones para tablas con múltiples criterios de clasificación y respuestas no tabuladas

Unidad 4. Introducción a los modelos lineales generalizados (18 h)

Objetivo: Presentar una familia de modelos lineales generalizados que contiene los modelos más importantes para respuestas categóricas y los modelos estándar para respuestas continuas a fin de proporcionar al alumno con el conocimiento de la estructura de estos modelos así como de las técnicas correspondientes de inferencia.

- 4.1 El modelo lineal generalizado
- 4.2 Modelos lineales generalizados para datos binarios
- 4.3 Modelos lineales generalizados para datos de conteos
- 4.4 Momentos y verosimilitud para modelos lineales generalizados
- 4.5 Inferencia para modelos lineales generalizados
- 4.6 Ajuste de modelos lineales generalizados
- 4.7 Modelos lineales generalizados y quasi-verosimilitud
- 4.8 Modelos aditivos generalizados

VI. METODOLOGÍA

El curso es teórico práctico y se desarrollará mediante la exposición de material por parte del profesor. Adicionalmente, se asignarán tareas extra-clase durante todo el semestre. Estas tareas requerirán que el estudiante aplique conceptos presentados en clase a diversos conjuntos de datos. Las tareas son individuales, pero es recomendable debatir con sus compañeros de clase acerca de las respuestas más acertadas a los ejercicios de la tarea.

Las clases serán tipo conferencia. Tanto la parte teórica como la parte práctica del curso serán presentadas en clase. Los recursos didácticos incluirán libros, pizarrón y el análisis de datos se ilustrará mediante el software estadístico SAS.

VII. EVALUACIÓN

Tres exámenes parciales	75 % de la calificación final
Tareas extra-clase	15 % de la calificación final
Asistencia	5 % de la calificación final
Participación	5 % de la calificación final

El estudiante debe cumplir con al menos 85 % de asistencia al curso.

VIII. BIBLIOGRAFÍA

Básica

Agresti, A. (2002), *Categorical Data Analysis*, sec. ed., Wiley.

Agresti, A. (2007), *An Introduction to Categorical Data Analysis*, sec. ed., Wiley.

Complementaria

Stokes, M. E., Davis, C. S. and Koch, G. G. (2003), *Categorical Data Analysis Using The SAS System*, sec. ed., SAS Institute and Wiley.

Lawal, B. (2003), *Categorical Data Analysis with SAS and SPSS Applications*, Lawrence Erlbaum Ass.