



UNIVERSIDAD DE COSTA RICA
FACULTAD DE CIENCIAS ECONOMICAS
ESCUELA DE ESTADISTICA

XS-0301: ESTADÍSTICA PARA BIÓLOGOS II

PROGRAMA
I Semestre 2017

Docente: Rebeca Sura Fonseca

Correo electrónico: cursos.sura@gmail.com

Oficina: Laboratorio de Profesores, Escuela de Estadística, 3er Piso, Facultad de Ciencias Económicas

Horario de clases: L: 9:00-11:50 (teoría)
M: 14:00-15:50 (lab. Grupo 2)
J: 10:00-11:50 (lab. Grupo 1)

Horas de consulta: M: 8:00-12:00

1. Descripción

Este curso supone el dominio de la estadística descriptiva y conocimientos de inferencia estadística incluyendo diversos tipos de pruebas para comparar dos grupos.

- Requisitos: XS-0300 Estadística para biólogos I.
- Correquisitos: Ninguno.
- Horas: Dos sesiones por semana - 5 horas semanales (3 teoría y 2 práctica).
- Créditos: 5.

2. Objetivo General

Proveer a los estudiantes de la carrera de Biología del conocimiento conceptual y los instrumentos estadísticos avanzados para la investigación científica en su campo de acción.

3. Objetivos Específicos

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

- Entender y aplicar la técnica del análisis de variancia y los procedimientos de comparaciones múltiples para diferentes tipos de experimentos.
- Seleccionar un conjunto de variables para formar un modelo de regresión apropiado mediante un análisis detallado de todos los predictores disponibles.
- Llevar a cabo los diagnósticos del modelo de regresión ajustado.
- Comprender el modelo de regresión logística para respuestas dicotómicas.
- Seleccionar y aplicar la técnica multivariada apropiada cuando se tienen conjuntos de datos complejos.
- Usar adecuadamente el lenguaje de programación estadístico R para llevar a cabo las diferentes pruebas y técnicas estudiadas en el curso.

4. Contenidos

Módulo I: Diseño de experimentos.

- 1.1 Análisis de variancia de una vía (diseño irrestricto aleatorio):
 - a. Fuentes de variación: entre tratamientos y dentro de tratamientos.
 - b. Comparación de fuentes de variación (tabla del ANDEVA).
 - c. Prueba de homogeneidad de variancias.
 - d. Pruebas robustas.
 - e. Comparaciones múltiples: contrastes de Duncan, Tukey, etc.
- 1.2 Potencia o poder de la prueba:
 - a. La importancia de la magnitud detectada por la prueba.
 - b. Determinación del tamaño de muestra para pruebas de hipótesis.
 - c. La potencia de una prueba realizada.
- 1.3 Otros diseños:
 - a. Arreglos factoriales.
 - b. Diseños con interacciones.
 - c. Bloques aleatorios / medidas repetidas.
 - d. Diseños anidados.
- 1.4 Pruebas no-paramétricas:
 - a. Prueba de Kruskal Wallis.
 - b. ANDEVA no-paramétrico con dos factores.
 - c. Análisis de variancia de Friedman.
 - d. Comparaciones múltiples.

Módulo II: Regresión.

- 2.1 Correlación lineal entre dos variables.
- 2.2 El modelo de regresión múltiple:
 - a. Estimación de los coeficientes.
 - b. Valores ajustados y residuales.
 - c. Prueba de hipótesis sobre los coeficientes.
 - d. Estimación de respuesta media y predicción de nuevas observaciones.
 - e. Coeficiente de determinación.
- 2.3 Diagnósticos.
- 2.4 Transformaciones, regresión polinomial y otras soluciones a problemas con los supuestos.
- 2.5 Regresión para respuestas de conteos (Poisson)
- 2.6 Regresión logística

Módulo III: Análisis multivariado.

- 3.1 Análisis de componentes principales (PCA).
- 3.2 Análisis discriminante.
- 3.3 Análisis de conglomerados (Cluster).

5. Metodología

- Se impartirán lecciones magistrales por parte del docente donde se explicarán los conceptos y sus aplicaciones.
- Se usará el laboratorio de computadoras de la escuela de biología para el desarrollo de 12 lecciones prácticas durante el semestre, las cuales corresponden a varios laboratorios por cada tema del programa.
- Durante las sesiones de laboratorio se utilizará el lenguaje de programación estadístico R para realizar ejercicios de la materia vista en clase, así como profundización en el análisis de los resultados.
- Se asignarán tareas y ejercicios para asegurar la “puesta en práctica” de los conceptos estudiados. Se asignará una tarea por cada dos o tres laboratorios de tal forma que en total serán 5 tareas.
- Los estudiantes realizarán un trabajo de análisis de datos reales utilizando lo aprendido en el curso. Para esto presentarán un anteproyecto donde expondrán los objetivos del estudio y un plan de análisis el cual será criticado por el profesor y devuelto para que se realicen las mejoras sugeridas. Al final de curso se espera que los estudiantes escriban un artículo donde se describa detalladamente el procedimiento de análisis de los datos seleccionados y expongan los resultados en la clase.
- Se usarán como textos: Biostatistical Analysis de Zar y Análisis multivariado de Catena et al.

6. Evaluación

Se realizarán tareas, un trabajo de análisis de datos y tres exámenes parciales en el laboratorio de cómputo. Además los estudiantes presentarán un trabajo de análisis de datos. En el siguiente cuadro se presenta el desglose porcentual de la nota:

Rubro	Ponderación
Primer examen parcial	20%
Segundo examen parcial	20%
Tercer examen parcial	20%
Tareas	20%
Trabajo de análisis de datos	20%
Total	100%

Si un estudiante faltara a algún examen por **causa justificada**, debe solicitar por escrito la reposición del examen indicando las razones de la ausencia, acompañada de los documentos justificantes. La misma debe entregarse ante el profesor que imparte el curso **a más tardar en cinco días hábiles del reintegro a lecciones**. No se recibirán exámenes ni tareas que incumplan con el tiempo establecido para su entrega.

7. Cronograma

Mes	Módulo	Día			Actividad	Evaluación		
		L	M	J				
MARZO	I. DISEÑO DE EXPERIMENTOS	13						
			15	16	Lab.0 - Intro a R			
		20						
			22	23	Lab.1 - ANOVA una vía			
		27						
ABRIL		I. DISEÑO DE EXPERIMENTOS		29	30	Lab.2 - Pruebas múltiples		
			3					
				5	6	Lab.3 - Factoriales		
			10			Semana Santa		
				12	13			
17								
MAYO	II. REGRESIÓN			19	20	Lab.4 - Bloques		
			24				Semana Universitaria	
				26	27	Lab.5 - Regresión simple		
			1			Día del trabajador		
			3	4		Examen 1		
MAYO		II. REGRESIÓN	8				Anteproyecto	
				10	11	Lab.6 - Predicciones		
			15					
				17	18	Lab.7 - Diagnósticos		
			22					
			24	25	Lab.8 - Poisson - transformaciones			
	29							
			31	1	Lab.9 - Logística			
JUNIO	III. ANÁLISIS MULTIVARIADO		5					
				7	8	Lab.10 -Discriminante		
		12						
			14	15		Examen 2		
		19						
			21	22	Lab.11 - Componentes principales			
26								
		28	29	Lab.12 - Cluster				
JULIO		III. ANÁLISIS MULTIVARIADO	3				Entrega y exposición del Trabajo final	
				5	6			
	10							
			12	13		Examen 3		

8. Referencias bibliográficas

Catena et al. (2003). *Análisis multivariado: un manual para investigadores.* Madrid: Editorial Biblioteca Nueva.

Johnson, D. (2000). *Métodos multivariados aplicados al análisis de datos.* International Thomson Editores.

Montgomery, D. (2005). *Diseño y Análisis de Experimentos.* 2ª ed. Limusa Wiley.

Neter et al. (1996). *Applied Linear Statistical Models.* 3ª ed. WCB McGraw-Hill.

Ramsey, F.L. y D.W. Schafer (2002). *The Statistical Sleuth: A Course in Methods of Data Analysis.* Duxbury, Australia: Thomson Learning.

Steel, R.G.D. y J.H. Torrie. (1985). *Bioestadística: principios y procedimientos.* Bogotá, Colombia: McGraw-Hill.

Zar, J.H. (1996). *Biostatistical Analysis.* 3a ed. Englewood Cliffs, N.J. : Prentice Hall.