



XS-0301: ESTADÍSTICA para BIÓLOGOS II, I Semestre 2023
Escuela de Estadística, Facultad de Ciencias Económicas, Universidad de Costa Rica

Docente: Hazel Quesada Leitón (G1), Jimena Murillo Montero (G2)

Correo electrónico: hazel.quesadaleiton@ucr.ac.cr, jimena.murillo@ucr.ac.cr

Oficina: Escuela de Estadística, Facultad de Ciencias Económicas

Horario de clases: L: 9:00-11:50 (teoría)

M: 14:00-15:50 (lab. G2)

J: 10:00-11:50 (lab. G1)

Horas de consulta: Grupo 1 (Hazel): J:8:00-9:50

Grupo 2 (Jimena): L: 13:00-14:50

1. Descripción

Este curso se ubica en el bloque de asignaturas correspondientes al quinto semestre del plan de estudios de la carrera de Biología. El curso se orienta a dotar a los estudiantes de herramientas estadísticas útiles en diseño de experimentos, regresión y análisis multivariado. Se pretende desarrollar el sentido crítico del estudiante, fomentando su capacidad para afrontar y resolver problemas biológicos a partir de un planteamiento estadístico. Se pretende introducir una concepción estadística para la solución de problemas biológicos reales, orientada a obtener una respuesta satisfactoria.

El curso tiene un componente teórico donde se exponen las bases conceptuales de los métodos estadísticos y un componente práctico donde se expone al estudiante a datos reales para el análisis en el computador. Para la aplicación de los principios y métodos a cubrir en Estadística para Biólogos II es necesario dominar los conceptos fundamentales de la estadística descriptiva e inferencial.

- 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.



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 de conglomerados (Cluster).

5. Metodología

- El curso se impartirá de manera presencial, haciendo uso de la plataforma institucional Mediación Virtual para colocar materiales de clase tales como documentos, presentaciones y evaluaciones del curso, por tanto, es un curso de baja virtualidad.
- Se impartirán lecciones magistrales por parte del docente donde se explicarán los conceptos y sus aplicaciones.
- Cada persona utilizará las computadoras del laboratorio donde se impartirán las lecciones o su propia computadora para desarrollar 12 lecciones prácticas durante el semestre.
- Durante las sesiones de laboratorio se utilizará el lenguaje de programación estadístico R para realizar ejercicios de la materia vista en clase.



- El profesor enviará los laboratorios con anticipación de tal forma que los estudiantes los hayan revisado previamente. Durante la sesión de laboratorio se evacuarán dudas sobre el uso del software y se profundizará en el análisis.
- 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 3 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.
- Se usarán como libros de texto: Biostatistical Analysis de Zar y Análisis multivariado de Catena et al.

6. Evaluación

Se realizarán tareas y tres exámenes parciales en su computadora. 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	27.5%
Segundo examen parcial	27.5%
Pruebas cortas (quices)	8%
Tareas	12%
Trabajo de análisis de datos	25%
Total	100%

Se realizarán quices para apoyar el proceso de aprendizaje, mismos que no se avisarán y pueden cubrir cualquier tema visto en clase (incluidas las lecturas asignadas). Dado su naturaleza, los quices no se repetirán (al no tener fechas preestablecidas). Según el artículo 15 del Reglamento Académico Estudiantil es posible realizar quices sin anunciar al estudiante la fecha con antelación, en la medida que se especifique en el programa del curso. Únicamente las evaluaciones estipuladas en el artículo 18 y que son anunciadas al estudiante son sujeto de reposición (o sea exámenes parciales y de ampliación) (oficio OJ-1196-2010 de la Oficina Jurídica). Las tareas deben ser entregadas el día y hora estipulada por el profesor(a). No se reciben ni evalúan tareas entregadas luego de la fecha/hora.

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.**

La causa debe estar contemplada dentro del Reglamento de Régimen Académico Estudiantil, aprobado por el Consejo Universitario en la sesión 4632 del 3 de mayo de 2001. **(sólo el profesor o la profesora del curso recibirá dicha justificación, ni el coordinador de la cátedra, ni la secretaría de la Escuela de Estadística recibirán las mismas).**

Se entenderá por causa justificada: 1. Enfermedad comprobada mediante dictamen médico 2. Choque en día y hora con otro **examen dentro de la U.C.R.** (siempre y cuando sea un examen de cátedra) (el estudiante deberá presentar una constancia con la firma del profesor y sello de la Unidad Académica respectiva, donde se indique el horario donde el estudiante realizó el examen). 3. Otra causa grave (muerte de pariente en 1er o 2o grado y causas fortuitas).

Según el oficio OJ-1306-2008 de la Oficina Jurídica, se entenderá como causa fortuita: “acontecimientos que no han podido perverse y cuyas circunstancias deben ser irresistibles o inevitables, siendo impotente el hombre para impedir su ocurrencia”. Finalmente, el estudiante sólo tendrá dos oportunidades para realizar su evaluación, fechas definidas en este documento. El reglamento citado sólo considera la realización de un examen de reposición de cada evaluación ordinaria programada con antelación, situación ratificada en el oficio OJ-693-2014 de la Oficina Jurídica.



El examen de AMPLIACION que incluye **toda** la materia del curso se realizará a todos aquellos estudiantes cuya nota final sea 6,0 o 6,5. El estudiante que obtenga 7,0 o más en este examen aprobará el curso con nota de 7,0.

7. 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.



8. Cronograma

Mes	Módulo	Martes	Miércoles	Jueves	Actividad
Marzo	Diseño de experimentos	11	13	14	Lab 1 ANOVA una vía
		18	20	21	Lab 2 pruebas de hipótesis
		25	27	28	Semana santa
		1	3	4	Lab 3 Factoriales/ Entrega 1
Abril	Diseño de experimentos	8	10	11	Lab 4 bloques
		15	17	18	Examen 1
		FERIADO			
	Examen Parcial 1	22	24	25	Lab 5 regresión lineal
		Semana universitaria			
Mayo	Regresión	29	1	2	Lab 6 Predicciones
		6	8	9	Lab 7 Diagnósticos
		13	15	16	Lab 8 Poisson/ Entrega 2
		20	22	23	Lab 9 Logística
		27	28	29	Lab 10 PCA
		3	5	6	Lab 11 Clusters
		10	12	13	Practica/ Entrega 3
Junio	Análisis multivariado	17	19	20	Examen 2
		24	26	27	Entrega trabajo final
	Examen Parcial 2	1	3	4	Exposiciones
		9	11	12	Ampliación
Julio	Exposiciones	Exposiciones	Entrega de notas		
	Examen de ampliación		Ampliación		

