



## PROGRAMAS DE CURSOS Y ASIGNATURAS

**TÍTULO DEL CURSO:** Modelos de regresión en ecología y ciencias ambientales

**AÑO ACADÉMICO:** 2024

**CARRERA:** Doctorado en Biología - Plan de Estudios Ord. N° 556/86, Modif. 557/10 y 807/17

**FECHA DE DICTADO:** desde: 5/08/2024 - hasta: 9/08/2024

**DOCENTE/S RESPONSABLE/S:** Pablo A. E. Alarcón

**DOCENTE/S COLABORADORES/S:** Luca Schenone

**CARGA HORARIA TOTAL:** 40

**FUNDAMENTACIÓN:** Una parte importante de la ecología se esfuerza por hallar “leyes” generales que permitan explicar y predecir el comportamiento de los sistemas naturales, y la construcción de modelos cuantitativos es parte esencial del proceso (Begon et al. 2006, Bolker 2008, Lehman et al. 2019). La ecología empírica del siglo XX (1900’s) se desarrolló especialmente sobre la base de modelos estadísticos estimados bajo inferencia clásica (frecuentista) aunque, en las últimas décadas, la perspectiva Bayesiana ha ganado terreno en los estudios de ecología (ej. Bolker 2008). La concepción Bayesiana de la inferencia estadística se diferencia en aspectos fundamentales de la inferencia estadística frecuentista, situación que ha derivado en un largo debate filosófico. Sin embargo, existe un consenso más generalizado respecto de las bondades prácticas de los métodos Bayesianos, entre las que se cuentan la facilidad para implementar modelos complejos (ej. modelos jerárquicos) e interpretar los modelos estimados. Este curso capacitará a los alumnos en la construcción, la implementación y el diagnóstico de un conjunto de modelos de regresión haciendo uso de un marco de inferencia estadística Bayesiana.

**PROGRAMA ANALÍTICO:** Unidad 1. Conceptos generales sobre la construcción de modelos en ciencia. El concepto de modelo. Tipos de modelos en ciencia: verbales vs. cuantitativos; determinísticos vs. estocásticos; fenomenológicos vs. mecanísticos; descriptivos vs. predictivos; generales vs. particulares; inductivos vs. deductivos.

Unidad 2. Conceptos generales sobre los modelos de regresión. Los modelos de regresión en el universo de modelos. Estructura general de los modelos de regresión: componente determinístico y componente estocástico. Funciones matemáticas y distribuciones de probabilidad. Tipos de preguntas que pueden ser exploradas a partir de los modelos de regresión.

Unidad 3. Estructura del Modelo Lineal General (LM). Función lineal de primer grado como componente determinístico. Distribución Normal como componente estocástico. Modelos de

la media y de la varianza. Predictores continuos y categóricos. Modelos aditivos y con interacción entre variables. Supuestos de un LM. Diagnóstico de un LM.

Unidad 4. Estructura del Modelo Lineal Generalizado (GLM). Función lineal de primer grado como componente determinístico. Distribuciones binomial, Poisson y gamma como componentes estocásticos. Función de enlace. Predictores continuos y categóricos. Modelos aditivos y con interacción entre variables. Supuestos de un GLM. Diagnóstico de un GLM.

Unidad 5. Estructura del Modelo Lineal y Mixto (LMM). Estructuras de organización jerárquica en la naturaleza. Conceptos de “no pooling”, “complete pooling” y “partial pooling”. Modelos de intercepto y pendientes variables. Modelos para diferentes niveles de jerarquía. Supuestos de un LMM. Diagnóstico de un LMM.

Unidad 6. Estructura del Modelo Lineal Generalizado y Mixto (GLMM). Función lineal de primer grado como componente determinístico. Distribuciones binomial, Poisson y gamma como componentes estocásticos. Función de enlace. Supuestos de un LMM. Diagnóstico de un Modelo Lineal Generalizado y Mixto.

Unidad 7. Marco de inferencia Bayesiana. Probabilidad condicional y probabilidad subjetiva (Bayesiana). Teorema de Bayes. Distribución posterior, distribución previa y función de likelihood. Aproximación analítica y numérica de la distribución posterior de un modelo. Cadenas de Markov Monte Carlo (MCMC): algoritmo de Metropolis y Gibbs sampler. Intervalos de credibilidad. Posterior predictive check.

**OBJETIVOS:** El objetivo del curso es asistir a los alumnos en el desarrollo de habilidades que les permitan construir, implementar y diagnosticar modelos estadísticos en estudios sobre la biodiversidad.

**ACTIVIDAD PRÁCTICA:** Desde el punto de vista organizacional, este curso contempla clases teóricas. Todas las clases se realizarán de manera virtual a través de Zoom. A continuación, se describen las principales características de cada una.

Clases teóricas. Las clases teóricas consistirán de presentaciones expositivas por parte del docente que introducirán y conectarán los distintos contenidos del programa. En concreto, estas clases tendrán como objetivo revisar los conceptos más importantes y necesarios para transitar las tres etapas fundamentales de la construcción de un modelo: 1) la construcción conceptual del modelo, 2) la implementación y la estimación del modelo en computadora y 3) el diagnóstico del modelo estimado. Para lograr este objetivo, las clases teóricas repasarán los conceptos fundamentales de álgebra y probabilidad que permitirán construir el andamiaje de los modelos de regresión. Seguidamente, estas clases repasarán conceptos relacionados con la inferencia estadística, y particularmente con la inferencia Bayesiana, y esbozarán la mecánica de los algoritmos utilizados para estimar un modelo a partir de datos. Finalmente, las clases teóricas harán hincapié en la importancia de diagnosticar los modelos estimados y expondrán la lógica de algunas estrategias utilizadas con este propósito. El curso contempla un total de diez clases teóricas durante los primeros cuatro días de curso. Cada clase constará de dos bloques de una hora aproximadamente separados por un receso de 30 minutos.

Clases prácticas. Las clases prácticas consistirán de ejercicios y situaciones-problema que los alumnos deberán resolver en forma individual y grupal. Estas clases tendrán como objetivo general que los alumnos se ejerciten en la construcción conceptual de los modelos de regresión e indaguen en los aspectos técnicos que les permitirán estimarlos y diagnosticarlos. En particular, los alumnos se ejercitarán en la construcción de modelos con aplicabilidad a sistemas ecológicos y en el manejo de los programas informáticos R y JAGS, ambos de acceso libre y gratuito. Las clases prácticas se intercalarán con las clases teóricas siempre que sea posible, de manera que los alumnos puedan prontamente aplicar los conceptos revisados para resolver situaciones concretas. El curso constará de un total de ocho clases prácticas de dos horas cada una durante los primeros cuatro días de curso.

**EVALUACIÓN Y CONDICIONES DE ACREDITACIÓN:** El curso incluirá cuatro instancias de evaluación parcial y una instancia de evaluación final. Las evaluaciones parciales (“parcialitos”) consistirán de un examen escrito que explorará mayormente aspectos conceptuales y se completarán durante los primeros 10 minutos de cada clase. El principal propósito de estas evaluaciones es promover en los alumnos el hábito de estudiar y revisar los conceptos trabajados, a fin de que ellos mismos puedan reconocer las temáticas en las que deben volcar mayor esfuerzo de estudio. La instancia de evaluación final consistirá de un trabajo práctico en la que los alumnos deberán implementar las técnicas estudiadas durante el curso sobre una base de datos reales. Este trabajo deberá presentarse oralmente 15 días posteriores a la finalización del curso. La aprobación del curso se logrará mediante la aprobación del 75% de los exámenes parciales y un 70 % de la evaluación final.

**MODALIDAD DE DICTADO: PRESENCIALIDAD COMBINADA:** se desarrolla al mismo tiempo de manera combinada en edificios, instalaciones y entornos virtuales institucionales.

- Estrategias pedagógicas: El curso implementará dos estrategias de aprendizaje principales. La primera consistirá en concebir la construcción de un modelo como un proceso que transita por tres etapas principales y sucesivas: 1) la construcción conceptual del modelo, 2) la implementación y la estimación del modelo en computadora y 3) el diagnóstico del modelo estimado. Esta estrategia debería permitir a los alumnos reconocer fácilmente el valor conceptual y práctico que los contenidos del programa tienen en el contexto de los modelos de regresión y, por lo tanto, facilitar su asimilación. La segunda estrategia consistirá en explorar técnicas progresivamente más sofisticadas usando continuamente a las técnicas más sencillas como punto de referencia y comparación. Para tal fin, el curso se iniciará con la “deconstrucción” del modelo de regresión lineal simple en sus partes componentes, ejercicio que facilitará modificar y/o extender su estructura hacia versiones más sofisticadas. En forma similar, las técnicas Bayesianas se estudiarán mediante el análisis comparativo con las técnicas de la estadística clásica (frecuentista) con la que los alumnos estarán probablemente mejor familiarizados. Es importante destacar que la inferencia estadística clásica se estudiará con el único propósito de establecer un marco de referencia conceptual; sin embargo, el curso no buscará necesariamente que los alumnos aprendan a aplicar estas técnicas.

- Carga horaria y actividades destinadas a las diferentes modalidades: 40 hs totales: 20 hs de clases teóricas y 20 hs de clases prácticas

- Interacciones docente-estudiantes y estudiantes-estudiantes previstas: Las clases teóricas serán clases expositivas pero que incentivarán la interrupción reiterada por parte de los

estudiantes para compartir ideas, comentarios, preguntas y/o puntos de vista. En las clases prácticas se espera que los alumnos trabajen de manera independiente, aunque los docentes estaremos atentos para responder preguntas y resolver inconvenientes técnicos. Los docentes intentaremos generar un ambiente relajado que incentive la participación, la cooperación entre los estudiantes y la búsqueda de soluciones creativas a los problemas planteados en las actividades.

- Mecanismos de seguimiento, supervisión y evaluación de las actividades: Al inicio de cada clase teórica se implementarán evaluaciones parciales (“parcialitos”) de manera que los alumnos puedan autoevaluarse en relación a la asimilación de conceptos generales. Asimismo, las clases teóricas y las clases prácticas serán intervenidas por parte del docente para incentivar la participación de los estudiantes y así supervisar el seguimiento que los alumnos hacen de los temas estudiados.

**BIBLIOGRAFÍA:** Bolker, B. M. Ecological models and data in R. Princeton University Press, 2008.

o Gelman, A., y J. Hill. Data analysis using regression and multilevel/hierarchical models. Cambridge university press, 2006.

o Kéry, M., y J. A. Royle. Applied Hierarchical Modeling in Ecology: Analysis of distribution, abundance and species richness in R and BUGS: Volume 1: Prelude and Static Models. Academic Press, 2015.

o Kruschke, J. Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan. Academic Press, 2014.

o Lehman, C., S. Loberg, y A. Clark. Quantitative Ecology: A New Unified Approach. University of Minnesota Libraries Publishing, 2019.

o Matthiopoulos, J. How to be a quantitative ecologist: the 'A to R' of green mathematics and statistics. John Wiley & Sons, 2011.

o McElreath, R. Statistical rethinking: A Bayesian course with examples in R and Stan. CRC press, 2020.

o Taper y Lele. The nature of scientific evidence: statistical, empirical and philosophical considerations. The University of Chicago Press, 2004.