



PROGRAMA PARA CURSOS

DENOMINACIÓN DEL CURSO: Modelos y Datos en Ecología

AÑO ACADÉMICO: 2025

CARRERA: DOCTORADO EN BIOLOGÍA

Plan de Estudios Ord. N° 556/86, Modif. 557/10 y 807/17

DOCENTE/S A CARGO: Juan Manuel Gurevitz

DOCENTE/S AUXILIAR/ES:

MODALIDAD: Presencialidad física: se desarrolla en edificios e instalaciones institucionales (presencialidad convencional).

FECHA PROPUESTA: 05/05/2025-09/05/2025

DESTINATARIOS: Estudiantes de posgrado y profesionales de ecología como de otras ramas de la biología y otras disciplinas (química, ingeniería, economía, sociología), con interés en entender y cuantificar la variabilidad en los datos a través de modelos de regresión y en explorar el enfoque bayesiano.

Si bien no es excluyente, se recomienda fuertemente contar con experiencia programando en R o similar (por ej., Matlab, C, C++, Python) o haber realizado algún curso sobre programación en R. El CRUB (Univ. del Comahue) brinda el curso de posgrado “Empieza con R”. También hay diversos cursos online, varios gratuitos. Cada estudiante deberá traer su computadora portátil.

El contenido del curso supone conocimientos matemáticos y estadísticos generales y alguna experiencia en análisis de regresión convencionales.

FUNDAMENTACIÓN: Los modelos de regresión son probablemente la herramienta que más utilizamos para evaluar la relación entre diferentes variables. Partiendo de la versión más simple, la regresión lineal, existen diversas variantes y sofisticaciones (LM, GLM, GLMM, GAM, GAMM y otros). Muchas veces cada variante se presenta como algo sustancialmente diferente. Sin embargo, existe un marco general que integra todos estos modelos y otros, dándonos una comprensión más integral y, a la vez, permitiéndonos una flexibilidad casi ilimitada para formular modelos que representen mejor nuestras ideas sobre los sistemas que estudiamos.

Luego de formular nuestro modelo, el ajuste a datos se puede realizar con diversas herramientas y bajo diferentes enfoques. El enfoque más habitual sigue siendo el frecuentismo. Sin embargo, el enfoque bayesiano es una alternativa cada vez más utilizada por resultar más versátil e intuitiva y por describir de forma más detallada la incertidumbre en los ajustes.

Al menos en el ámbito de la biología, la formulación de modelos (matemáticos, de regresión) y la estadística son temáticas que generan reticencia y resultan tediosas, muchas veces por resultar difíciles, intrincadas o poco comprensibles. Sin embargo, esto no necesita ser así. Una buena (y totalmente factible) comprensión de los conceptos centrales y de las herramientas necesarias nos permite abordar con fluidez y facilidad los modelos y la estadística, ambos fundamentales para mejorar nuestra comprensión de la realidad de una manera robusta. Es así que en este curso nos proponemos lograr que la formulación de modelos y su ajuste a datos con estadística bayesiana resulten algo accesible, práctico y fluido. Para ello, siguiendo ciertos lineamientos generales, nos ejercitaremos en cómo representar mejor la pregunta conceptual o ideas que tengamos sobre nuestro sistema de estudio, analizando y reflexionando sobre diferentes representaciones posibles. Junto a esto, partiendo de nuestras intuiciones, veremos en qué consiste el enfoque bayesiano y sus implicancias para la interpretación de nuestros modelos y para el ajuste a datos. A partir de nociones básicas e intuitivas de estadística y matemática comprenderemos cómo funciona un algoritmo de estimación y nos entrenaremos en utilizar el lenguaje de programación Stan (desde R) como herramienta para ajustar nuestros modelos a datos bajo un enfoque bayesiano. Utilizaremos scripts muy accesibles y explicados detalladamente, permitiendo ejecutar los análisis sin necesariamente contar con un dominio de programación en R y Stan.

Si bien proviene de una tradición en ecología, este curso puede resultar de interés y utilidad para otras ramas de la biología y otras disciplinas (por ej., química, ingeniería, economía, sociología). Durante la cursada, serán centrales los aportes que cada cursante haga a partir de sus intereses y temas de trabajo, tanto para poder brindar soluciones específicas como para lograr un panorama más amplio de las aplicaciones posibles.

OBJETIVOS: El curso busca contribuir al desarrollo de las siguientes capacidades:

- Formalizar en diferentes tipos de modelos de regresión las ideas o preguntas sobre un sistema.
- Comprender las nociones básicas de las herramientas estadísticas para conectar modelos con datos bajo un enfoque bayesiano.
- Implementar computacionalmente la formulación de modelos y su ajuste a datos bajo un enfoque bayesiano.
- Comunicar y discutir los modelos desarrollados y ajustados a datos.

PROGRAMA ANALÍTICO: - Los modelos como representación del pensamiento.

Generalidades de modelos fenomenológicos y mecanísticos. Modelos de regresión.

Formulación de modelos a partir de preguntas, hipótesis o conceptos. Lineamientos para formalizar ideas en modelos. Modelando la variabilidad.

- Determinismo y aleatoriedad. Nociones de probabilidad, frecuentismo, bayesianismo.

Distribuciones de probabilidad habituales, continuas y discretas.

- Modelos (no) lineales generalizados. Formulación matemáticas habituales. Variables continuas, discretas y categóricas. Componentes determinístico y estocástico. Función de enlace. Modelos como generadores de datos. Simulación de datos a partir de modelos.

- Introducción a modelos jerárquicos, lógica, notación. Fuentes de variabilidad y su cuantificación; variabilidad de proceso y variabilidad de observación. Modelos mixtos, efectos aleatorios.

- Conexión entre modelos y datos. Optimización, ajuste de modelos a datos, sobreajuste y generalidad. Teorema de Bayes. Función de verosimilitud, distribuciones previa y posterior. Estimación de incertidumbre.

- Muestreo de distribuciones posteriores. Montecarlo basado en cadenas de Markov (MCMC), Metropolis-Hastings, Hamiltonian Montecarlo.
- Evaluación de modelos, análisis de residuos, sobre/subdispersión, chequeos predictivos previo y posterior.
- Implementación de modelos de regresión en R. Nociones básicas de R y de programación. Ajuste bayesiano de modelos utilizando Stan desde R. Exploración gráfica de modelos y datos en R.

ACTIVIDAD PRÁCTICA / SALIDA DE CAMPO: El dictado del curso consistirá en clases teóricas y clases prácticas. En las clases teóricas se presentarán los conceptos centrales de cada tema. En las clases prácticas, los estudiantes aplicarán a través de la implementación informática los conceptos vistos en las clases teóricas. Se usará intensivamente el programa R (con sus diversos paquetes); para análisis bayesianos se introducirá el uso de Stan en R. Todos estos programas y paquetes son de uso libre y gratuito. Cada estudiante deberá traer su computadora portátil con los programas instalados.

EVALUACIÓN Y ACREDITACIÓN: La evaluación se compondrá de parcialitos diarios y una evaluación final. Los parcialitos constarán de un breve cuestionario sobre los temas visto cada día y servirán para repasar los temas vistos y detectar los puntos que requieran refuerzo. La evaluación final se basará en la elaboración de un trabajo donde cada estudiante elija un set de datos (propio o de bases públicas de datos), defina una pregunta y la aborda a través de elaborar un modelo de regresión y de su ajuste a los datos. Este trabajo podrá entregarse hasta 15 días después de terminado el curso.

La aprobación del curso se logrará con una nota final mayor o igual a 7 (siente). Esta nota final se compondrá en un 40% del promedio de los parcialitos y el restante 60% de la nota del trabajo final.

BIBLIOGRAFÍA BÁSICA: Bolker, B. 2008. Ecological Models and Data in R. Princeton University Press.

Gelman, A. and J. Hill. 2006. Data Analysis Using Regression and Multilevel/Hierarchical models. Cambridge University Press.

Kruschke, J. 2015. Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. Academic Press.

McElreath, R. 2020. Statistical Rethinking: A Bayesian Course with Examples in R and Stan. 2nd Edition. CRC Press.

(El material estará disponible en formato digital)

BIBLIOGRAFÍA COMPLEMENTARIA U OPTATIVA: Gelman, A., J. Carlin, H. Stern, D. Dunson, A. Vehtari, and D. Rubin. 2014. Bayesian Data Analysis. 3rd edition. Chapman & Hall/CRC.

Stan Development Team. 2024. Stan User's Guide. Version 2.36. https://mc-stan.org/docs/2_36/stan-users-guide/

Stan Development Team. 2024. Stan Reference Manual. Version 2.36. https://mc-stan.org/docs/2_36/reference-manual/

APOYO TÉCNICO REQUERIDO: Proyector. Zapatillas para conectar laptops. Pizarrón y marcadores. Conexión a internet.



DEPARTAMENTO DE POSTGRADO

Universidad Nacional del Comahue
Centro Regional Universitario Bariloche

Quintral 1250

Tel: 0294 – 4423374 / 4428505- Interno 454

deptopostgradocrub@gmail.com



SOPORTE: PEDCO

CARGA HORARIA TOTAL: 40

Horas Asincrónicas:

CRONOGRAMA: Se adjunta.

Juan M. Gurevitz