

# **Introducción a los métodos estadísticos bayesianos en Ecología**

## **Objetivos:**

Proporcionar una adecuada comprensión de los principios fundamentales de los métodos de análisis bayesianos y adquirir la capacidad de utilizarlos en forma crítica y razonada estos métodos en sus actividades de investigación.

## **Docente:**

Pablo Inchausti, profesor titular de Ecología de la Universidad de la República, Uruguay.

## **Modalidad:**

El curso tendrá 6h de clase diarias (dos sesiones de 3h) durante 5 días de clase (30h en total). El curso será dictado en Español. Se empleará el programa de análisis estadístico gratuito R (así como de la interfase R-Studio) que se ha convertido en una referencia mundial para el análisis de datos. Se espera que los estudiantes utilicen sus laptops durante los prácticos de este curso. Todos materiales (clases, datos, scripts, libros y artículos) se pondrán a disposición de los estudiantes antes del comienzo del curso.

## **Requisitos:**

Este curso presupondrá un conocimiento básico sobre el concepto de probabilidad, de las principales funciones de distribución de probabilidades (binomial, Poisson y normal), de nociones de estimación puntual y por intervalo (i.e. intervalos de confianza) de parámetros, test de hipótesis estadísticas y su relación con los intervalos de confianza, test t para diferencias de medias, regresión simple y Análisis de Varianza. Estos temas tratados en otros cursos básicos de estadística **NO** serán cubiertos en este curso. Asimismo, los estudiantes deberán tener una familiaridad mínima (i.e. importación de datos, generación de graficos simples, etc.) con los programas R y R-Studio como para interactuar con los mismos sin mayores problemas.

## **Cupos y destinatarios:**

Este curso está principalmente destinado a estudiantes de Postgrado en Ecología, Evolución, Ciencias Ambientales y Agronomía. Se dictará con un mínimo de 5 estudiantes y habrá un cupo máximo de hasta 40 estudiantes (debido a la logística de realización de los prácticos). Los estudiantes participantes en el curso serán seleccionados en base a un CV y a una carta de motivación detallada donde se indique claramente su formación o experiencia previa en Estadística y análisis de datos.

**Costo:** A\$6000 para quienes no sean estudiantes de posgrado de la Facultad de Ciencias Exactas y Naturales de la Universidad de Buenos Aires.

**Fecha y lugar:** Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Argentina. Inicio: Lunes 17 a Viernes 21 Agosto 2022.

## **Programa del curso:**

### **Programa analítico**

1. **Introducción general.** Origen y desarrollo histórico de los marcos conceptuales para la inferencia frecuentista y bayesiana. Diferencias en la interpretación de la

probabilidad, variabilidad de los parámetros, interpretación y evaluación de precisión de estimación, uso de información previa y adicional. Conceptos de Probabilidad condicional y marginal. Teorema de Bayes y noción de causalidad inversa.

2. **Elementos básicos del análisis bayesiano.** Componentes del método bayesiano de estimación: función de verosimilitud, funciones de probabilidades previas y posteriores. Dificultades y polémica histórica acerca de la distribución previa. Principales tipos de distribuciones previas: vagas, informativas y no informativas. Principales distribuciones de probabilidad empleadas para las distribuciones previas.
3. **Análisis bayesiano I:** Métodos analíticos para obtener la distribución posterior empleando probabilidades previas conjugadas. Métodos de simulación de las funciones de probabilidades posteriores basados en la optimización estocástica global (cadenas Markovianas Monte Carlo; MCMC). Principales métodos de simulación: muestreo de Gibbs, algoritmo de Metropolis-Hastings (JAGS) y enfoque Hamiltoniano (Stan) en R. Visualización e interpretación de las distribuciones posteriores de los parámetros. Ejemplos con el Modelo Lineal General. Diagnósticos de convergencia de los algoritmos. Uso del factor de Bayes para el test de hipótesis y el análisis comparado de la evidencia de los modelos ajustados. Criterios de información (DIC y WAIC) y su uso en la selección de modelos estadísticos.
4. **Análisis bayesiano II:** Los Modelos Lineales Generalizados (GLM) como marco teórico general para el análisis de datos univariados en escalas binaria (Binomial), de conteos (Poisson y Binomial Negativa) y continua (Beta, Normal, Gama, Lognormal). Componentes de los GLM: predictor lineal de las variables explicativas, función de conexión y errores aleatorios. Ajuste de GLM con métodos bayesianos e interpretación de parámetros. Validación de modelos estadísticos a través del análisis de residuos y la simulación de las distribuciones predictivas posteriores.
5. **Modelos bayesianos jerárquicos.** Estructuras jerárquicas en la naturaleza y en los datos. Definición de efectos fijos y efectos aleatorios y su interpretación biológica. Los modelos jerárquicos bayesianos como generalización de los modelos lineales generalizados mixtos (GLMM) frecuentistas. Formulación y estimación de GLMM con métodos bayesianos e interpretación de sus parámetros. Uso de Redes bayesianas (gráficos acíclicos dirigidos; DAG) para representar dependencias entre variables aleatorias y parámetros. Uso de variables latentes para modelar la incertidumbre de observación y errores de medición en los modelos jerárquicos.

### **Evaluación:**

Los alumnos serán evaluados a través de un proyecto final de análisis de datos, sobre el cual entregarán un informe individual con la interpretación detallada de los resultados. Se aprobará el curso de acuerdo a los requerimientos académicos de la Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires. Este informe será enviado por e-mail dos semanas después de la fecha final del curso. Además de la nota del curso, los estudiantes recibirán por e-mail la corrección comentada de su informe final. Cada estudiante admitido en el curso se compromete formalmente a realizar la evaluación final del mismo.

### **Bibliografía Básica:**

Bolker, B. 2008. Ecological Models and Data in R. Princeton University Press. Princeton.  
Fletcher D. 2018. Model Averaging. Springer-Verlag. New York.  
Gelman A. et al. 2014. Bayesian data analysis, 3rd edition. Chapman & Hall/CRC. Boca Raton, USA.

- Gelman A. & J. Hill. 2007. Data Analysis Using Regression and Multilevel/Hierarchical models. Cambridge University Press.
- Hoobs N. & Hooten M. 2015. Bayesian Models a statistical primer for ecologists. Princeton University Press.
- Kery M. 2010. Introduction to WinBUGS for Ecologists. Academic Press. New York.
- Kery M. & Schaub M. 2012. Bayesian Population Analysis using WinBUGS. Elsevier, New York.
- Kruschke J. 2010. Doing Bayesian Data Analysis A Tutorial with R and BUGS. Academic Press. New York.
- Lambert J. 2018. A Student's Guide to Bayesian Statistics. Sage Publishers. London.
- McGrayne S. 2010. The Theory That Would Not Die. Yale University Press. New Haven.
- McCarthy M. 2007. Bayesian Methods for Ecology. Oxford University Press. Oxford.
- McElreath R. 2015. Statistical Rethinking: A Bayesian Course with Examples in R and Stan. Chapman & Hall/CRC. Boca Raton, USA.
- Royle J. & Dorazio R. 2008. Hierarchical Modelling and Inference in Ecology: the Analysis of Data from Populations, Metapopulations and Communities. Academic Press. New York
- Wang S. et al. 2018. Bayesian Regression Modeling with INLA. Chapman & Hall/CRC Boca Raton, USA.