



Curso Estadística Avanzada y Aplicaciones CURE Rocha

Docentes: Dr. Mathias Bourel (Laboratorio de Probabilidad y Estadística, IMERL, Facultad de Ingeniería) y Dra. Carolina Crisci (PDU Modelización y Análisis de Recursos Naturales, CURE)

Lugar: CURE, sede Rocha (miércoles salón 9, viernes salón 6)

Días y horario: viernes de 13:30 a 17:00, además habrá clases en el mismo horario los miércoles 8 y 29 de abril, 13 de mayo, 17 de junio y 8 de julio.

Fecha de inicio: miércoles 8 de abril de 2015

Fecha de finalización: viernes 10 de julio de 2015

Carga horaria total: 112 horas (45 horas Teóricos + 18 horas Prácticos + 40 horas Actividades individuales de preparación de informes + 3 horas Defensa de evaluaciones + 6 horas Seminarios)

Modalidad: presencial

Aprobación: Entrega de ejercicios prácticos y de un trabajo final de aplicación de técnicas con datos de interés para el participante. El trabajo final tiene defensa oral obligatoria.

Cursos previos: curso básico de Estadística

Solicitar inscripción escribiendo a carocrisci@gmail.com

Objetivos del curso

El curso de Estadística Avanzada y Aplicaciones está destinado a docentes y estudiantes de los diversos grupos de trabajo y carreras, en particular de posgrado, del CURE. Se tratarán temas que pretenden dar herramientas para fortalecer el análisis estadístico en trabajos de grupos de investigación, así como de estudiantes en formación.

El curso constará mayormente de exposiciones por parte de los docentes sobre técnicas estadísticas de interés para el análisis de datos. Se dará un espacio importante a la ilustración de las técnicas mediante la presentación de aplicaciones y se promoverá la

participación activa de los participantes. Habrá instancias prácticas en donde se aplicarán las técnicas expuestas utilizando el software R.

En caso de que surja la necesidad por parte de los asistentes, se dictarán clases introductorias de R de modo de brindar herramientas para la realización de los trabajos prácticos y del trabajo final así como clases de consulta cerca de la fecha de entrega de dicho trabajo.

Por último, se contará con exposiciones (seminarios) por parte de investigadores nacionales sobre algunas aplicaciones de las técnicas vistas en el curso en distintas áreas de interés de los participantes.

Programa

Módulo 1: Análisis de ordenación

Repaso de Algebra Lineal

Distribución Normal Multivariada

Análisis de Componentes Principales

Clustering

Análisis de Correspondencia

Módulo 2: Análisis clásicos de Regresión y Clasificación

Regresión lineal simple y múltiple

Modelos Lineales Generalizados - GLM

Análisis Discriminante

Módulo 3: Técnicas avanzadas de Regresión y Clasificación

Introducción al Aprendizaje Automático

Modelos Aditivos Generalizados (GAM)

Árboles de Clasificación y Regresión (CART)

Métodos de agregación de modelos: Bagging, Boosting, Stacking

Bosques Aleatorios - (Random Forest)

Máquinas de Vectores de Soporte (SVM)

Bibliografía

Bourel, M. (2012) Model aggregation methods and applications. *Memorias de trabajos de difusión científica y técnica*, Vol. 10, p. 19-32, 2012.

Crisci C, Ghattas B, Perera G (2012). A review of supervised machine learning algorithms and their applications to ecological data. *Ecological Modelling* 240: 113-122.

De'Ath, G. et al (2000). Classification and regression trees: a powerful yet simple technique for ecological data analysis.

Everitt y Torsten (2010). A Handbook of Statistical Analyses Using R, 2nd edition, Brian S., Chapman & Hall/CRC.

Hastie T., Tibshirani R and Friedman J (2011). *The Elements of Statistical Learning: Data Mining, Inference and Prediction.* (5th. Edition). Springer Seris in Statistics.

Izenman, A.J. (2008). *Modern Multivariate Statistical Techniques. Regression, Classification and Manifold Learning.* Springer Texts in Statistics.

James, G., Witten, D., Hastie, T., Tibshirani, R. (2013) An introduction to Statistical Learning with Applications in R, Springer.

Legendre, P., Legendre, L. (1998). Numerical Ecology. Elsevier Science

Peña, D. (2002). Análisis de datos multivariantes, Daniel Peña, Mac Graw Hill