

Programa Tentativo del Curso:
“Astroestadística”,
Primer Cuatrimestre, Año 2010

Parte 1: Probabilidad

Introducción a la probabilidad. Variables aleatorias discretas y continuas. Funciones distribución importantes.

Generadores de numeros (pseudo)aleatorios. Método de Inversión y rejección. Análisis básico de datos: estimadores, intervalos de confianza. Histogramas, Bootstrap Resampling. Testeo de hipótesis e (in)dependencia de datos.

Test de chi square, Kolmogorov-Smirnov, Tasa de descubrimientos Falsos, (FDR). Aplicaciones: separación señal-ruido en imágenes astronómicas profundas.

Estadística de Bayes. Priors y distribuciones posteriores. Estimación Bayesiana. Métodos de Máxima Entropía. Bases de Datos y Software: Organización de bases de datos relacionales.

Structured Query Language (SQL). Integración con software estadísticos. Optimización. Utilización de bibliotecas para correlaciones.

Observatorios Virtuales y desafíos en el era de la abundancia de datos: VISTA, DES and LSST surveys. Aplicaciones: Minería de datos en bases de datos Sloan Digital Sky Survey(SDSS) y Millennium simulations. VIsiVo y VOtables.

Parte 2: Inferencia Estadística:

Estimación de Distribuciones en Varias Variables: La distribución multivariada normal.

Estimación del vector media y la matriz de covarianza.

Las distribuciones y el uso de los coeficientes de correlación de la muestra. Clasificación de las observaciones. La distribución de la matriz de covarianza.

Tests de Independencia de Variables: Test de la hipótesis general lineal, análisis multivariado de la varianza. Test de independencia. Test de hipótesis de igualdad del vector media y de la matriz de covarianza. Aplicaciones: Uso básico de un lenguaje de programación estadístico: R, estimación de la matriz de covarianza dado un modelo teórico: funciones ventana.

Un ejemplo de inferencia: clustering. Inferencia exacta por enumeración completa. Maximum Likelihood y Clustering.

Aplicaciones: Descomposición en halos de la distribución de galaxias. Método de Karhunen-Loeve. Marginalización exacta. El método de Laplace.

Comparación de modelos. Análisis de componentes independientes.

Tópicos de inferencia random. Teoría de decisiones. Inferencia Bayesiana y teoría de muestreo.

Problemas de Inversión: El problema general de la inversión discreta. Criterios de mínimos cuadrados, de valores mínimos absolutos. Métodos Monte Carlo en inversiones. Problemas de inversión de funcionales.

Aplicaciones: Reconstrucción de estructuras tridimensionales de objetos astronómicos.

Parte 3: Modelos Observacionales:

Estimación de Modelos, Cuadrados Mínimos, Descomposición en Valores Singulares. Método de Levenberg-Marquardt

Información de Fisher y la desigualdad de Cramer-Rao. Forecasting. Transformadas: Ortogonales, Fourier y Wavelets. Arquitecturas: polinomios, funciones ortogonales. Redes Neuronales: introducción, aprendizaje (supervisado) y deconvolución.

Regularización. Aplicaciones: Estimación de redshifts fotométricos y clasificación morfológica y espectroscópica automática de galaxias.

Búsqueda Multidimensional y Optimización. Simulated Annealing. Algoritmos genéticos.

Filtrado y Estimación de Estado:

Matched Filters. Filtros de Wiener y Kalman. Aplicaciones: Reconstrucción del campo de densidad de galaxias; detección de cúmulos de galaxias y galaxias satélites de la Vía Láctea y Andrómeda.

Modelos Monte Carlo: Modelos Monte Carlo eficientes. Muestreo Monte Carlo exacto.

Métodos variacionales.

Aplicaciones: Monte Carlo Markov Chain en la práctica. Uso de COSMOS-MC (Superbayes) en estimación de parámetros cosmológicos.

Bibliografía del Curso: "Astroestadística", Primer Cuatrimestre, Año 2010

Libros:

The Nature of Mathematical Modeling, N. Gershenfeld, 2003 Cambridge University Press.

An Introduction to Multivariate Statistical Analysis, T. Anderson, 1984 John Wiley.

Information Theory, Inference and Learning Algorithms, David MacKay, 2003 Cambridge University Press.

The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Trevor Hastie, Second Edition (Springer Series in Statistics)

Numerical Recipes: The Art of Scientific Computing, Third Edition, 2007 Cambridge University Press

The Practical SQL Handbook, J. Bowman, 1996 Addison Wesley.

Foundations of Database Systems, Elmasri and Navathe, 2006 Addison Wesley

Introducción a la Minería de Datos, A. Orallo 2004. Pearsons Prentice Hall.

Bayesian Core A Practical Approach to Computational Bayesian Statistics, J. M. Marin & Ch. P. Robert. 2007, Springer-Verlag, New York.

Digital Signal Processing Handbook, V Madisetti and D. William, 1998, IEEE Press

Inverse Problem Theory, Albert Tarantola, 2005 SIAM.

Artículos y Reviews específicos, entre los más importantes podemos citar:

Statistical techniques in cosmology

Authors: Alan Heavens

arXiv:0906.0664

A practical guide to Basic Statistical Techniques for Data Analysis in Cosmology

Authors: Licia Verde

arXiv:0712.3028

Indexing the Sphere with the Hierarchical Triangular Mesh

Authors: Alexander Szalay et al

CS/0701164

Bayesian reconstruction of the cosmological large-scale structure: methodology, inverse algorithms and numerical optimization

Authors: F. Kitaura and T. Ensslin

Astro-ph/0705.0429

Sistemas de galaxias y sus halos de materia oscura,

Mariano Domínguez, 2006 Tesis Doctoral FaMAF, UNC

Se planea realizar trabajos practicos utilizando el software estadístico R (<http://www.r-project.org>), Cosmos-MC (<http://cosmologist.info/cosmomc/>) y VIsvO (<http://visivo.oact.inaf.it/>) corriendo sobre plataformas linux pertenecientes al IATE.

