

CFT 1: Aplicaciones Estadísticas al Diseño Experimental y Técnico y al Análisis de Datos

Fecha: 14-18/09/2015

Horario: 10:00-14:00

Metodología didáctica: curso teórico-práctico

Coordinadores: Jacobo de Uña, Universidade de Vigo (jacobo@uvigo.es)
Ángeles Saavedra, Universidade de Vigo (saavedra@uvigo.es)

Profesores: Jacobo de Uña (Universidade de Vigo)
Ángeles Saavedra (Universidade de Vigo)

DESCRIPCIÓN

Facilitar conocimiento y herramientas estadísticas para problemas de aplicación en ciencias marinas, tecnología y gestión.

Contenidos

Multiple linear regression:

Covariance and correlation. Simple correlation, partial correlation, multiple correlation. The multiple linear regression model. Interpretation of the model. Estimation of the model and hypothesis testing. Explained and residual variance. ANOVA table. Multiple R squared. Adjusted R squared. Residuals and residual analysis. Variable selection: extra sum of squares. Log-linear regression. Polynomial regression. Model with interaction. Prediction.

Time series:

- Time series data: Time series. Plots, trends and seasonal variation. Decomposition of series.
- Correlation: Stationary time series. Autocorrelation. The correlogram.
- Forecasting strategies.
- Stationary models: The ARMA process.
- Non stationary models: ARIMA and SARIMA processes.

Advanced regression and multivariate analysis: Regression with categorical response. Binary and multinomial logistic regression. Nonparametric regression. Principal components analysis. Cluster analysis: K-means and hierarchical clustering. Discriminant analysis: Fisher linear discriminant rule, logistic discrimination. ROC analysis in classification problems.

Spatial statistics:

- Types of spatial dependence
- Spatial correlation : Fitting a variogram model.
- Spatial prediction: The kriging methods.

Referencias

- Peña, D (2002) Análisis de Datos Multivariantes. McGraw-Hill (in Spanish)
- Applied Logistic Regression, 2nd Edition. David W. Hosmer, Stanley Lemeshow, Wiley (2000)
- Kernel smoothing, Wand and Jones, Chapman & Hall (1995)
- Analyzing Multivariate Data (Lattin J, Carroll JD, Green PE, 2003). Thomson.
- Biostatistical Analysis (Zhar, 1996). Prentice Hall.
- Estadística. Modelos y métodos. Vol. 2: Modelos lineales y series temporales (Peña, 1999). Alianza Universidad Textos (in Spanish)
- Cowpertwait, P.S.P. and Metcalfe, A.V. Introductory Time Series with R, Springer New York, 2009.
- Diggle, P.J. and Ribeiro, P.J. Model-based geostatistics, Springer New York, 2006
- Zuur, A.F.; Ieno, E.N. and Smith, G.M. Analysing ecological data, Springer New York, 2007

Distribución de la carga lectiva

Las 21 horas presenciales (habrá descansos de 30 minutos cada una de las cuatro sesiones) se dedicarán a la impartición de los contenidos, la realización de ejercicios prácticos, y la resolución de dudas. Además, se estima que el alumno deberá dedicar 27 horas más al estudio personal (consulta del material docente suministrado, con antelación al comienzo del curso y a lo largo del mismo –por la tarde, ya que el horario del curso es de mañana). El examen final tendrá una duración de 2 horas.

Sistema de evaluación

El alumno será declarado apto/no apto en función de los resultados que obtenga en la prueba teórico-práctica (examen) final. Esta prueba permitirá asimismo otorgar una nota numérica al estudiante.