WORKSHOP: YOUNG RESEARCHERS IN BIOSTATISTICS

11th of June, 2014
Aula 1, Faculty of Medicine
University of Santiago de Compostela

Scientific Committee
Carmen Cadarso-Suárez
Carmen Carollo Limeres
Xosé Luis Otero Cepeda

Organizing Committee
Ipek Guler
Roberto Domínguez Gómez
Ana Bouzas Lorenzo
Programme

10:00h – 10:30h


Arantzazu Arrospide Elgarresta
Unidad de Investigación AP-OSIs Gipuzkoa

10:30h – 11:00h

“MULTIVARIATE TECHNIQUES IN ECOLOGY: THE INFAUNA ASSOCIATED TO A CWC HABITAT (FACIES OF ISIDELLA ELONGATA), INFLUENCES OF FISHING ACTIVITY AND NATURAL VARIABILITY”

Valeria Mamouridis
Institut de Ciències del Mar (ICM-CSIC)

11:00h – 11:30h

“EFFECTS OF DEEP-SEA FISHERIES ON THE DYNAMICS OF BATHYAL FOOD WEBS (NW MEDITERRANEAN)”

Valeria Mamouridis
Institut de Ciències del Mar (ICM-CSIC).
11:30h – 12:00h

“BIVARIATE GAUSSIAN DISTRIBUTIONAL REGRESSION: AN APPLICATION ON DIABETES”

Nadja Klein
Chair of Statistics
Georg-August-University Göttingen

12:00h – 12:30h

“USING SCREENING DATA IN THE STUDY OF BREAST CANCER RISK IN CENTRAL PORTUGAL”

Elisa Duarte
Unit of Biostatistics
Faculty of Medicine, University of Santiago de Compostela

12:30h – 13:00h

“IMPUTATION OF MISSING DATA IN ACCELERATED FAILURE TIME (AFT) MODELS”

Moisés Castro Cacabelos
Clinical Epidemiology Unit
Hospital Clinico Santiago de Compostela, Spain

13:00h – 13:30h

“COMPARING THE PREDICTIVE PERFORMANCE OF DIFFERENT REGRESSION MODELS FOR LONGITUDINAL AND TIME-TO-EVENT DATA”

Ipek Guler
Unit of Biostatistics
Faculty of Medicine, University of Santiago de Compostela
In the Basque Country (Spain), mammographies have been done in biennial basis to women in their fifties and sixties since 1996. The main objective of this project was the evaluation of the impact of the Screening program in terms of health benefit and harms in the Basque women population since 1996.

A discrete event simulation model was built to represent the natural history of breast cancer in women invited to the breast cancer screening program in the Basque Country during the period 1996-2011. The disease progress was described in three main states (healthy, preclinical and clinical). We assumed that all women would be diagnosed at the beginning of the clinical stage unless they had been diagnosed previously through the screening program. Data collected during the 15 years of the screening program allowed the model’s calibration and validation. Age-specific incidence and breast cancer mortality rates were calculated for the screening and non-screening scenarios. The number of false positive results and over diagnosed cancers were calculated to assess screening harms.

Since the program started working, 5,267 cancers were detected by screening among 341,880 women who attended the screening which represents the 83% of the invited ones. Among 13,477 women submitted to the reference hospital for additional tests 39.1% were diagnosed breast cancer. One out of five screen-detected cancers was over diagnosed. Focusing on year 2011, breast cancer incidence in the screened population increased more than a 20% for women aged 50-69 and decreased for those aged 70 in comparison with the non-screened scenario. Overall breast cancer mortality rates, on the other hand, decreased a 16.3% by year 2011.

Fifteen years after the screening program started our results support an important decrease in breast cancer mortality with reasonable harms.

KEYWORDS: breast cancer, screening, discrete event simulation.
Deep-sea communities associated to the so-called cold water corals (CWC), show unique complexity and species diversity, because corals act as potential areas of foraging, refuge and breeding for many deep-sea species, enhancing their ecological dimensions. In early ‘90, the coral Isidella elongata formed pristine forests on soft sediments in the north-western Mediterranean but from 1996 the habitat has been subjected to fishing activity.

During the present study we found poor density colonies of the coral (decreased from 255 colonies/ha to 0.9 colonies/ha). We analysed infauna samples from two different habitats, mud and coral mud habitats, through multivariate analysis. We also investigated possible mesoscale gradients related to environmental variability. Despite the start of the fishing, we found that samples were segregated mainly according to the habitat, but also high variability within the same habitat has been observed. The coral mud habitat showed higher heterogeneity of species composition and biomasses and was dominated by detritivores (mainly crustaceans), while in mud habitat we found higher importance of carnivorous than in coral mud habitat. The higher quality of organic source in coral mud habitat and other favorable environmental variables with respect to the mud habitat permit to identify different infaunal assemblages.
We explored the dynamics of a food web (who eats whom) in the bathyal ecosystem of the NW Mediterranean at 600-650m depth using real data from 8 cruises performed during years 2007 and 2011. This ecosystem is historically subjected to the fishery of red shrimp. We considered two inputs of carbon: the organic matter from the vertical fall through the water column and from the advective transport through the slope, and four outputs (loss of carbon): the burial process in sediments, the dissipation through respiration, and the loss due to the fishing activity on target (red shrimp) and on no-target species (fish and few invertebrates, called by-catch). Internal components of the food web are: the total organic matter in sediment, the macrobenthos, the zooplankton-micronekton and the suprabenthos, that rely directly on the organic matter in sediments and inputs, and the megafauna components (megabenthos, megaichthyofauna and the red shrimp), that prey on previous compartments. The most of the carbon flows through the macrobenthos, that play a key structural role in the community. The actual fishing effort and its low changes makes components to oscillate around their mean values, while its removal affects substantially the megabenthos (big invertebrates), that shows an initial decrement of its biomass and a recovery after long term (20 years). Changes in fishing activity also produce alternations of top predators' biomass as a result of competition.
So far, different markers reflecting the blood sugar levels are analysed separately and used for the diagnosis and control of diabetes. It is known that such markers (glycated proteins) have a high correlation which is expected to be dependent on clinical or biochemical factors. Since neglecting correlations between two proteins or heteroscedasticity of the distribution may lead to misclassification of patients, we aim in identifying important covariates on the variance and correlation parameters with the help of structured additive distributional regression for multivariate responses.
USING SCREENING DATA IN THE STUDY OF BREAST CANCER RISK IN CENTRAL PORTUGAL

Elisa Duarte¹, Bruno de Sousa², Carmen Cadarso-Suarez¹, Vitor Rodrigues³, and Thomas Kneib⁴

1. Unit of Biostatistics, Department of Statistics and Operations Research, School of Medicine, University of Santiago de Compostela, C/ San Francisco s/n, 15782-Santiago de Compostela, Spain
2. Faculty of Psychology and Education Sciences, University of Coimbra, Rua do Colégio Novo, Apartado 6153, 3001-802 Coimbra, Portugal
3. Faculty of Medicine, University of Coimbra, Rua Larga, 3004-504 Coimbra, Portugal
4. Chair of Statistics, Georg-August-Universität Göttingen, Platz der Göttinger Sieben 5, 37073 Göttingen, Germany

The screening data can be seen as a repository from which information can be extracted for spatial-temporal studies of variables interpreted as potential risk factors. When the information of the cancer diagnostic is provided, the screening data is useful in the exploration of the relationships between variables identified as risk factors and the probability of having the disease.

The present study aims to analyze these relationships, from the data contained in the 278282 records of women who are part of the Breast Cancer Screening Program in central Portugal, provided by the Portuguese Cancer League (LPCC).

The study combines variables that are intrinsic and extrinsic to those women who are part of the screening program. The former variables include year of birth, breast cancer family history, age of menarche and menopause, and other reproductive factors. The latter variables consider socioeconomic and geographical factors as part of the model.

Structured additive regression models were used in order to combine this wide range of covariates and to simultaneously explore possible spatial correlations.

Data shows that women born between 1940 and 1945 are associated with higher probabilities of developing breast cancer, with family history contributing positively to this probability. As for spatial effects, there is a marked increase in breast cancer risk along the east-west direction. Also of note is the contribution of early menarche and late menopause ages to the increased risk of the disease. This result concurs with the premise that women with longer fertility periods due to an early menarche and a late menopause may have a higher risk of breast cancer.
El principal problema que vamos a tratar aquí es el de datos faltantes, que surge con mucha frecuencia en la realidad, por ejemplo, en el campo de la investigación médica. Veremos los distintos métodos existentes para tratarlos y luego, teniendo en cuenta que nos interesa hacer un análisis de supervivencia, veremos cuál es el mejor. Para ello, haremos un estudio de simulación y veremos con cuál obtenemos mejores resultados.
Many follow-up studies produce different types of outcomes and commonly both longitudinal measurements and time to event are of interest to study the association between them. Beside some alternative methods like an extending Cox model with longitudinal covariates, and two stage models, joint modeling approaches with shared random effects have become very popular because of several advantages of the latter method (Rizopoulos, 2010; Phillipson et al, 2012).

This paper compares the predictive accuracy performances of different model approaches to study longitudinal and time to event processes together, with focus on survival processes. The predictive performance of the models assessed through use of time dependent ROC curves (Heagerty et al, 2005). Statistical approaches were applied on a biomedical database.