

## V Seminario de Inferencia Estadística en Áreas Pequeñas

Viernes 7 de julio de 2017.

Aula de Grados Víctor Gulías de la Facultad de Informática, A Coruña.

### PROGRAMA:

10:00-11:00h. **A Unified Approach to Goodness-of-fit Tests with Application to Small Area Estimation.** *Jiming Jiang, University of California, Davis, USA*

Abstract: We develop a method originally proposed by R. A. Fisher into a general procedure, called tailoring, for deriving goodness-of-fit tests that are guaranteed to have a chi square asymptotic null distribution. We apply the method to small area estimation for detecting potential model misspecification. Three tests are proposed using the tailoring method. We evaluate performance of the tests both theoretically and empirically. Our empirical results suggest that the proposed tests are more accurate in size and have higher power than existing tests.

This work is joint with Mahmoud Torabi of the University of Manitoba, Canada

11:00-12:00h. **Classified Mixed Logistic Model Prediction.** *Thuan Nguyen, Oregon Health & Science University, USA*

Abstract: We develop a classified mixed model prediction (CMMP) method for clustered binary data by extending a method proposed by Jiang *et al.* (2016) for continuous outcome data. By identifying a class, or cluster, that the new observations belong to, we are able to improve the prediction accuracy of probability associated with a mixed effect as well as binary outcome of future observation over the traditional method of logistic regression (LR). Furthermore, we propose an improvement of CMMP by incorporating covariate information in match the class. This allows the new method, called classified mixed logistic model prediction (CMLMP), to outperform the traditional mixed logistic model prediction without matching the class. We prove consistency of CMLMP method, and demonstrate finite-sample performance of CMLMP via simulation studies. An application to medical data is considered.

This work is joint with Hanmei Sun and Yihui Luan of Shandong University, China, and Jiming Jiang of the University of California, Davis, USA.

11:30-12:00h. Coffe-Break

12:30-13:30h. **Generalized Akaike information to estimate the employed people by economic activity.** *Esther López-Vizcaino, Instituto Galego de Estatística, Spain.*

Abstract: The question of model selection has received a lot of attention in the literature in the past and also in recent years due to the increasing complexity of modeling approaches. However, in small area estimation (SAE) this is still a problem that is tentatively studied. In this work, we introduce and validate a Mixed Generalized Akaike Information Criterion, xGAIC. The novelty of the proposal is twofold, on one side, xGAIC is derived using a mixed log-likelihood that focus in the random effect and the

variability between the areas, and, on the other side, the penalty is a GDF measure, which is calculated by bootstrap.

To study the performance of xGAIC, we consider three popular mixed models in small area inference: a Fay-Herriot model, a monotone model and a penalized spline model. A simulation study shows the good performance of xGAIC in comparison with a more classical AIC based on a conditional likelihood.

Also, a Small Area-specific estimation approach that borrows strength across areas and across time is presented to obtain Labor Force Estimators by economic activity. Often, for a given area and model, the time trend pattern is more abrupt than desired. On the other side, when time series models are considered too smoothed patterns are generated and prevent major trend changes being detected. The unequal behavior of employment in different economic activities regarding to the sample variability, the relationship with the auxiliary and the temporal evolution, give us the idea of considering different small area estimation approaches in different activities, even more after concluding, from observing the temporal patterns, that neither the direct estimators nor the standard small area model-based estimators, behave well homogeneously across activities.

We propose an approach based on obtaining the Aggregated Mixed Generalized Akaike Information Criterion statistic across time; and then using, for each area, the corresponding component.

This work is joint with M.J. Lombardía of Universidade da Coruña, Spain, and C. Rueda of Universidad de Valladolid, Spain.