

Bilbao, December 8, 2017

Josu NAJERA ZULOAGA will defend his doctoral thesis

Josu NAJERA ZULOAGA joined [BCAM-Basque Center of Applied Mathematics](#) as a PhD Student (Predoc Severo Ochoa 2014) within the [Applied Statistic](#) Research Line in 2015, coming from the UPV/EHU-University of the Basque Country, where he obtained his bachelor degree in Mathematics and a master degree at the joint programme by UPV/EHU, University of Zaragoza and University of Oviedo. His research is supported by the Spanish Ministry of Economy and Competitiveness MINECO: BCAM Severo Ochoa accreditation SEV-2013-0323, grant SVP-2014-068806 since 2014

His PhD thesis has been directed by [Inmaculada AROSTEGUI](#) and [Dae-Jin LEE](#), [BCAM](#) Researchers.

The defense of his doctoral thesis:

Hierarchical modelling of patient-reported outcomes data based on the beta-binomial distribution

will take place on **December 13, 2017, 12:00h**, Salón de Grados, Science and Technology Faculty ([UPV/EHU Leioa, Basque Country, Spain](#))

Abstract:

The beta-binomial distribution does not belong to the exponential family and, hence classical regression techniques cannot be used when dealing with outcomes following the mentioned distribution. In this thesis we propose and develop regression models based on the beta-binomial distribution for the analysis of U, J or inverse J-shaped discrete and bounded outcomes. In fact, although this thesis is focused on the analysis of patient-reported outcomes (PROs), which usually follow the mentioned distributional shapes, proposed models can also be extended to several fields.

First of all, we make a review and comparison of existing beta-binomial regression approaches in independent data context, concluding that the marginal approach is the most adequate. However, PRO studies are usually carried out in a longitudinal framework, where patients' responses are measured over time. This leads to a multilevel or correlated data structure and consequently, we extend the marginal beta-binomial regression approach to the inclusion of random effects to accommodate the hierarchical structure of the data. We develop the estimation and inference procedure for the model proposal. Furthermore, we compare the performance of our proposal with similar approaches in the literature, showing that it gets better results in terms of reducing the bias of the estimates. We apply the model to a longitudinal Chronic Obstructive Pulmonary Disease study carried out at Galdakao Hospital, reaching clinically and statistically relevant results about the evolution of the patients over time.

PROs are usually obtained using rating scale questionnaires consisting of questions or items, grouped into one or more subscales, often called dimensions or domains. Therefore, we also propose a multivariate regression model based on the beta-binomial distribution for the joint analysis of all the longitudinal dimensions provided by different questionnaires. PROs are usually obtained using rating scale questionnaires consisting of questions or items, grouped into one or more subscales, often called dimensions or domains. Therefore, we also propose a multivariate regression model based on the beta-binomial distribution for the joint analysis of all the longitudinal dimensions provided by different questionnaires.

Finally, it is worth mentioning that we have implemented all the proposed regression models in the PROreg R-package which is freely available at CRAN.

Josu NAJERA ZULOAGA, defiende su tesis doctoral

Josu NAJERA ZULOAGA se unió a [BCAM-Centro Vasco de Matemática Aplicada](#) como estudiante de doctorado de la Línea de Investigación [Applied Statistic](#) en 2015, procedente de la UPV/EHU Universidad del País Vasco / Euskal Herriko Unibertsitatea, donde obtuvo su diploma de licenciatura y maestría. Ha sido becado por el MINECO-Ministerio de Economía y Competitividad: BCAM-Acreditación Severo Ochoa SEV-2013-0323, grant SVP-2014-068458.

Su tesis de doctorado ha sido dirigida por [Inmaculada AROSTEGUI](#) y [Dae-Jin LEE](#), Investigadores de [BCAM](#).

La defensa de su tesis doctoral:

Modelización jerárquica de datos de respuestas percibidas por el paciente basada en la distribución beta-binomial

será el **13 de Diciembre de 2017, a las 12:00h**, Salón de Grados, de la Facultad de Ciencia y Tecnología ([UPV/EHU Leioa, País Vasco, Spain](#))

Resumen:

La distribución beta-binomial no pertenece a la familia exponencial y, por lo tanto, técnicas clásicas de regresión no deben ser utilizadas cuando procedemos al análisis de datos que siguen la mencionada distribución. En esta tesis, proponemos y desarrollamos modelos de regresión basados en la distribución beta-binomial para el análisis de respuestas discretas y acotadas que muestran formas de U, J o J-invertida. De hecho, aunque esta tesis esté enfocada al análisis de respuestas percibidas por el paciente (RPP), las cuales suelen mostrar este tipo de distribuciones, los modelos propuestos pueden ser útiles en diferentes ámbitos.

Para empezar, realizamos una revisión y estudio de comparación de dos enfoques diferentes para el desarrollo de regresiones basadas en la distribución beta-binomial para el análisis de datos independientes. Sin embargo, estudios de RPP son usualmente llevadas a cabo en un entorno longitudinal donde, a lo largo del tiempo, se realizan diferentes mediciones de cada paciente. Estas repetidas observaciones constituyen

una estructura de correlación, por lo que extendemos el óptimo modelo de regresión basado en la distribución beta-binomial para el análisis de datos correlados. Así mismo, comparamos idoneidad del modelo de regresión que proponemos con enfoques similar existentes en la literatura, mostrando que nuestra propuesta obtiene mejores resultados reduciendo el sesgo de las estimaciones. Adicionalmente, aplicamos nuestro modelo a un estudio longitudinal llevado a cabo en el Hospital de Galdakao donde la calidad de vida de pacientes con Enfermedad Pulmonar Obstructiva Crónica (EPOC) es medida a lo largo del tiempo. Concluimos en resultados de relevancia clínica sobre factores de riesgo y la evolución de los pacientes.

Las RPP son amenudo medidas por medio de cuestionarios que se componen de preguntas escaladas que son usualmente agrupadas subescalas dando lugar a diferentes dimensiones. Por lo tanto, en este trabajo proponemos así mismo un modelo de regresión basado en la distribución beta-binomial para el análisis conjunto de diferentes dimensiones proporcionados por el mismo cuestionario.

Por último, cabe destacar que toda la metodología propuesta y desarrollada en esta tesis ha sido implementada en un paquete del software estadístico R, el cual está libremente disponible en CRAN.