

3rd *BIOSTATNET* WORKSHOP

WOMEN AND BIostatISTICS

This workshop is dedicated to **María Eugenia Castellanos Nueda**, coordinator of the Valencia–VABAR Node of the BIostatNET network.

María Eugenia, was an outstanding researcher and colleague. She combined intellectual clarity and remarkable dedication with humility and generosity. Her research focused on the development and application of Bayesian methods for data modelling, ranging from the quantification of model uncertainty to advanced methodologies for complex data. Throughout her career, she published extensively through national and international collaborations, making significant contributions in biostatistics. In all aspects of her professional and personal life, she left an enduring mark, remembered with admiration and affection.

November 28, 2025

Madrid (also online)

Facultad de Estudios Estadísticos,
Universidad Complutense de Madrid
Avda. Puerta de Hierro s/n,
Ciudad Universitaria, 28040 Madrid

Inscription:

<https://rb.gy/n8tv8t>



PROGRAMME

Morning Session

09:30–10:00 Opening of the Workshop. BIOSTATNET.

10:00–10:50 Modelos bayesianos en investigación biomédica: una trayectoria compartida con María Eugenia Castellanos. **Alicia Quirós Carretero** (Valencia–VABAR BIOSTATNET Node). Department of Mathematics, Universidad de León.

10:50–11:10 Bayesian Latent Variable Selection Approach. Building the Grouped Bayes Factor. **Carolina Mulet** (Valencia–VABAR BIOSTATNET Node). Department of Informatics and Statistics, Universidad Rey Juan Carlos.

11:10–11:30 Spatio-temporal modeling for record-breaking events: An application to temperature daily series. **Ana C. Cebrián** (Zaragoza BIOSTATNET Node). Department of Statistical Methods, Universidad de Zaragoza.

11:30–12:00 Coffee break

12:00–12:20 How smoothing works in multivariate disease mapping models. **Jaione Etxeberria** (Navarra Spatial Statistics BIOSTATNET Node). Department of Statistics, Computer Science, and Mathematics, Universidad Pública de Navarra.

12:20–12:40 Exploring the association between underreported mental disorders and atmospheric pollution. **Leire Unamuno** (Catalunya–SEA BIOSTATNET Node). Centre de Recerca Matemàtica, Barcelona (CRM).

12:40–13:00 Fair models in Health and Biostatistics: strategies to mitigate bias. **Sandra Benítez Peña** (Madrid Semipar BIOSTATNET Node). Department of Statistics, Universidad Carlos III de Madrid.

13:00–13:20 A novel Bayesian method for responsible logistic regression. **Pepa Ramírez Cobo** (Andalucía Occidental BIOSTATNET Node). Department of Statistics, Universidad de Cádiz.

13:20–13:40 Matemáticas en la inoculación de la viruela. **M^a Jesús López Herrero** (Madrid UCM BIOSTATNET Node). Department of Statistics and Data Science, Universidad Complutense de Madrid.

13:40–14:00 Prediction models for complex survey data: application to health records. **Inmaculada Arostegui** (País Vasco BIOSTATNET Node). Department of Mathematics, UPV/EHU and Basque Center for Applied Mathematics, BCAM.

14:00–16:30 Lunch

Afternoon Session

16:30–17:00 **María Eugenia Castellanos: Pasión por el conocimiento y la incertidumbre.**
Carmen Armero and Anabel Forte Deltell (Valencia–VABAR BIOSTATNET Node). Departament d'Estadística i Investigació Operativa, Universitat de València.

17:00 – 18:00 (approx.)

Short tributes and interventions in memory of María Eugenia Castellanos:

- **Asunción M. Mayoral and Javier Morales Socuellanos**
- **Gonzalo García Donato**
- **Abraham Duarte**
- **M^a José Nueda Roldán**
- **María Isabel González Vasco**

SPEAKER 1

Alicia Quirós Carretero (Valencia-VABAR BIOSTATNET Node). Department of Mathematics, Universidad de León.

Title

Modelos bayesianos en investigación biomédica: una trayectoria compartida con María Eugenia Castellanos.

Abstract

En esta charla presento una síntesis de algunos trabajos realizados en colaboración con la profesora María Eugenia Castellanos, centrados en el desarrollo y aplicación de metodologías bayesianas en el ámbito biomédico. A lo largo de nuestra colaboración, que comenzó en 2009, realizamos algunos trabajos aplicados, de los que he seleccionado tres: un análisis bayesiano de supervivencia en pacientes con cáncer de pulmón, un diseño adaptativo para un ensayo clínico y una metodología para selección de variables con datos perdidos y su aplicación a datos de cardiología. Más allá de su excelencia académica, esta presentación pretende también poner en valor el legado científico y humano de María Eugenia Castellanos, cuya contribución a la estadística bayesiana en contextos aplicados fue tan sólida como generosa.

Biographical Sketch

Alicia Quirós Carretero es Profesora Titular y Subdirectora del Departamento de Matemáticas de la Universidad de León. Licenciada en Matemáticas por la Universidad Autónoma de Madrid y doctora en Estadística por la Universidad Rey Juan Carlos, inició su carrera investigadora en 2003 y desde entonces se ha especializado en la aplicación de métodos estadísticos al análisis médico, con colaboraciones en neuroimagen, oncología y cardiología. Es miembro del nodo Valencia de BIOSTATNET desde 2014 y ha sido colaboradora de María Eugenia Castellanos desde 2009. Desde junio de 2025, es coordinadora del Grupo de Trabajo en Bioestadística de la Sociedad Española de Estadística e Investigación Operativa. En marzo de 2025 se convirtió en la primera mujer española en obtener la Acreditación Estadística Europea, un reconocimiento de excelencia otorgado por la Federación Europea de Sociedades de Estadística. Ha desempeñado su labor docente en múltiples instituciones, incluyendo la Universidad Rey Juan Carlos, la Universidad de Castilla - La Mancha, la Universidad de Valencia, la Universidad de Cagliari (Italia) y la Universidad Carlos III de Madrid.

SPEAKER 2

Carolina Mulet (Valencia-VABAR BIOSTATNET Node). Department of Informatics and Statistics, Universidad Rey Juan Carlos.

Title

Bayesian Latent Variable Selection Approach Building the Grouped Bayes Factor.

Abstract

Under the Model Uncertainty framework, we are often unsure about which model is the "true" one that generated the data. The Bayesian perspective allows us to quantify how probable each subset of variables is and draws conclusions to incorporate that model uncertainty in the decision making.

Usually, in applied studies, we are interested in hypothetical constructs rather than in measurable variables while explaining the response of interest. These are commonly referred to as latent variables, and we can reconstruct their idiosyncratic behaviour through means of several observable variables or proxies. In this variable selection problem, we study how to measure the importance of the latent variable. In particular, we emphasise the necessity of considering a hierarchical prior over the model space, which derives in a weighted average of the usual Bayes factors, the grouped Bayes factor. Moreover, we discuss the role of the model prior probability taking into account the relation among the observed indicators, improving the estimated evidence given by the data available.

Finally, we show the importance of considering the structure of the latent variable in a real problem. We compare the standard and the latent variable approach in the variable selection procedure of physical and lung functions while explaining the Maximum Inspiratory Pressure (MIP).

Joint work with María Eugenia Castellanos and Gonzalo García-Donato.

Biographical Sketch

Carolina Mulet holds a degree in Mathematics and a Master's in Biostatistics from the University of Valencia. She is currently a scientific-technical staff member at Universidad Rey Juan Carlos, where she focuses on Bayesian approaches to variable selection within the AI4DDS project (Artificial Intelligence for Data-Driven Solutions). She has previous research experience at ISPA and the University of Valencia, which led her to begin her PhD in 2024 under the supervision of Gonzalo García-Donato and María Eugenia Castellanos.

SPEAKER 3

Ana C. Cebrián (Zaragoza BIOSTATNET Node). Department of Statistical Methods, Universidad de Zaragoza.

Title

Spatio-temporal modeling for record-breaking events: An application to temperature daily series.

Abstract

The occurrence of record-breaking temperature events is one of the evidences of climate change. In this context, we present a Bayesian framework to investigate both the occurrence and the magnitude of record-breaking temperatures across years for any given day of the year, within a space-time setting. To this end, we develop two models: one for the occurrence of records and another for the record values, conditional on a record-breaking event having occurred. For the occurrence model, we propose a hierarchical conditional specification for the indicators, incorporating explicit trends, necessary autoregressive terms, additional covariates, and both spatial and strong daily random effects. For the record values, rather than modeling them directly, we model the increments relative to the previous record, conditional on the occurrence of a record-breaking event. This approach allows us to treat increments as conditionally independent and to avoid the order constraint inherent in record values. The analysis is based on sixty-one years of daily maximum temperature data across Spain. The fitted models provide evidence of the impact of climate change on record-breaking temperatures, and quantify them identifying spatial and seasonal differences.

Biographical Sketch

Ana C. Cebrián holds a Bachelor's and a Ph.D. in Mathematics from the University of Zaragoza. After completing a postdoctoral stay at the Catholic University of Leuven (Belgium), she obtained a position as Associate Professor in the Department of Statistical Methods at the University of Zaragoza. Her main line of research focuses on the modeling and development of statistical inference tools applied to the analysis of climate data. In particular, she works on the study of extreme events using different approaches (exceedances over threshold, upper-tail quantiles of a distribution, and records), with special emphasis on phenomena such as droughts and heat waves. More recently, she has specialized in the space-time analysis of extreme events within Bayesian models, in addition to implementing the developed methodology in open-source R libraries. Her research has been supported by multiple national and regional projects, in several of which she has served as principal investigator.

SPEAKER 4

Jaione Etxeberria (Navarra Spatial Statistics BIOSTATNET Node). Department of Statistics, Computer Science, and Mathematics, Universidad Pública de Navarra.

Title

How smoothing works in multivariate disease mapping models.

Abstract

Understanding the degree of smoothing induced by models in disease mapping is crucial as insufficient smoothing results in noisy maps dominated by random variation, while excessive smoothing can obscure the detection of both high- and low-risk areas. In a previous research Retegui et al., (2025), we quantified the degree of smoothing in univariate spatial models, showing that it is influenced by both the choice of spatial prior and the underlying spatial structure, particularly as the number of regions increases.

Multivariate models continue to play a crucial role in disease mapping, as they enable the joint analysis of multiple diseases. Although many approaches have been proposed in this field, the coregionalization framework remains popular for modeling dependence among diseases in spatial mapping. However, it can become computationally intensive as the number of diseases increases. To address this challenge, Botella-Rocamora et al. (2015) introduced M-models, which reduce computational burden by sacrificing some parameter identifiability.

This study presents some novel metrics to quantify smoothing in M-models. In particular, a theoretical measure based on conditional generalized variance along with empirical methods to assess disease-specific smoothing. We evaluate these metrics using simulated and real datasets.

Joint work with G Retegui, A Gelfand and MD Ugarte.

Biographical Sketch

Dr. Jaione Etxeberria holds a degree in Mathematics from the University of the Basque Country (2006) and earned her Ph.D. from the Public University of Navarre in 2012 (with an outstanding doctoral award), specializing in Statistics. Currently, she serves as an Associate Professor in the Department of Statistics, Computer Science, and Mathematics at the Public University of Navarre. Her research is focused on disease mapping, where she has developed various spatial and spatio-temporal methods with applications in epidemiology. More information in <https://www.unavarra.es/pdi/?uid=7857&dato=descripcion&languageId=2>

SPEAKER 5

Leire Unamuno (Catalunya-SEA BIOSTATNET Node). Centre de Recerca Matemàtica, Barcelona (CRM).

Title

Exploring the association between underreported mental disorders and atmospheric pollution.

Abstract

Growing evidence shows a significant association between exposure to pollutants such as ozone (O_3), nitrogen dioxide (NO_2), and particulate matter (PM_1 , $PM_{2.5}$, PM_{10}) and mental health problems, especially in children and adolescents. Although emerging research suggests a similar impact in adults, evidence remains less conclusive due to limitations in registries and exposure data. These findings highlight the need for further research and stronger public health measures, particularly in urban areas with high pollution levels.

We model the incidence of mental disorders in Catalonia between 2017 and 2022 across 379 health areas. The true incidence is assumed to follow a Gaussian distribution, but the observed data are subject to underreporting, represented through a Gaussian mixture model with parameters for both the probability and intensity of underreporting. Each observation is modeled using fixed covariates and smooth functions that capture non-linear, spatial and temporal effects.

For inference, we apply the Integrated Nested Laplace Approximation (INLA), a computationally efficient alternative to MCMC for latent Gaussian models, providing accurate results at lower cost. Moreover, we will previously conduct a Compositional Data Analysis for pollutants.

Joint work with Amanda Fernández-Fontelo, Pedro Puig and David Moríña.

Biographical Sketch

Leire Unamuno Celayeta es graduada en Matemáticas por la Universidad del País Vasco (UPV-EHU) y máster en Estadística e Investigación Operativa por la Universidad Politécnica de Catalunya y la Universidad de Barcelona (UPC-UB). Actualmente es técnica de investigación (research technician) en el Centre de Recerca Matemàtica (CRM) y profesora asociada en la Universidad Autònoma de Barcelona (UAB).

SPEAKER 6

Sandra Benítez Peña (Madrid Semipar BIOSTATNET Node). Department of Statistics, Universidad Carlos III de Madrid.

Title

Fair models in Health and Biostatistics: strategies to mitigate bias.

Abstract

The rise of machine learning and data science has brought a critical challenge to the forefront: model fairness. In this talk, we will explore what it means to evaluate and mitigate biases in algorithms, review the available metrics and methods, and present examples of applications in biomedicine and biostatistics, where algorithmic fairness is crucial for obtaining more reliable and responsible conclusions. We will also introduce new approaches aimed at advancing the development of fairer and more robust models.

Biographical Sketch

Sandra Benítez is an Assistant Professor in the Department of Statistics at UC3M and holds a PhD in Mathematics from the University of Seville. Her research focuses on developing machine learning models that are fairer, more interpretable, and efficient. She has published many articles in JCR-indexed journals and has participated in numerous research projects and contracts. In addition, she actively engages in scientific outreach initiatives.

SPEAKER 7

Pepa Ramírez Cobo (Andalucía Occidental BIOSTATNET Node). Department of Statistics, Universidad de Cádiz.

Title

A novel Bayesian method for responsible logistic regression.

Abstract

In this talk, we introduce a novel variant of the logistic regression model—specifically designed to produce solutions that are fair and interpretable, key concerns in the domain of responsible machine learning. The adopted framework is the mean-field variational inference and both fairness and sparsity shall be controlled through the prior distributions. Numerical experiments on synthetic and real datasets demonstrate the competitiveness of our approaches in terms of fairness, sparsity, and predictive accuracy.

Joint work with R. Jimenez-Llamas and E. Carrizosa.

Biographical Sketch

Pepa Ramírez Cobo is a faculty member in the Department of Statistics at the University of Cádiz (Spain), where she currently holds the position of full professor. Her research interests include Statistical Learning, Bayesian methods, and the statistical estimation of stochastic processes. She is the co-author of 38 research articles published in high-impact international journals, such as *Biostatistics*, *Bayesian Analysis*, *European Journal of Operational Research*, *Reliability Engineering and System Safety*, *Computers and Operations Research*, *Annals of Applied Statistics*, and the *Journal of Applied Statistics*, among others. Her work has led to significant contributions in the areas of statistical learning—particularly in sparse, cost-sensitive, and fair learning methods—and in statistical inference for stochastic processes, both from Bayesian (MCMC- or ABC-based) and optimization perspectives. In terms of mentorship, she has co-supervised 8 PhD dissertations. Regarding her participation in research projects, she has been involved in 24 national and international projects.

SPEAKER 8

M^a Jesús López Herrero (Madrid UCM BIOSTATNET Node). Department of Statistics and Data Science, Universidad Complutense de Madrid.

Title

Matemáticas en la inoculación de la viruela.

Abstract

Las epidemias son tan antiguas como la humanidad, pero su modelización matemática es mucho más reciente. Hasta mediados del siglo XVIII, no aparece una publicación documentada en la que se utiliza un modelo para representar la expansión de una epidemia, en este caso de viruela.

En esta charla hablaremos de epidemias, dioses, castigos, héroes, y, como no, de los avances terapéuticos y científicos que contribuyeron al desarrollo de modelos matemáticos con los que predecir, y controlar, en lo posible, la expansión de un patógeno en un grupo de individuos.

Biographical Sketch

M^a Jesús López Herrero es profesora de la Universidad Complutense de Madrid, donde ha trabajado durante más de 35 años. Su investigación está orientada al estudio y modelización de sistemas biológicos y de telecomunicaciones desde un punto de vista estocástico. El marco teórico se encuadra dentro de los procesos Markovianos (tanto en tiempo discreto como continuo) y la metodología empleada se basa en resultados analítico-matriciales. Forma parte del grupo de investigación UCM "Modelos estocásticos", que ha recibido financiación procedente de programas competitivos nacionales e internacionales. Es autora de más de 40 artículos de investigación referenciados y publicados en revistas de reconocido prestigio.

SPEAKER 9

Inmaculada Arostegui (País Vasco BIOSTATNET Node). Department of Mathematics, UPV/EHU and Basque Center for Applied Mathematics, BCAM.

Title

Prediction models for complex survey data: application to health records.

Abstract

Complex survey data, common in social and health sciences, use stratified and multistage designs with sampling weights. Since standard statistical methods assume simple random sampling, applying them directly to such data often leads to biased results and requires methodological adaptations.

Our work investigated logistic regression models for binary outcomes under complex survey designs. Regarding parameter estimation, a simulation study comparing alternative estimation strategies supported the inclusion of sampling weights to improve parameter accuracy. Model selection procedures were addressed through LASSO regression. A novel design-based cross-validation method was proposed, which accounts for complex sampling features. Simulation results demonstrated substantial improvements over traditional cross-validation. Predictive performance was evaluated using the area under the receiver operating characteristic curve. Estimators that incorporate sampling weights were introduced to obtain unbiased discrimination measures and optimal classification cut-off points. The implementation as R packages provide researchers with practical tools.

These results could be useful beyond survey data. For instance, applying complex sampling-designs to health data that are systematically registered.

Joint work with Amaia Iparragirre and Irantzu Barrio.

Biographical Sketch

Inmaculada Arostegui has a degree and a PhD in Mathematical Sciences from the UPV/EHU, and Master's degree in Health Sciences in Biostatistics from the Johns Hopkins University. She is full professor in the Department of Mathematics in the UPV/EHU, statistical advisor at Galdakao University Hospital and research collaborator at the Basque Center for Applied Mathematics – BCAM. Her research focuses on the development, validation and implementation of statistical methodology. Her work emphasizes interdisciplinarity and transactional research, in particular social impact and patients' benefit, with a commitment to facilitating the results of clinical research to be transferred to clinical practice.

SPEAKERS 10 and 11

Carmen Armero and Anabel Forte Deltell (Valencia-VABAR BIOSTATNET Node).
Departament d'Estadística i Investigació Operativa, Universitat de València.

Title

María Eugenia Castellanos: Pasión por el conocimiento y la incertidumbre.

Abstract

Con esta comunicación queremos homenajear a nuestra querida amiga y compañera María Eugenia Castellanos. Recorreremos algunos aspectos de su trayectoria científica y personal. Visitaremos Villarrobledo, ahí empezó todo, València, Elche, Pittsburgh, Cagliari, Rio de Janeiro y Madrid. Recordaremos brevemente algunos de sus trabajos metodológicos en modelos jerárquicos bayesianos, especialmente en temas de bondad de ajuste, imputación múltiple y análisis de valores extremos, y en comparación de modelos y selección de variables. Finalmente, daremos una mirada muy rápida a sus contribuciones aplicadas en Antropología, Cibernética, Educación, Etnobotánica, Etnofarmacología, Genética, Histoquímica y Transporte aéreo, pero nos detendremos un poquito más en sus trabajos en Medicina en relación con el Análisis de supervivencia en modelos multiestado.

Biographical Sketch

Anabel Forte Deltell es licenciada en Matemáticas (2005) y Estadística (2006) por la Universitat de València, donde también obtuvo su doctorado en Matemáticas en 2011. Tras un período posdoctoral en la Universitat Jaume I de Castellón, regresó a la Universitat de València, donde es profesora titular. Su investigación se centra en técnicas estadísticas para la selección de modelos, área en la que ha colaborado con grupos de Europa, Estados Unidos y Australia, publicando más de 40 trabajos en libros y revistas científicas. Actualmente, dirige el Máster en Bioestadística de la Universitat de València y es guionista y presentadora de las Píldoras para comprender y utilizar la Información de la Sociedad Española de Estadística e Investigación Operativa. En el ámbito de la divulgación, coordina el proyecto StatWars, participa en programas de radio como Raíz de 5 en Radio 5, y escribe en su blog BayesAna: Estadística (casi) por todas partes, galardonado como mejor blog de ciencia en los Premios 20Blogs 2021. Es también autora del libro *¿Cómo sobrevivir a la incertidumbre?* (NextDoor Publishers, 2022).

Carmen Armero es Catedrática de Estadística e Investigación Operativa en la Universitat de València, asociada en Biomatemáticas y Estadística de Escocia (BioSS) y directora del Grupo de Investigación Bayesiano de València, VABAR. Su investigación siempre se ha desarrollado en el marco de la inferencia bayesiana, inicialmente en sistemas con

congestión y de espera colas y más tarde en análisis de supervivencia y modelos longitudinales. También le gusta explorar la aplicación de la inferencia bayesiana en materias como la arqueología, la biomedicina, la dosimetría y el deporte. Actualmente es la investigadora principal de la Universitat de València en un proyecto europeo sobre redes bayesianas en la conducción semiautónoma de vehículos. Ha publicado más de 90 trabajos en revistas científicas, ha participado en la presentación de más de 250 ponencias en congresos nacionales e internacionales y colaborado en más de 50 proyectos científicos, como miembro del grupo investigador o como responsable del proyecto.