Frequentist and Bayesian compositional model to analyse longitudinal microbiome data

I. Creus-Martí[1], *A. Moya*[2], *F.J. Santonja*[3]

[1]icreus@alumni.uv.es, Institut de Biologia de Sistemes (I2Sysbio), Universitat de València-CSIC. Departament d'Estadística i Investigació Operativa, Universitat de València.

[2]Andres.Moya@uv.es, Institut de Biologia de Sistemes (I2Sysbio), Universitat de València-CSIC. Fundació per al Foment de la Investigació Sanitaria y Biomèdica de la Comunitat Valenciana (FISABIO). CIBER en Epidemiologia y Salut Pública (CIBEResp)

[3] francisco.santonja@uv.es, Departament d'Estadística i Investigació Operativa, Universitat de València.

Recent studies emphasize the importance of studying longitudinal microbiome data. We have develop a Frequentist and a Bayesian model to analyse the relative abundance of longitudinal microbiome data taken into account the compositional feature of the microbiome data. In both models the relative microbial abundance follow a Dirichlet distribution.

In the Frequentist model, we divide all the Dirichlet parameter between the same selected Dirichlet parameter. The logarithm of this quotient is equal to a regressive expression that use the alr transformation of the taxa and a balance as covariates. The model provides information about the importance of a given taxa and of the rest of the community to define the abundance of that taxa.

In the Bayesian model, the logarithm of the Dirichlet parameters are equal to a regression with the principal balances as covariates. We must mention that principal balances are a compositional tool which chooses the groups of bacteria that maximise the variance and provides information about the relationship between these groups. With this formulation the number of parameters is reduced and we are able to model more microbial taxa. In addition, this model gives information about the dynamics between groups of bacteria are obtained, focusing on which community groups affects the taxa.

We use these two models to analyse microbial taxa time series.

Keywords: Compositional, longitudinal, microbiome.