A Regression Model on Quantiles Extracted from Scans of Asthmatic Patients

Marie-Félicia Beclin*[1,2], Nicolas Molinari[2,3], and Pierre Lafaye De Micheaux[4]

[1]Institut Desbrest de santé publique – Université de Montpellier, Institut National de la Santé et de la Recherche Médicale - INSERM – France

[2]Médecine de précision par intégration de données et inférence causale – PreMeDICaL project team, INRIA – France

[3]Institut Desbrest de santé publique – Université de Montpellier, Institut National de la Santé et de la Recherche Médicale - INSERM – France

[4]UNSW Sydney – Australie

Résumé

We are focused on evaluating the effectiveness of Benralizumab, a drug used to treat asthma, by analyzing tomographic scans taken during expiration and inspiration, before and after one year of treatment. The underlying medical hypothesis is that patients who have benefited from the treatment shows visible improvements in the thoracic scans taken during expiration post-treatment. This means that the patient exhale more efficiently, leading to a more effective emptying of the lungs. This improvement is reflected by higher Hounsfield unit values, with a rightward shift in the post-treatment histogram compared to the pre-treatment histogram.
Irpino and Verde (4) proposed an extension of classical linear regression, applied to quantile functions rather than to real observations. We generalize this approach by deriving the estimation laws of the model parameters using a maximum likelihood method. Starting from the space of quantile functions, we define quantile polynomials before focusing on the specific case of linear regression and explicitly formulate the maximum likelihood estimators. We also propose estimators and confidence intervals. This model was implemented in Python and applied to a real dataset comprising 44 patients treated with Benralizumab.

Although this approach offers certain advantages, particularly its simplicity and alignment with the needs of practitioners, it also has limitations. These include the loss of spatial information and the assumption of linear relationships between voxel distributions. Further work is needed to develop a more robust and generalizable regression method, such as those proposed by Chen et al (1) or Ghodrati et al (3). However, our method retains the advantage of being easy to implement and providing a clear clinical interpretation. We also explore Fréchet regression (5), which allows for the modeling of objects in a metric space based on clinical covariates. We propose a method to quantify the importance of covariates in the model, which we applied to our dataset on expiration histograms before treatment. This approach motivate an extension of our initial model by incorporating clinical covariates.

Moreover, the registration of inspiration and expiration images (2) allows for voxel-by-voxel

*Intervenant

correspondence, paving the way for a generalization of the initial approach. Specifically, it allows for the calculation of the bivariate inspiration-expiration distribution. These histograms contain more information, as they provide insight into the joint distribution during inspiration and expiration. Using the parametric response map (PRM), we can quantify the lung volume affected by asthma in each patient. These PRM data have proven to be better predictors of treatment response than the more commonly used biological biomarkers. Therefore, it would be relevant to explore distribution-to-distribution regression in the context of multivariate distributions. Future work will include predicting 2D histograms post-treatment from the corresponding pre-treatment histograms and the patient's clinical covariates.

References :

(1) Y. Chen, Z. Lin, and H.-G. Müller. Wasserstein regression. Journal of the American Statistical Association, 118(542):869–882, 2023-04-03. Publisher: Taylor & Francis.

(2) C. J. Galbàn, M. K. Han, J. L. Boes, K. A. Chughtai, C. R. Meyer, T. D. Johnson, S. Galbàn, A. Rehemtulla, E. A. Kazerooni, F. J. Martinez, and B. D. Ross. Computed tomography–based biomarker provides unique signature for diagnosis of COPD phenotypes and disease progression. Nature Medicine, 18(11):1711–1715, 2012.

(3) L. Ghodrati and V. M. Panaretos. Distribution-on-distribution regression via optimal transport maps. Biometrika, 109(4):957–974, 01 2022.

(4) A. Irpino and R. Verde. Linear regression for numeric symbolic variables: a least squares approach based on wasserstein distance. Advances in Data Analysis and Classification, 9(1):81–106, 2015.

(5) A. Petersen and H.-G. Müller. Fréchet regression for random objects with euclidean predictors. The Annals of Statistics, 47(2):691–719, 2019.