

Methods for the recovery of missing data in medical research

Danya Facchinetti¹, Federico Tavola¹,
Marie Claire Cantone¹, and Augusto Giussani²

¹*University of Milan, Italy*

²*Helmholtz Zentrum München - German Research Center for
Environmental Health (GmbH), Germany*

Abstract

The data collected in medical analysis are useful to describe physiological processes. Due to the clinical environment, in which logistics and experimental protocols have to follow first the patient needs, data collection can't usually be the same within a population study as for the time of sample collection. Therefore the data collected of the single patients of a population can't be directly comparable to each other. This aspect limits the capability to define a statistical model able to describe the process studied.

In order to bypass this problem, different statistical approaches have been studied to reconstruct the missing data in every data set related to the patients, so that at the end every data set would have the same number of data at the same time point, which make them directly comparable as a synthesis of real and simulated data. The different approaches studied in this work are the smoothing exponential, the moving average, regressive and autoregressive models.

Once the different data sets are directly comparable and then analyzed, a model is defined and then validated, applying it to the initial data sets without the simulated data.

In this work it has been considered, as an example, the typical case of a PET (Proton Emission Tomography) study in order to define the biokinetic of the radiopharmaceutical, due to dosimetric purposes, in which starting from real data sets collected from plasma, urine and organ images, missing time point data have been simulated.

Keywords

Missing data, Time series analysis, PET, SPECT.

References

- Box, G.E.P., Jenkins, G.M., and Reinsel G.C. (2008). *Time Series Analysis*. New Jersey: Wiley.
- Kedem, B. and Fokianos, K. (2002). *Regression Models for Time Series Analysis*. New Jersey: Wiley.
- Molenberghs, G. and Kenward, M.G. (2007). *Missing Data in Clinical Studies*. West Sussex: Wiley.
- Santamaria, L. (1981). *Analisi delle serie storiche*. Bologna: Il Mulino.