Politecnico di Milano
Scuola di Ingegneria Industriale e dell'Informazione
Applied Statistics
July 6th, 2021

Problem n.4

The file `colours.txt` reports 84 measurements of weekly revenues y [k€] for restaurants in Milan (all equipped with a delivery service), collected in the period February-May 2021. The dataset also reports the UTM coordinates s_i of the restaurants and the indication on the 'colour' of the region on the week of the measurement, according to the Covid-19 regulations. Consider for the revenue $y(s_i)$, $i = 1, ...84$, the following model

$$y(s_i) = a_{0g} + \delta(s_i),$$

with $\delta(s_i)$ a stationary residual with spherical model without nugget, and $g = 1, 2, 3$ the grouping induced by the variable *colour* ($g = 1$ for *yellow*, $g = 2$ for *orange*, $g = 3$ for *red*).

a) Assuming $a_{0g} = a_0$ for $g = 1, 2, 3$, estimate the parameter a_0 of the model via generalized least squares. Report the model estimated for $\delta(s_i)$, and discuss the model assumptions.

b) Assuming $\alpha_{0i} \neq \alpha_{0j}$, for $i \neq j$, estimate the parameters a_{0g} of the model via generalized least squares. Report the model estimated for $\delta(s_i)$, and discuss the model assumptions.

c) Which model do you deem more appropriate to describe the data? Comment on your choice.

d) Provide three point predictions $y^*(s_0)$ for the revenues of a restaurant located in the Isola district at location $s_0 = (514811.55, 5037308.54)$, for a yellow, orange and red week.

Upload your results here:
`https://forms.office.com/Pages/ResponsePage.aspx?id=K3EXCvNtXUKAjjCd8ope612LHtvIHvFEsEi2L6mhPglUQkJWQk1QTUlRQ1hESzQwT1pRU1ZWQVpaQy4u`