Politecnico di Milano
Scuola di Ingegneria Industriale e dell'Informazione
Applied Statistics
January 17th, 2024

Problem 4: Spatial modelling of the sales of cold beverages

A study is conducted to analyze the impact of weather conditions, holiday periods, and geographic location on the sales of cold beverages by bars from May to July 2023 in Milan. The dataset `beverages.txt` reports the sales of cold beverages of various bars of Milan at different days in the considered period. It includes the UTM geographical coordinates s_i of the bars, a categorical variable `holiday` indicating if the considered day is a holiday (Saturday, Sunday or bank holiday) or not, the recorded average daily temperature $\texttt{temp}_i$ (in degrees Celsius) during this period and the `sales` of cold beverages during that day $y(s_i)$ [k€/day]. Consider the following model:

$$y(s_i) = b_{0,j} + b_1 \texttt{temp}_i + \delta(s_i) \tag{3}$$

where $\delta(s_i)$ represents stationary residuals with spherical variogram with nugget and $j = 0, 1$ the grouping induced by the variable `holiday` ($j = 0$ for `holiday` = `FALSE`, $j = 1$ for `holiday` = `TRUE`).

a) Report a plot of the fitted variogram. Indicate the estimate of the range and the sill.

b) Estimate the parameters $b_{0,0}$, $b_{0,1}$ and b_1 using the generalized least squares method.

c) Using the model, provide an estimate of the total sales that will be realised by *La Spritzeria* in the month of July 2024 considering a constant temperature of 30°C.

d) Due to the geographical location of the bars, they can be categorized into central (`central=TRUE`) or peripheral (`central=FALSE`). Modify the model in Eq. (3), including this categorical effect, as follows:

$$y(s_i) = b_{0,j,k} + b_{1,k} \texttt{temp}_i + \delta(s_i)$$

where $\delta(s_i)$ represents stationary residuals with spherical variogram *without* nugget and k the grouping induced by the variable `central`.

Provide an estimate of the parameters $b_{0,0,0}$, $b_{0,0,1}$, $b_{1,0}$ and $b_{1,1}$.

How would you describe the effect of the `central` variable (during working days)?

Upload your results here: `https://forms.office.com/e/rv3AKrvmHs`