**Exam of Applied Statistics 15/06/2020**

**Exercise 4**

First of all I predict the revenue using the first observation of the dataframe as a new observation to estimate the two parameters using GLS which are a0 = 56.00482 and a1 = 25.08727. To do so I fitted a variogram obtained with a Gaussian model and a sill of 505.6451 and a range of 484.6377. Then I use this model to perform the prediction and estimate the parameters.

Immagine che contiene mappa, testo, tavolo, uomo

Descrizione generata automaticamente

I fit a linear model for population using distance from Duomo, I check if residuals are normal using shapiro test and obtaining a pvalue of 0.4431 so I can consider them as so, moreover they are homoschedastic and the leverage effect is ok. So I estimate the population using distance of 1666.318 (euclidean distance between duomo and brera) and I obtain 6132.345 as population of Brera. I predict using this population with kriging and I obtain a revenue 94.96324 with a variance of 224.7885. Here is the diagnostic plot for the linear model, we can see that they are okay:

Immagine che contiene mappa, screenshot

Descrizione generata automaticamente

The variance of the prediction is very high and it is not fully informative of the true uncertainty, since I did an universal kriging.