

First Name and Family Name:
ID Number:

Problem n.1

The file `sunchair.txt` collects the prices [€] of 62 sun chairs of comparable characteristics sold by a famous e-commerce company, in 4 different periods of the year: Mar-May, Jun-July, Aug-Oct, Nov-Feb. Assume the prices of different chairs to be independent.

- a) Through an appropriate statistical test, verify if there exist a significant variation in the mean prices of a sun chair during the year. Introduce and verify the needed assumptions.
- b) Use four Bonferroni intervals (global level 95%) to describe the dynamic of the mean price of a sun chair. Based on your analyses, suggest the best period to buy a sun chair and its expected price.

Problem n.2

The province of Ascoli Piceno is well-known for the *olive all'ascolana* (i.e., breaded and fried stuffed olives), hereafter named just *olives*. The file `olives.txt` reports the total weight [g], the filling weight [g] and the weight of the fried breading of 40 olives served in the restaurant *Dalla Luigina*, and of 36 olives served at the *Caff  Muletti*, in Ascoli Piceno.

- a) Is there a significant difference (level 95%) in the mean of the total weight, the filling weight and the breading weight of the olives served in the two restaurants? Introduce and verify the needed assumptions.
- b) Provide T^2 intervals for the mean difference between the total weight, the filling weight and the breading weight of the olives served *Dalla Luigina* and at *Caff  Muletti*. Comment the results.

Problem n.3

Recent studies in the Knossos area have discovered numerous fragments of amphoras. The positions of these fragments (file `knossos.txt`) with respect to the centre of Knossos Palace (considered as origin of the coordinate system), suggest the existence of a second site of archeological interest.

- a) Identify two clusters of locations through a hierarchical clustering method (Euclidean distance and complete linkage). Report the estimate of the mean within the groups, their size, and compute the cophenetic coefficient.
- b) Assume the identified groups to be independent. Having introduced and verified the needed assumptions, test the hypothesis according to which only one archeological site exists in the Knossos area. Write a report of max 3 lines to the archeologists summarising the results of the analysis.

Problem n.4

The Venice lagoon is affected by important tide phenomena, and particularly by the so-called *acqua alta*. Consider the following model for the sea level H [cm] at *Punta della Salute* (in Venice) at 17:00 of each day:

$$H = \beta_0 + \beta_1 \cdot \sin\left(\frac{2\pi}{28}t\right) + \beta_2 \cdot \sin\left(\frac{\pi}{365}(t - t_0)\right) + \beta_3 \cdot t + \varepsilon$$

where $t \in [1, 365]$ is the day of the year, $t_0 = 82$ indicates the vernal equinox, and $\varepsilon \sim N(0, \sigma^2)$. Interpret the first term as due to the effect of the moon motion (astronomical tide), the second term as associated with seasonal effects, the third as due to the global increase of the sea level. The file `tide.txt` reports the sea level measured at 17:00 on 203 days of 2015.

- a) Estimate the five parameters of the model. Report the estimates of β_i , $i \in \{1, 2, 3\}$, and σ .
- b) Having introduced and verified the appropriate assumptions, perform two statistical tests to verify if
 - the mean sea level is influenced by the periodic components;
 - the mean sea level depends on the global increase of the sea level.
- c) Based on point (b), propose a reduced model and estimate its parameters.
- d) Based on model (c), provide two prediction intervals (global level 90%) for the sea level at 17:00 of 20th September 2017 (day 263 of 2017) and of 1st December 2017 (day 335 of 2017). Comment the results knowing that in Venice high-water is expected whenever the sea level is higher than 90 cm.