

Question 1 - Binomial two-way classification model ¹

This question concerns an experiment testing four biological insecticide effects (termed B, \dots, E). In the experiment, each batch containing a known number of insects (varying from 20 to 24 insects) was treated with one of the insecticides or a control (labelled A). The number of dead insects was determined after an incubation period. The batches were incubated in temperature regimens chambers: "C" and "H". The design of the experiment was such that for each combination of insecticide treatment and temperature regimen, there were 20 replicates.

The question related to this exercise is whether there are differences between the insecticides and/or the temperature regimen regarding the probability of survival for a seedling.

The data frame 'Exam2024.01' contains the data for this question. This data frame contains 240 observations (the batches) and 4 variables: "Insecticide" (indicating the insecticide), "TempReg" (representing the temperature regimen used), "N" (the number of insects in the current batch) and "NumberDead" (the number of dead insects after the incubation period).

¹Final Version!

Question 2 - Binomial regression model ²

In this question, we will analyse some results of an experiment involving the germination of seeds of a cactus under different temperatures (from 1 to 40 Celsius degrees) and subjected to a treatment (scarified or not). The experiment consisted in counting the number of germinated seeds after 36 hours from batches with 150 seeds each. That is, each batch of 150 seeds was placed on humidified filter paper in a Petri dish and was incubated at constant temperature for 36 hours; after the incubation, the number of germinated seeds was counted and registered. For each of the 40 different temperatures studied, 40 batches were used: 20 of those batches were subjected to a scarification treatment (i.e. a mechanic treatment with abrasion to remove part of the first layer, the testa, of the seed) and 20 batches were not scarified. .

The question related to this exercise is whether the probability of germination is affected by the treatment and/or the temperature.

The data-frame 'Exam2024.02' contains the data for this question.

The question related to this exercise is to describe how the probability of germination is affected by the temperature and/or the treatment. (Hint: Consider the Arrhenius law of temperature dependence of the rate of a reaction).

²Final Version!

Question 3 -Poisson one-way classification model

3

Here, we will analyse the data of a study to compare the abundance of pollen grains of *Magonia pubens* (a tree common in the central Savannah in South America) in six different geological eras, which is a measure of the predominance of this plant in the ecosystems in the past. In the study, samples of sediments at 20 different places in the bottom of a lake were collected at different depths. By dating the sediments at different depths using Carbon-14, it is possible to associate depth regions with different past Eras. Here, we used six different Eras (termed "A", "B", ..., "F"). Twenty samples (collected at different places) for each Era were analysed, and the number of pollen grains of *Magonia pubens* found in each sample was recorded. The question is whether the data provides evidence that the abundance of the pollen grains of *Magonia pubens* vary from era to era.

The data frame "Exam2024.03" contains the data for this question.

³Final Version!

Question 4 - Poisson regression

The experiment to be analysed in this question regards the determination of the survival capacity of two species of *Penicilium* (a fungus) in soil: *P. restrictum* and *P. simplicissimum*. The experiment consisted of inoculating a portion of soil (from the Brazilian Caatinga) with *P. restrictum* and another portion of the same soil with *P. simplicissimum*. The inoculation was performed with suspensions containing approximately the same quantity of inoculum of *Penicilium*. The two soil portions were kept separated and were incubated for six months under similar conditions. The following procedure was then applied to the soil portion inoculated with *P. restrictum* in order to determine the abundance of this fungus in the soil after the incubation period: a known quantity of soil was suspended in a given amount of water (containing a dispersant) and, after a series of dilutions, a portion of the diluted suspension was plated in a Petri dish containing a cultivation medium⁴ and after an incubation period the colonies of *P. restrictum* formed in the Petri dish were counted and registered. Ten different dilution factors were used in such a way that the amount of soil transferred to the Petri dishes varied from 0.0002g to 0.002g/ Petri dish. For each of the ten dilution factors used, ten replicates were made, so that $10 \times 10 = 100$ Petri dishes were plated. The same procedure was applied in parallel to the portion of soil inoculated with *P. simplicissimum*. The data-frame "Exam2023.04" contains the data for this question. Note that the dataset contains 200 observations corresponding to 10 dilution factors times 10 replicates times 2 separate soil portions inoculated with different species.

You are asked to model the number of colony-forming units that were present in the soil after incubation for each of the two species, compare the abundance (*i.e.* the mean number of CFU per gram of soil) after incubation of the two species and compare the effect of the competition of the two species (if any).

⁴See the chapter 4 of the course lecture notes for details of a similar experiment with another fungus.

Question 5 - Gaussian two-ways classification model ⁵

This question regards the analysis of an experiment on the effect of three different treatments for weeds (represented by the variable "Treat" in the data) on the concentration of NO_3 in the leaf tissues at the yield of four different cultivars of Amaranth (a plant producing grains). The question is whether the treatments affect the concentration of NO_3 in the leaf tissues of the cultivars and whether the effect (if any) of these treatments are similar in the four cultivars.

The data frame "Exam2024.05" contains the data for this question.

⁵Final Version!

Question 6 - Gaussian regression model ⁶

Here we analyse an experiment on the degradation of a pesticide in soil. A certain amount of an organochloride pesticide was added to a portion of soil, which was divided in 240 bags. The soil bags were kept under constant temperature and humidity conditions (similar to all the bags). After 1 month, the contents of the pesticide (in ppm) were determined and registered separately in 30 bags; these bags were then discarded. The same procedure was repeated at 2, 3, . . . , 8 months.

The task here is to find a reasonable statistical model describing the decay in the pesticide concentration in the observed period (1 to 8 months).

The data frame "Exam2024.06" contains the data for this question.

⁶Final Version!