Master AETPF - UE 9.16 - Advanced statistics

Evaluated exercise: writing a scientific report

1   Introduction

The main objective of this work is to get experience in writing two important sections of any scientific report: the "Material and Methods" section and the "Results" section. To do this, you will have to

- understand the data structure

- describe the data

- choose a statistical technique that is appropriate for your data

- perform the statistical analysis

- provide the results of the analysis

You must keep in mind that you are writing this report for people who are not familiar with the data and the analysis. Consequently, you must provide them enough information to understand what you did. On the other hand, the style must be concise and too much information often brings more confusion than anything else. So, try to give only what is necessary. Do not hesitate to compare with scientific literature.

2   Data

You will be given a dataset that comes from the Lake Matapedia Experimental Forest (LMEF). The LMEF originates from a clearcut that was carried out in the mid 1930s. In 1989, a network of 23 permanent plots was established in order to monitor the growth of individual trees. Each fixed-radius plot covers an area of 400m^2. Within the plots, all the trees with diameter at breast height (dbh, 1.3m in height) greater than 1cm were tagged and their dbh was measured. The plot were measured again in 1994, 2000 and 2006.

For each individual tree, the successive measurements were paired in order to create non overlapping intervals. The first measurement provides the initial conditions in which the tree was growing whereas the second measurement provides its status (dead/alive) and its diameter increment in case it survived. The data can be found in the file *LMEFData.csv*.

This file contains the following fields:

Variable	Description	Units
plotID	unique id for each plot	-
treeID	unique id for each tree	-
dateYr	date	year
densityHa	plot-level tree density	stems ha^{-1}
basalAreaM2ha	plot-level basal area	m^2 ha^{-1}
species	species group of the tree	-
dbhCm	tree dbh	cm
BALM2Ha	basal area of all trees with dbh greater than the subject	m^2 ha^{-1}
dead	a binary variable (1= dead, 0=alive)	-
diMm	dbh increment over the growth interval	mm

You may choose to work on either mortality rates or dbh increments.

3 Hypothesis

Those data make it possible to better understand the factors that affect mortality rates and diameter increments. Among others, we can hypothesize that

1. the mortality rates are different among the species

2. the larger the tree, the greater their probability of mortality is

3. the more intense the competition is, the greater the probability of mortality is

 and

1. diameter increments are different among the species

2. larger the trees have smaller diameter increments

3. the more intense the competition, the smaller the diameter increment is

IMPORTANT: Depending on whether you work on mortality rates or diameter increments, your statistical analysis should make it possible to test one of these two groups of hypothesis.

4 Writing the report

This section provides some hints about what kind of information should be given in each section of the report.

4.1 Material and methods

First, describe the data you have. A summary of your data in the form of a Table is often helpful. It should provide minimum, mean, and maximum values for different fields, especially the fields that contain the response variable and the explanatory variables.

In a second step, you should mention which statistical technique you plan to use. The choice of a statistical technique must be supported by some justifications, e.g. the nature of the response variable and the assumptions.

4.2 Results

In this sections, you must provide your final model. You do not need to provide any graph or result regarding the statistical assumptions behind your analysis (i.e. normality, homogeneous variances, and independence). At this point, the reader trust you. However, those assumptions must be checked in your R script.

The resulting parameter estimates and their significance level should be provided in a Table. Some graphs of predictions are also useful to better understand the influence of each explanatory variables on the response.

IMPORTANT: do not provide any interpretation of the results. Statements like "this result was expected because..." or " this is in accordance with some previous work" should appear in the discussion section which is out of the scope of this work.

5 Expected outcome of this work

At the end of this work, you have to provide the two sections of the report (4 pages maximum including Tables and Figures) in a .doc file as well as the R script that contains your statistical analysis. You will be evaluated on both files.