Disease in cucumbers

January 30, 2024

1 Background on data

A greenhouse experiment was carried out to investigate how the spread of a disease in cucumbers depended on variety, climate and amount of fertilizer. Two climates were used: (1) change to day temperature 3 hours before sunrise, and (2) normal change to day temperature. Three amounts of fertilizer were applied: normal (2.0 units), high (3.5 units), and extra high (4.0 units). The two varieties were Aminex and Dalibor. At a certain time the plants were 'standardized' to have equally many leaves, and then (on day 0, say) the plants were contaminated with the disease. On 8 subsequent days the amount of infection (in percent) was registered. From the resulting curve of infection two summary measures were calculated (in a way not specified here), namely the rate of spread of the disease, and the level of infection at the end of the period.

There were 3 blocks each consisting of 2 sections, a section being a part of the greenhouse. Each section consisted of 3 plots, which were each divided into 2 subplots, each of which had 6-8 plants. Thus there were a total of 36 subplots. Results were recorded for each subplot. The experimental factors were randomly allocated to the different units as follows: the 2 climates were allocated to the 2 sections within each block, the 3 amounts of fertilizer were allocated to the 3 plots within each section, and finally the 2 varieties were allocated to the 2 subplots within each plot. Thus, in summary, there were

- 3 blocks
- 2 sections per block (given 2 different climates)
- 3 plots per section (given 3 different amounts of fertilizer)
- 2 subplots per plot (with 2 different varieties)

and variables

- block, numbered 1-3
- section, numbered 1-6
- plot, numbered 1-18

- climate, numbered 1-2
- fert, values 2.0, 3.5 and 4.0
- variety, values aminex, dalibor
- rate, rate of spread of the disease, real number
- level, level of infection at the end of the period, real number

The data were provided by Eigil de Neergård, Department of Plant Pathology, The Royal Veterinary and Agricul- tural University, Denmark.

2 Theoretical part

Consider a four-way analysis of variance with systematic factors given by all cross combinations of climate, fert, variety and random factors block, section, plot.

1. How can variables section, plot and subplot equivalently be formulated as cross factors involving block, climate, fert and variety ?
2. Write down an orthogonal decomposition of $\mathbb{R}^{36}$ involving the above mentioned systematic and random factors (here it may be relevant to use the result of the first question).
3. Construct the factor structure diagram and compute dimensions of the relevant 'V' spaces. Identify the strata for each random effect.
4. Derive the orthogonal decompositions of the covariance matrix, the data vector and the mean vector.
5. Derive the factorization of the likelihood function over the four strata.
6. Construct F-tests for testing the three-way interaction and for testing the main effect of climate. Provide the degrees of freedom.
7. Derive the distribution of the contrast estimate for variety. Specify a confidence interval based on the t-distribution.
8. Construct F-tests for testing whether block, section and plot variances are zero. Provide the degrees of freedom.

3 Practical part

We still consider a four-way analysis of variance with systematic factors given by all cross combinations of climate, fert, variety and random factors block, section, plot.

The aim of the practical part is to assess how the variables fert, climate and variety influence the variables rate and level, either as main effects or by taking part in interactions. We will do that by computing F-tests for interactions and main effects while respecting the hierarchical principle. That is, consider first the three way interaction. If this is not significant then proceed to two-way interactions and soforth.

Note that this is a kind of exploratory analysis. We do not have control of the overall probability of a Type I error since we conduct multiple tests and the number of tests is not known a priori.

1. Why is it relevant to include block, section and plot as random factors ?
2. Check the model using residual plots, qq-plots etc.
3. Assess the systematic factors and their interactions using F-tests as described above. Use both `aov` and `lmer`+`lmerTest` and check that you get the same results with the two procedures.
4. Compute contrast estimate and contrast confidence interval for variety (assuming variety does not enter in any interactions).
5. Estimate the block, section and plot variances. Use again both `aov` and `lmer`.
6. Use F-tests to assess whether any of the variances could be zero.