NTNU – Trondheim
Norwegian University of
Science and Technology

Department of Biology

Examination paper for BI 3051 Evolutionary Analysis

Academic contact during examination: C. Pelabon

Phone: 90824159

Examination date: 8/12/2016

Examination time (from-to): 09:00 / 13:00

Permitted examination support material: authorized calculator

Other information:

Language: English

Number of pages (front page excluded):2

Number of pages enclosed: 3

Informasjon om trykking av eksamensoppgave

Originalen er:

1-sidig X 2-sidig □

sort/hvit X farger □

Checked by:

Date Signature

(20%) Question 1: linear models

A) What are the assumptions of linear models?

B) What are the alternative statistical models you can use if these assumptions are not fulfilled? After clearly explaining which problematic aspects of the data the models can account for, you will provide one or two examples (possibly taken from your field of research) to illustrate these problems.

(30%) Question 2 – Interpretation of models

We study the effect of temperature (in degree C) on time to maturation (in days) in five different populations of an insect species adapted to different ecological conditions. Note that the covariate (temperature) has been mean centered (around the grand mean which equals 25 degrees C). The outputs of the model are presented below.

Explain rapidly the different outputs presented and give a rapid interpretation of these results in biological terms including a graphical representation.

Output 1

```
> model1<-lm(Timemat~temp*factor(pop))
> anova(model1)

Analysis of Variance Table
Response: Timemat
                       Df      Sum Sq  Mean    Sq F value     Pr(>F)
temp                   1       343.44  343.44 26.9823       4.144e-06
factor(pop)            4       195.31  48.83  3.8361        0.008753
temp:factor(pop)       4       59.96   14.99  1.1776        0.332535
Residuals             48       610.96  12.73
```

Output 2

```
> model2<-lm(Timemat~temp+factor(pop))
> anova(model2)

Analysis of Variance Table
Response: Timemat
                 Df     Sum     Sq Mean          Sq F value     Pr(>F)
temp             1      343.44  343.44           26.6186        3.935e-06
factor(pop)      4      195.31  48.83            3.7844         0.008942
Residuals        52     670.92  12.90
```

Output 3

```
> summary(model2)
Call:
lm(formula = Timemat ~ temp + factor(pop))
Residuals:
    Min      1Q  Median      3Q     Max
-6.9453 -2.3021 -0.1252  2.2788  7.8689
Coefficients:
                Estimate    Std. Error          t value         Pr(>|t|)
(Intercept)      13.3031    0.9951              13.369          < 2e-16 ***
Temp            -0.7503     0.1569              -4.783          1.47e-05 ***
factor(pop)2    -0.4507     1.3552              -0.333          0.74078
factor(pop)3    -3.3990     1.4904              -2.281          0.02670 *
factor(pop)4    -2.0439     1.6634              -1.229          0.22470
factor(pop)5    -4.6637     1.4131              -3.300          0.00175 **

Residual standard error: 3.592 on 52 degrees of freedom
Multiple R-squared: 0.4454,    Adjusted R-squared: 0.392
F-statistic: 8.351 on 5 and 52 DF,  p-value: 7.52e-06
```

(15%) Question 3 – Statistical power

A) What is the power of a statistical test?

B) When you are conducting an experiment, for example to test the effect of a chemical on the metabolic rate of some birds, how can you increase the power of your statistical test?

C) Is it always desirable?

(35%) Question 4 – Study design and statistics

Remark: In this exercise we consider plants functional traits as examples of traits to measure. However, this is only an example and you don't need to know more about functional trait that what is written here to answer the question.

Plants functional traits are defined as measurable species characteristics with explicit connections to individual performance. These traits reflect plant ecological strategy and vary consistently along climate gradients. Therefore, they are often considered as potential predictors of plant community response to climate change. In a study aiming at better understanding the variation of functional traits with specific environmental variable, we conduct a study where traits are measured in different species from two types of plant along specific environmental gradients.

We sampled 6 species from two different life forms (grass and shrub) that are widely distributed, so that they be easily found in contrasting conditions. In this study we focus on two environmental factors: average summer temperature and summer precipitation. In order to estimate the effects of these two factors on the variation of functional traits, we select for each species 5 replicated series of 12 sites. The 12 sites are chosen with one of three mean summer temperatures (~6.0, 9.0, and 10.5°C) and one of four mean annual precipitations (~600, 1200, 1900, and 2800 mm), so that each combination of temperature and precipitation is represented in the sampling procedure. Eventually, we obtain 5 replicated sites for each combination of temperature and precipitation. Other environmental variables are relatively consistent across sites (calcareous soil, southwest aspect, slope of about 20°, and comparable grazing and land-use history).

For each sites, we select three subpopulations (square of 50 cm × 50 cm for grass species, and 10 m × 10 m for shrubs). In each subpopulation, three random individuals are measured. For each plant we measure the plant height, the leaf area and the specific leaf area (leaf area / leaf dry mass). For leaf trait measurement, we collect one leaf per individual for grass species and 10 for shrubs; trait values were averaged at the individual level for the analyses.

A) In this study, what are the predictor and response variables?

B) For each of the response variable, what is the total sample size?

C) How would you analyze these data? Explain clearly the type of model that you want to do and what will be the hypothesis tested by this model. If the models appear too complex, suggest ways to simplify the analysis. To do this, you will pay particular attention to the design of the sampling.

D) Are all observations independent? If not what are the source of dependency?