Ciências ULisboa

ECOLOGIA NUMÉRICA

Exame teórico – 1ª Época 2018/2019

Nome:..... **Nº aluno:** ..... **Classificação:** .....

1. A t-student test was used to test the null hypothesis that 2 different cetacean species have the same number of newborn per year. Two random samples of females were collected, one for each species, and the mean observed number of newborns per year per female was 0.371 for *Bichus inventadicus* and 0.373 for *Bichus nonexistenticus*. The P-value of the test was 0.0032.

1.1 (cotação 1) What can you conclude about the sample size used? Support your answer.

.....

.....

.....

.....

1.2 (cotação 1.5) What does the value 0.0032 represent (i.e., what does a P-value represent)? Support your answer.

.....

.....

.....

.....

1.3 (cotação 1) What would be your interpretation of these results, regarding the statistical and biological significance? Support your answer.

.....

.....

.....

.....

2. There are several measurement scales for random variables.

2.1 (cotação 1) What is the difference between a random variable measured in an ordinal versus a nominal scale?

.....

.....

.....

.....

2.2 (cotação 0.5) Present a biological example of each type.

.....

.....

.....

.....

2.3 (cotação 0.5) In R, what is the class of objects one can use to store these types of variables?

.....

.....

3. In Numerical Ecology we work with random variables. There are two kinds of random variables, that we can charactyerize via their probability mass function and probability density function.

3.1 (cotação 1) What is a random variable?

.....

.....

3.2 (cotação 1) Present two properties of a probability density function, usually represented by f(x).

.....

.....

3.3 (cotação 1) Provide a biological example of each variable type, and provide the family of distributions more sensible to model such a random variable.
Example 1, type 1:.....

Family to model example 1:.....

Example 2, type 2:.....

Family to model example 2:.....

4. A common objective in numerical ecology is to compare means across different samples, testing the null hypothesis that all samples come from a population with a equal mean value.

4.1 (cotação 1) What are the assumptions of the usual parametric analysis used for that purpose?
.....

.....

.....

.....

4.2 (cotação 0.5) Whan the assumptions do not hold, what alternative procedure can you use ?
.....

.....

.....

.....

4.3 (cotação 1) if the null hypothesis was not rejected, what would you conclude? Assuming the null hypothesis is rejected, what should you do next?

.....

.....

.....

.....

5. A GLM was used to evaluate the relation between the concentration of a stress related hormone, as a function of some environmentla variables measured in the capture location. The corresponding R output is shown below.

```
Call:
glm(formula = ys ~ profundidade + largura + o2 + temperatura,
    family = Gamma(link = log))

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-1.80585  -0.48991  -0.06917   0.28533   1.62909

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.010893   0.733024   2.743  0.00665 **
profundidade  0.190921   0.015452  12.356  < 2e-16 ***
largura      -0.066017   0.021883  -3.017  0.00289 **
o2           -0.003998   0.008347  -0.479  0.63247
temperatura   0.091596   0.010730   8.537 3.88e-15 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Gamma family taken to be 0.3517987)

    Null deviance: 148.758  on 199  degrees of freedom
Residual deviance:  75.893  on 195  degrees of freedom
AIC: 1782.5

Number of Fisher Scoring iterations: 5
```

5.1 (cotação 1) Choose a significance level, explain your choice, and based on it say which variables seem to be significant to explain the stress hormone concentration.

.....

.....

.....

5.2 (cotação 1.5) What the estimated mean value of the hormone for a fish captured at a place with profundidade 6 meters, largura 8 meters, o2 of 75.3 % and temperatura of 15.3 ºC (assuming the units used here were those used in the model). Present the equation used for this.

.....

.....

.....

5.3 (cotação 0.5) Where do you expect to find more individuals, in a place not very largo or in a very largo place? Explain your reasoning.

.....

.....

.....

6. (cotação 1.5) CTo understand if the invertebrate community of puddles from the west coast of Portugal in the tide area are organized mostly (1) according to a latitudinal gradient or (2) with a gradient depending on the low tide to the high tide areas, we collected a sample from 50 locations along the Portuguese coast, regarding the invertebrate species abundances. In each location 10 puddles were sampled. Describe how you could implement an ordination analysis to answer this ecological question.

.....

.....

.....

.....

.....

.....

.....

.....

.....

7. The following output represents a PCA over the environmental variables collected at several locations along a river. The colors of each location corresponds to a group obtained by a clustering methods. The number associated with each point is proportional to the distance to the moth of the river. Adapted from Borcard, D.; Gillet, F. & Legendre, P. 2011. Numerical Ecology with R, Springer.

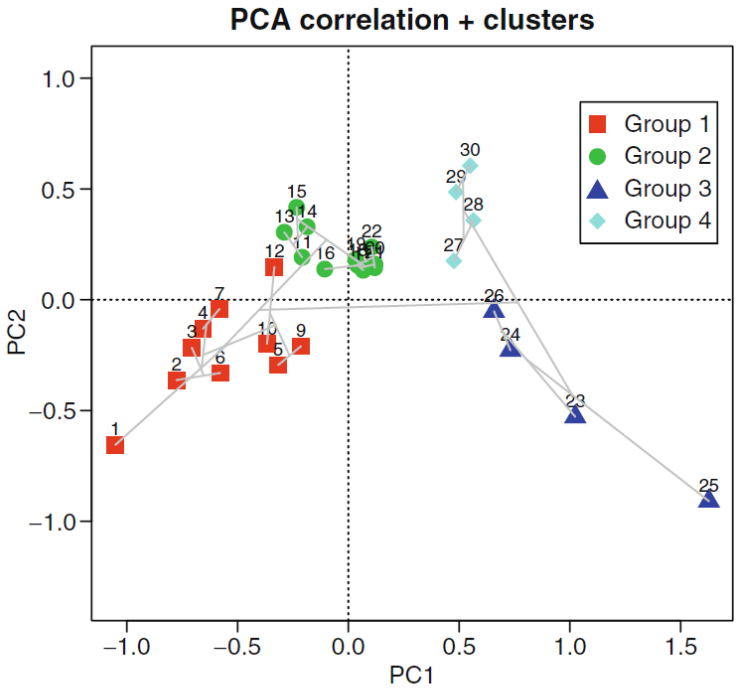

Fig. 5.3 PCA biplot (scaling 1) of the Doubs environmental data with overlaid clustering results

7.1 (cotação 1.5) Based on this analysis, how would you describe the organization of the places as a function of its environmental covariates?

.....

.....

.....

.....

.....

.....

7.2 (cotação 1) Considering the analysis above, what is more sensible, that the matrix of correlations or the matrix of variance covariance was used? Support your reasoning.

.....

.....

.....

.....

.....

.....

7.3 (cotação 1) If you are told that the percentage of variation explained by the first two axis was 92%, and that the analysis considered 15 environmental variables, comment on the analysis performance in terms of dimensionality reduction.

.....

.....

.....

.....

.....

.....

8. (cotação 1) While comparing the weights of 2 species, *Petrofagus decametrus* e *Petrufagus tetrametrus*, the weight of 10 animals per species were recorded along a 30 week period. A biologist whom had a 10 in Numerical Ecology presented the following comparison output resulting from comparing the 300 pairs of values to one biologist whom had a 20, telling him this was proof that the two species had different weights. The Biologist whom had 20 was not convinced. Why?

```
        Welch Two Sample t-test

data:  x1 and x2
t = -2.211, df = 596.82, p-value = 0.02742
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.32982146 -0.01951443
sample estimates:
mean of x mean of y
0.5344414 0.7091094
```

.....

.....

.....

.....

……

……

1.1 (cotação 1)
1.2 (cotação 1.5)
1.3 (cotação 1)
2.1 (cotação 1)
2.2 (cotação 0.5)
2.3 (cotação 0.5)
3.1 (cotação 1)
3.2 (cotação 1)
3.3 (cotação 1)
4.1 (cotação 1)
4.2 (cotação 0.5)
4.3 (cotação 1)
5.1 (cotação 1)
5.2 (cotação 1.5)
5.3 (cotação 0.5)
6. (cotação 1.5)
7.1 (cotação 1.5)
7.2 (cotação 1)
7.3 (cotação 1)
8. (cotação 1)