

Proposta A

Morphological measures between similar species

The function goal is demonstrated differences in morphological dimensions between similar species. For this example it will use in measures used in frogs.

Home: table/data.frame containing information on measures of each species. In total 11 morphological variables will evaluate: (1) Snout-vent length (SVL); (2) head width (HW); (3) head length (HL); (4) Upper eyelid width (UEW), (5) horizontal eye diameter (ED); (6) eye-nostril distance (END); (7) inter-orbital distance (IOD); (8) inter-narial distance (IN); (9) tympanum diameter (TD); (10) tibia length (TL); (11) foot length (FL).

Function operation=

â€¢ Variance and standard deviation in dimensions of each species

â€¢ Differences between means of each dimensions for the species

â€¢ To assess normality in data distribution:

o Normal â€” PCA

o Not normal â€” KruscalWalis

â€¢ Making graphics

o Multivariate anova/PCA

â€¢ Output â€” table with the descriptive information (means, variance and standard deviation), and graphics PCA.

With this function are identified quantitative differences between similar species.

Comentários Vitor Rios

Jhon, your formatting is confusing, please reformat your proposal, and try to describe in words what your function does. From what I could understand you are going to calculate descriptive statistics for each variable, is that it?

comment

Hi Vitor,

The purpose of the function is to analyze the data so descriptive and exploratory.

input: table or data.frame containing the morphological dimensions (eleven) for each species.

Vitor Rios

John, please rewrite you proposals. they are appearing with strange characters for me, and the formatting is confusing.

comment

Function

Input: table containing the morphological dimensions for each species ([table.txt](#)).

The purpose of the function is to analyze the data so descriptive and exploratory.

1) Mean and standard deviation of each species.

2) Testing the existence of differences between measurements for each species

1. Test normality

Normal (PCA)

Not normal (KruscalWalis)

3) Make graphics PCA

4) Output - table with the descriptive information (means, variance and standard deviation), and graphics PCA.

Proposta B

Percentage of purines and pyrimidines in a DNA sequence

Two functions for know the percentage of purines and pyrimidines in DNA sequence of diffrente taxas. The function goals are identify the frequency and perform analysis to determine the presence of differences significant between taxas. In the end, these results will be shown by graphs.

Home: file.txt containing the sequence of DNA.

First funtion: purines

Operation of the function:

â€¢ Sum the amount of purines ("A" and "G") that occur along a DNA sequence.

â€¢ Divide this sum over the total bases containing the DNA sequence and multiplies this value by

100.

Second function: pyrimidine

â€¢ Sum the amount of pyrimidines ("C" and "T") that occur along a DNA sequence.

â€¢ Divide this sum over the total bases containing the DNA sequence and multiplies this value by 100.

Comentários Vitor Rios

This function could be very basic our very complex, depending on your file input. It is basically an application of the table() function, is that correct?

in general, please revise the formatting of your proposals so that we can understand better

comment

Hi Vitor, This function differs from the table function (X). In my proposal the input file is a file in FASTA format, which contains the sequence for each species

(<http://www.ncbi.nlm.nih.gov/BLAST/blastcguihelp.shtml>). Executing this function will be known frequencies of purines ("A" and "G") and pyrimidines ("C" and "T") in each input files (species).

comment

Input: file in FASTA format containing the sequence of DNA ([js_2.txt](#)).

The purpose of the function knows the percentage of purines (A" and "G") and pyrimidine ("C" and "T") in DNA sequence of different taxas.

First function:

1) Sum the amount of purines ("A" and "G") that occur along a DNA sequence. 2) Divide this sum over the total bases containing the DNA sequence and multiplies this value by 100. 3) Make graphics (Histograms showing results)

Second function:

1) Sum the amount of pyrimidine ("C" and "T") that occurs along a DNA sequence. 2) Divide this sum over the total bases containing the DNA sequence and multiplies this value by 100. 3) Make graphics (Histograms showing results)

Vitor Rios

John, both your functions are very simplistic, unless you intend to

code analyses by hand, and not simply call functions which do analyses for you, especially the first one. Also, if you are going to do an “automated” analysis, you *really* should test for the other premisses of both ANOVA and PCA, normality is not the only one. The second one is just an application of the `table()` function on a string of characters, unless you decide to do the counting by hand, again.

From:
<http://ecor.ib.usp.br/> - **ecoR**

Permanent link:
http://ecor.ib.usp.br/doku.php?id=05_curso_antigo:r2015:alunos:trabalho_final;jhon.sarria:propostas_para_trabalho_final 

Last update: **2020/08/12 06:04**