

Mariana Rodrigues dos Santos



Mestranda em Oceanografia com ênfase em Oceanografia Biológica

O título da minha tese é: Ocorrência, Distribuição e Abundância das larvas de Serranidae na região entre cabo Frio (RJ) e cabo de Santa Marta Grande (SC). Sob a orientação de: Prof. Dr. Mário Katsuragawa.

Meus exercícios

Exercício 1 [resolucao_lista1.r](#)

Exercício 2 [resolucao_lista2.r](#)

Exercício 3 [exercicio_lista3.r](#)

Exercício 4 [resolucao_lista4.r](#)

Exercício 5 [exercicio_aula5_final.r](#)

Exercício 6 [exercicio_lista6.final.r](#)

Exercício 7 parcial: [exercicio_lista7.r](#)

Exercício 7 final: [exercicio_lista7_final.r](#)

Exercício 9: [resolucao_lista9.r](#)

Trabalho Final

Proposta

Realizar uma análise exploratória de dados, comparando os dados abióticos (temperatura, salinidade, profundidade do local) e abundância das larvas de peixes pertencentes à família Serranidae, de acordo com cada sub-família identificada. Minha planilha inicial é estruturada da seguinte forma: as linhas representam as amostras e as colunas representam a data, hora, latitude e longitude do ponto coletado, profundidade local, Profundidade de coleta, Temperaturas (0 e 10m), Salinidade (0 e 10m), Total de larvas pertencentes à família, Total de larvas de Anthiinae, Serraninae, Epinephelinae e relação d/v*. Primeiramente a planilha será transformada em um data.frame e será determinada a abundância total de acordo com as três sub-famílias. Posteriormente será realizado testes de correlação entre os dados abióticos e a abundância e serão gerados gráficos do tipo boxplot e

espalhogramas com linhas de tendência como dados de saída, além de um data.frame com os resultados obtidos.

*Relação d/v = razão entre a profundidade máxima de coleta e volume de água filtrado pela rede.

Plano B

Gerar uma função para estimar a abundância total e relativa das larvas de peixe pertencentes à família Serranidae de acordo com a classificação em sub-famílias. Para a determinação da abundância das larvas (dada em larvas/m²) será utilizado a fórmula proposta por Tanaka (1973): $N = x \cdot (d/v)$ em que o N = abundância; x = número de larvas coletadas; d = profundidade máxima de coleta e v = volume de água filtrado pela rede. Será necessário determinar o volume de água filtrado através da equação: $v = a \cdot n \cdot c$, sendo v = volume filtrado; a = área da boca da rede; n = número de rotações do fluxômetro; c = coeficiente de calibração do fluxômetro. Apesar de ser aplicado a uma família, esta função poderá ser aplicada a outros organismos pertencentes ao domínio zooplanctônico marinho (copepodes, ovos, larvas, etc).

Comentários

PI

A primeira proposta me parece muito específica para seu conjunto de dados e pergunta. Se é isso, tente generalizar. A segunda proposta me parece ser um cálculo muito simples de densidade (larvas/volume). Pense em uma maneira de torná-la mais desafiadora (cálculo de intervalos de confiança? Estimativa e biomassa?).

MARIANA

Olá Paulo, tudo bem? Minha primeira proposta realmente, do jeito que está escrito parece específica para meu conjunto de dados e proposta, pois tomei como base meu trabalho para tentar montar minha função. Entretanto, poderá ser aplicada de forma generalizada para qualquer tipo de organismo marinho. Gostaria de saber qual das duas propostas seria mais interessante a ser trabalhada. Vou tentar generalizar. Obrigada pelas dicas. Mariana

PROPOSTA A MELHORADA

Com o objetivo de melhorar a visualização da função do cálculo de correlação de Pearson e Spearman entre mais de duas variáveis, minha proposta é criar uma função que concatena os dois fatores de correlação citados acima e como resultado, além da tabela de correlação apresentada, uma página de gráfico mostra uma matriz de espalhogramas, com as retas não paramétricas de tendência. Um data.frame com as variáveis que se pretende calcular a correlação será o dado de entrada dessa função. No return da função, uma lista contendo o sumário do data.frame inicial, a

tabela de correlação (ou Pearson ou Spearman - de acordo com a necessidade do operador) e o gráfico de matriz de espalhogramas com as linhas de tendência serão os outputs dessa função. Como forma de testar a função, utilizarei meu banco de dados de comparação entre os fatores abióticos (temperatura, salinidade, biovolume) com a abundância das larvas de Serranidae. Apesar desse teste ser realizado com meus dados, essa função pode ser realizada com qualquer conjunto de variáveis que se deseja analisar sua correlação. Dentro da própria função, uma análise preliminar dos dados, retirando possíveis NA's existentes será realizada.

TRABALHO FINAL

HELP

```
cor.ps          package=unknown  R Documentation

**Description:
Calculo dos coeficientes de correlação de Pearson/Spearman de um data.frame
com duas ou mais variáveis, mostra matriz de espalhograma e análise
exploratória de dados omitindo os valores de NA quando presentes e mostrando
um boxplot dos dados iniciais.

**Usage:
cor.ps(x, Pearson=TRUE, NAs=TRUE)

**Arguments:
x: data.frame dos dados que se deseja calcular o coeficiente de correlação
de Pearson/Spearman
Pearson: lógico. Se TRUE, calcula o coeficiente de correlação de Pearson. Se
FALSE, calcula o coeficiente de correlação de Spearman.
NAs: lógico. Se TRUE, indica presença de NAs nos dados e retira para
calcular o coeficiente de correlação. Se FALSE, indica ausência de NAs nos
dados e ele calcula o coeficiente de correlação sem retirar os valores de
NAs.

**Details:
O input dessa função deve ser um data.frame contendo duas ou mais variáveis,
todas com mesmo comprimento.
x - no primeiro argumento da função, quando deseja calcular o coeficiente de
correlação de algumas variáveis do data.frame, deve-se indexar o valor
correspondente ao das colunas que se queira calcular a correlação. (Ex:
dados[,3:6] - dos dados desejo calcular a correlação das colunas 3 a 6).
Como análise exploratória dos dados iniciais, além do sumário, um boxplot
dos dados é gerado com o intuito de melhorar a visualização dos dados, bem
como identificar presença de outliers que possam interferir nos resultados e
uma matriz de espalhograma dos dados originais.
Matriz de espalhograma dos resultados - representação gráfica dos resultados
da matriz de correlação de Pearson/Spearman

**Value:
Return da função em formato de lista:
sumario: summary dos dados do data.frame original
```

resultado: result do coeficiente de correlação de Pearson/Spearman

****Warning:**

Para calcular o coeficiente de correlação de algumas das variáveis do data.frame, utilizar indexação, como mostrado nos detalhes do help. Para calcular o coeficiente de correlação, os dados das variáveis devem ter mesmo comprimento.

****Author(s):**

Mariana Rodrigues dos Santos santos.mariana@usp.br

****References:**

- * Triola, M. F. Introdução a Estatística. 9ed. Editora LTC. Rio de Janeiro. 2005
- * Dalgaard, P. Introductory Statistics with R. Springer. 2002

****Examples:**

```
#Baixar o banco de dados:
dados.brain <-
read.csv("http://ecologia.ib.usp.br/bie5782/lib/exe/fetch.php?media=dados:an
imais.txt", sep=";", dec=",", header=TRUE, as.is=TRUE)
# Teste Pearson:
cor.p <- cor.ps(dados.brain, Pearson=TRUE, NAs=TRUE)
# Teste Spearman
cor.s <- cor.ps(dados.brain, Pearson=FALSE, NAs=TRUE)
```

FUNÇÃO

```
cor.ps <- function(x, Pearson=TRUE, NAs=TRUE)
{
if (NAs==TRUE)
{dados=na.omit(x)
cat("valores NA retirados\n")
x11()
pairs(dados)
x11()
boxplot(dados)}
else
{dados=x
x11()
boxplot(dados)}
if(Pearson==TRUE)
{cor1=cor(dados)
result=cor1
cor1.dataframe <- as.data.frame(cor1)
x11()
matrix=pairs(cor1.dataframe)}
else
{cor2=cor(dados,method="spearman")
result=cor2
```

```
cor2.dataframe <- as.data.frame(cor2)
x11()
matrix2=pairs(cor2.dataframe)
}
return(list(sumario=summary(x), resultado=result))}
```

ARQUIVOS [dados.correl.csv](#) [cor_ps.r](#)

From:

<http://lbtrop.ib.usp.br/> - **Laboratório de Ecologia de Florestas Tropicais**

Permanent link:

http://lbtrop.ib.usp.br/doku.php?id=cursos:ecor:05_curso_antigo:alunos2012:alunos:trabalho_final:santos.mariana:start 

Last update: **2020/07/27 18:46**