

Marinella Coutinho Jacinto Pucci

Doutoranda em Oceanografia, Instituto Oceanográfico/IOUSP.

O título da minha tese é “Interações tróficas da comunidade do sistema sublitoral da Baía do Araçá, município de São Sebastião (SP): uma abordagem pelo uso de isótopos estáveis”, e sou orientada pela Profa. Lucy Satiko Hashimoto Soares.

[exec](#)

Resposta da Marinella

Oi Thais e Vitor, percebi a besteira que fiz depois de postar as propostas, mas fiquei com receio de mudar a forma devido ao prazo de postagem... Confesso o erro! 😊 Tentei especificar melhor as propostas agora...

Oi Marinella, o ideal é você escrever suas propostas aqui no wiki mesmo, nada de escrever em word e pedir para que o visitante baixe o arquivo só pra ler. Assim sua função fica acessível pra qualquer um. O site da disciplina é mto visitado, por alunos do país todo...buscar uma função que está dentro de um arquivo word não é nada prático =] ---

[Thais Lopes](#)

Comentários Vitor Rios

Além de postar no wiki, esclareça melhor o que cada função vai fazer, como é o calculo dos índices, se vc vai implementar o calculo ou chamar um pacote específico

Trabalho Final

Proposta A

Tradicionalmente, isótopos estáveis de carbono e nitrogênio têm sido utilizados em estudos de tramas tróficas de ecossistemas terrestres e aquáticos, pois permitem a identificação das fontes de matéria orgânica e da posição trófica dos organismos nos ecossistemas.

Assim, minha função será uma análise gráfica de duas variáveis quantitativas, os valores das razões de isótopos estáveis de carbono e nitrogênio ($\delta^{13}\text{C}$ e $\delta^{15}\text{N}$) para o estudo de interações tróficas. De maneira geral, um gráfico de dispersão com as médias dos valores de $\delta^{13}\text{C}$ (‰) e de $\delta^{15}\text{N}$ (‰) das espécies permite a identificação de (possíveis) fontes de carbono no sistema, enquanto um gráfico com os valores dos níveis tróficos possibilita a visualização da posição trófica das espécies.

O objeto de entrada da função é uma tabela com: nº do indivíduo/amostra, espécie, época de coleta (e/ou local de coleta), e valores de $\delta^{13}\text{C}$ (‰) e de $\delta^{15}\text{N}$ (‰).

Os objetos de saída são: um gráfico de dispersão com as médias dos valores de $\delta^{13}\text{C}$ (‰) e de $\delta^{15}\text{N}$ (‰) das espécies; um vetor com os valores estimados do nível trófico de cada espécie e a

representação gráfica desses níveis tróficos.

O cálculo do nível trófico será implementado na função. O nível trófico de cada espécie pode ser estimado através dos valores médios de $\delta^{15}\text{N}$ de cada espécie, de acordo com Vander Zanden & Rasmussen (2001): $\text{NT} = [(\delta^{15}\text{N}_{\text{espécie}} - \delta^{15}\text{N}_{\text{bivalve}}) / \text{TEF}] + 2,0$ onde o valor de TEF representa o fator de enriquecimento médio entre um nível trófico e outro, e 2,0 representa o nível trófico médio do consumidor primário. Vander Zanden & Rasmussen (2001) propuseram a substituição do valor isotópico de um produtor primário por um consumidor primário, como os bivalves, já que esse grupo processa grande quantidade de material em curto período de tempo e apresentam longevidade maior do que o fitoplâncton, com ciclo de vida mais curtos. Além disso, amostras puras de fitoplâncton são de difícil obtenção.

Para verificar se existem diferenças significativas entre as amostras coletadas em diferentes épocas (e/ou locais), também será aplicado um teste de ANOVA.

Proposta B - alternativa

O estudo da alimentação de peixes permite a compreensão e identificação da trama trófica em um ecossistema aquático. A análise de conteúdos estomacais é um método tradicional e muito utilizado, pois fornece informações detalhadas sobre a dieta de uma determinada espécie.

Minha função será o cálculo do Índice de Importância Relativa (IIR) para a análise da importância das presas na dieta de uma espécie de peixe; e a confecção de um cluster, com base no índice, para comparações intraespecíficas na alimentação da espécie (como diferenças sazonais ou locais).

O objeto de entrada dessa função é uma tabela com os dados das categorias de presas (ocorrência, massa e número - todos em %) e o local de coleta (e/ou época do ano).

Os objetos de saída são: uma coluna na tabela de entrada com os valores estimados de IIR (%) para cada categoria de presas da espécie; e um cluster de similaridade com base no IIR das presas por local de coleta (e/ou época do ano). O cálculo do Índice de Importância Relativa (IIR%) (Pinkas *et al.*, 1971) será implementado na função e pode ser estimado pela fórmula: $\text{IIR\%} = (\text{N\%} + \text{M\%}) * \text{O\%}$. O objetivo de combinar ocorrência, número e massa das presas em uma fórmula é balancear valores associados a grandes presas, presentes em menor número, com os valores de exemplares menores e, geralmente, mais numerosos.

A análise de agrupamento será feita por similaridade (distância euclidiana), através da função `hclust` do pacote `stats`. Caso haja formação de grupos no cluster, será aplicado teste de ANOVA para verificar se a diferença entre os grupos formados é significativa.

Comentários Vitor Rios

Gostei mais da segunda função. Na hora da comparação dos cluster, lembre que é necessário verificar as premissas da anova para escolher o teste adequado (paramétrico/não paramétrico e tal)

Help da Função

`cluster.IIR`

`package:unknown`

R Documentation

Cálculo do Índice de Importância Relativa (IIR total e IIR%) das presas;

análise de agrupamento para a dieta de peixes por uma variável categórica (local, época ou classe de tamanho).

Description:

A função calcula o Índice de Importância Relativa das presas na dieta de uma espécie de peixe, a partir de ocorrência, número e massa. Gera uma matriz de distância, um cluster e um dendrograma, considerando o IIR% e os níveis da variável categórica. Testa diferenças na composição das presas por nível da variável categórica.

Usage:

```
cluster.IIR <- function(presas, x="euclidean", y="average")
```

Arguments:

presas: data.frame com quatro colunas, com os seguintes nomes e ordem:

- 1ª coluna: (sem nome específico) apresenta as presas,
- 2ª coluna: O% (Frequência de Ocorrência das presas),
- 3ª coluna: N% (Frequência Numérica das presas),
- 4ª coluna: M% (Frequência Gravimétrica das presas),
- 5ª coluna: Categoria (Variável categórica e seus níveis: diferentes locais de coleta, ou épocas de coleta ou classes de tamanhos dos peixes).

x: índice utilizado para a elaboração da matriz de distância no objeto dist. 0 default da função é a distância euclidiana.

y: método de agrupamento utilizado para a construção do cluster. 0 default da função é o método "average" (método aglomerativo UPGMA).

Details:

Os nomes e ordem das colunas no objeto de entrada da função (data.frame) devem ser mantidos.

Para mais informações sobre cálculos das frequências de ocorrência, numérica e gravimétrica das presas, consultar Hyslop (1980) e Pinkas et al. (1971).

Em x, outros índices podem ser utilizados para a confecção da matriz de distância, desde que respeitadas suas premissas. Consultar no tópico "See Also" abaixo.

Em y, outros métodos de agrupamento podem ser utilizados, de acordo com os objetivos e o tipo de matriz de distância. Consultar no tópico "See Also" abaixo.

Value:

A função calcula o Índice de Importância Relativa das presas na dieta de um peixe. A partir da ocorrência, número e massa, o IIR balanceia valores de exemplares grandes e menos numerosos, com os menores e mais numerosos. São gerados valores de IIR total e IIR% das presas nos diferentes níveis da variável categórica.

Para isso, a função fornece como objetos de saída:

1) data.frame que integra os dados originais e os calculados pela função. A tabela será salva como "Tabela Alimentação.txt" no diretório de trabalho.

2) gráfico do tipo dendrograma, salvo como "Dendrograma1.jpg" no diretório de trabalho, para a identificação de possíveis grupos de acordo com os níveis da variável categórica.

3) lista com:

comp1: data.frame que integra os dados originais e os calculados pela função.

comp2: matriz de distância calculada pelo índice indicado e com base nos valores do IIR% das presas.

comp3: cluster gerado através da matriz de distância e agrupado pelo método indicado.

comp4: resultado do teste de normalidade de Shapiro para o IIR total em cada nível da variável categórica.

comp5: sumário do teste de ANOVA, realizado com base nos valores de IIR total.

Warning:

A função considera apenas uma variável categórica e seus diferentes níveis (diferentes locais de coleta, ou diferentes épocas de coleta, ou diferentes classes de tamanhos).

Não serão enviadas mensagens de advertência caso as premissas para o teste de ANOVA não sejam cumpridas. O teste de ANOVA será executado independentemente do resultado do teste de Shapiro.

Para a variável categórica que apresente apenas 2 níveis, um teste t-student deve ser executado além da função.

Note:

A confecção da matriz de distância e do cluster requerem a instalação dos pacotes stats e vegan.

Author(s):

Marinella C. J. Pucci (mpucci@usp.br)

References:

Hyslop, E.J. 1980. Stomach contents analysis – a Review of methods and their application. J. Fish. Biol., 17: 411-429.

Pinkas, L.; Olephant, M.S. & Iverson, I.L.K. 1971. Food habit of albacore, bluefin tuna and bonito in California Waters. Fish Bull., 152: 1-105.

Gotelli, N.J. & Ellison, A.M. 2011. Princípios de estatística em Ecologia. Porto Alegre: Artmed. 528p.

See Also:

Para informações complementares, ver: `vegdist {vegan}`, `hclust {stats}`, `rect.hclust {stats}`.

Examples:

Como exemplo, usar a planilha "dieta.csv": `read.table ("dieta.csv", header=TRUE, sep=",")`.

Código da Função

```
# Para utilizar a função, é necessária a elaboração de uma planilha inicial
# de dados em formato csv. ou txt. Para
# mais informações, consultar o help da função cluster.IIR

cluster.IIR <- function(presas, x="euclidean", y="average") # Cria a função
cluster.IIR
{
  library(vegan) # Carrega o pacote "vegan", considerando que já tenha sido
  instalado.
  library(stats) # Carrega o pacote "stats", considerando que já tenha sido
  instalado.
  IIR.total <- (presas$N+presas$M)*presas$O # Calcula o IIR total a partir
  das informações de ocorrência, número e massa
  # fornecidas no objeto de entrada (data.frame) e disponibiliza os
  resultados em um vetor.
  dados<- data.frame(presas, IIR.total) # Cria um objeto do tipo data.frame
  com os dados originais do objeto de entrada e uma nova coluna
  # com o IIR total calculado.
  soma.total <- tapply(X=dados$IIR.total, INDEX=dados$Categoria, FUN=sum) #
  Calcula a soma do IIR total para cada nível da variável categórica -
  # coluna Categoria (local de coleta, época de coleta, classe de tamanho)
  soma.IIR <- soma.total[dados$Categoria] # Cria um VETOR de repetições dos
  valores do IIR total dos níveis da variável categórica.
  dados2 <- data.frame(dados, soma.IIR) # Cria um objeto do tipo data.frame,
  com as informações da tabela DADOS e uma nova coluna
  # com a soma de IIR total para cada nível da variável categórica.
  IIR.perc <- (dados2$IIR.total/soma.IIR)*100 # Calcula o IIR% de cada
  categoria de presas (linha) e disponibiliza os resultados
  # em um vetor.
  tabela.alimentacao <- data.frame(dados2,IIR.perc) # Cria um objeto do tipo
  data.frame, com as informações da tabela DADOS2 e uma nova
  # coluna com o IIR% para cada categoria de presas. Essa tabela contém
```

```
todos os dados: aqueles apresentados no objeto de
# entrada e os calculados pela função e inseridos em novas colunas.
write.table(tabela.alimentacao, file = "Tabela Alimentação.txt", sep = "
", dec = ".", row.names = TRUE, col.names = TRUE) # Salva a
# tabela.alimentacao em formato .txt no diretório de trabalho, com o nome
Tabela Alimentação, com nomes das colunas e linhas e
# separação por espaços.
tabela.cluster <- data.frame(tabela.alimentacao$IIR.perc) # Para a
confeção do cluster, cria um objeto do tipo data.frame apenas
# com os valores do IIR%.
dist <- vegdist(tabela.cluster, x) # Cria o objeto dist, um matriz de
distância com os dados fornecidos em tabela.cluster. O argumento X
# é o índice de distância utilizado. O default da função é a distância
euclidiana. Outro índice pode ser utilizado desde que sejam
# respeitadas suas premissas.
cluster<-hclust(dist,y) # Cria o cluster para o objeto dist. O argumento y
é o método de agrupamento. O default da função é o método aglomerativo
UPGMA.
# Outros métodos podem ser utilizados, de acordo com os objetivos e o tipo
de matriz de distância.
# Abre o dispositivo jpg e pede para salvar o gráfico a ser gerado no
diretório de trabalho, com o título de dendrograma1, em pixels, com tamanho
e
# definições já estabelecidos.
# Plota o dendrograma a partir dos dados de cluster, com o título de
Dendrograma preliminar. As legendas de grupos são colocadas alinhadas na
# posição -0.1 e pede-se que a validade dos dados seja conferida (embora
check=TRUE seja o default).
# Insere no dendrograma uma linha de corte e retângulos nos grupos
formados, na cor vermelha, na altura de 40 (usualmente, considera-se a
formação
# de grupos com 60% de similaridade para trabalhos de alimentação de
peixes).
# Fecha o dispositivo jpg.
jpeg(filename = "Dendrograma1.jpg", width= 480, height = 480, bg=
"transparent",units="px", pointsize=12, quality=100)
dendrograma.preliminar<-plot(cluster, main="Dendrograma preliminar",
xlab="presas/nível da variável categórica", ylab='distância', hang= -0.1,
check= TRUE)
dendrograma.preliminar<- rect.hclust(cluster, h=40, border = "red")
dev.off()
shapiro <- tapply (dados$IIR.total, INDEX=dados$Categoria,
FUN=shapiro.test) # Aplica o teste de Shapiro para verificar a normalidade
dos dados do IIR total
# para cada nível da variável categórica.
res.anova <- aov(IIR.total~Categoria, data=dados) # Aplica o teste de
ANOVA para verificar se há diferença na composição da dieta entre os
diferentes níveis da
# variável categórica. O resultado é gerado mesmo se as premissas de
normalidade não for verificadas.
```

```
summary(res.anova) # Gera o sumário do teste de ANOVA.  
resultado <- list(tabela.alimentacao, dist, cluster, shapiro,  
summary(res.anova)) # Cria um objeto do tipo lista com os seguintes objetos:  
tabela.alimentacao (todos os dados),  
# dist, cluster, shapiro e summary (res.anova).  
return(resultado) # Retorna o objeto resultado.  
} # Encerra a função.
```

[help_da_funcao_cluster.iir.txt](#)

[script_funcao_cluster.iir.r](#)

[dieta.csv](#)

From:

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

Permanent link:

http://ecor.ib.usp.br/doku.php?id=05_curso_antigo:r2015:alunos:trabalho_final:mpucci:start

Last update: **2020/08/12 09:04**

