



Carolina Freitas Schlosser

Mestranda em Ecologia pela Universidade de São Paulo, sob orientação dos professores Marcio Martins e Leandro Tambosi. Meu projeto de pesquisa tem como objetivo investigar o efeito da cobertura vegetal sobre a comunidade de anfíbios em uma região de cerrado.

[exerciciospreparatorios_1_a_3.r.txt](#)

[exercicios5.r](#)

[exercicio.8.txt.r](#)

Trabalho final:

PLANO A

Essa função tem como objetivo geral, através de um mapa e um `data.frame`, informar sobre a abundância de uma espécie e os locais de registro dela, que, a depender da espécie, também podem ser lidos como locais de potencial distribuição (já que outrora ela já foi vista e registrada ali)

A função tem como entrada um argumento de `data.frame` e um argumento de espécie, e retorna ao usuário:

- 1) os sites (armadilhas de pitfall) onde uma determinada espécie animal foi registrada;
- 2) a abundância da espécie registrada em cada site (pitfall) em que ela ocorre;
- 3) a abundância total da espécie (somatória das abundâncias dos sites/pitfalls);
- 4) a abundância relativa da espécie nos sites em que ela foi encontrada;
- 5) a abundância relativa da espécie na comunidade de espécies presentes no `data.frame`.

Para tanto, ao fornecer para a função o nome da espécie, ela irá retornar um mapa e um `data.frame`.

Argumentos de entrada:

função (dados, espécie)

a) dados - `data.frame` com as seguintes colunas:

Espécie, site (nome do ponto de pitfall), latitude, longitude, abundância, fitofisionomia (do site). Neste `data.frame`, cada linha corresponde a presença de uma espécie em um ponto de pitfall.

Exemplo de `data.frame`:
espécie Site (local) Latitude Longitude Abundância Fitofisionomia X CS1 x y
10 Campo Sujo X CS2 xa ya 4 Campo Sujo Y SS1 xb xb 1 Cerrado Stritu Sensu Z SS2 xc xc 3 Cerrado
Stritu Sensu Z CS1 x y 0 Campo Sujo

b) espécie

Verificando os parâmetros:

O objeto dados é um data.frame? Se não for, retorna uma mensagem de erro.

O objeto espécie é uma coluna do data.frame? Se não, retorna uma mensagem de erro.

O argumento fornecido ao parâmetro “espécie” corresponde a pelo menos uma linha do data.frame (a espécie buscada está no data.frame)? Se não, retorna uma mensagem de erro.

Pseudo-código:

Cria objeto dados, dando nome ao dataframe. Comandos de instalação dos pacotes. Comandos para carregar os pacotes.

Comando para calcular a abundância total de todas as espécies em todos os sites. Comando para calcular a abundância total de cada espécie em cada site.

Abre um loop, no qual:

1) Calcula a abundância relativa de cada espécie em relação a toda comunidade relacionada no data.frame. 2) Calcula a abundância relativa de cada espécie para cada local em que ela ocorra.

Cria objeto (mapa ou gráfico) no qual serão plotados círculos nos sites (coordenadas xy) de ocorrência da espécie chamada.

Puxa todas as linhas do data.frame (pitfalls) em que a espécie chamada está presente. Todas as coordenadas correspondentes serão plotadas no mapa.

Cada valor de abundância absoluta será atribuído um tamanho de círculo.

Os círculos serão então plotados no mapa, nas coordenadas correspondentes aos sites onde a espécie ocorre. Cada site terá dois círculos: Um correspondente a abundância total de organismos registrados no site e outro correspondente a abundância da espécie naquele site.

Saída:

1) Um mapa, no qual estarão plotados 2 os círculos nos sites de registro (pontos de pitfall) da espécie: 1 círculo cujo tamanho é proporcional à abundância daquela espécie naquele local, e outro círculo cujo tamanho é proporcional à abundância total de todas as espécies registradas naquele local. Com esses dois círculos, o usuário pode ter uma informação visual sobre a quantidade de indivíduos registrados naquele local e a abundância relativa da espécie de interesse naquele local.

2) Data.frame com: Abundância total de todas as espécies em todos os sites; Abundância total da espécie; Abundância relativa da espécie para toda a comunidade (relacionando a abundância total da espécie à abundância total de indivíduos de todas as espécies presentes no data.frame); Sites de ocorrência da espécie (site names); Abundância total de todas as espécies nos sites em que a espécie chamada ocorra; Abundância da espécie em cada site; Fitofisionomia correspondente a esses sites; Abundância relativa da espécie nos sites em que ela foi amostrada (essa informação também será dada visualmente, no mapa).

Observações/Contextualização:

Tal função deverá ser muito útil para fornecer ao usuário essa informação visual sobre os locais de ocorrência de uma espécie e a abundância registrada, conquanto que o `data.frame` tenha um formato compatível com aquele lido pela função.

Imagine que interessante para uma pesquisadora (or) ou até mesmo uma pessoa leiga, ao fazer uma trilha, e tendo disponível esse `data.frame`, poderá saber em que locais uma espécie de interesse já foi vista, e onde ela foi vista com maior frequência (maior abundância registrada). E ainda tem um mapa, para ajudar na visualização da informação! 😊

Preciso estudar os pacotes de mapas, caso eu queira plotar os círculos em mapas, o que acredito ser o mais adequado e bonito. Caso isso não seja possível, plotarei em um gráfico, cujos eixos representam latitude e longitude, de modo que fique com a aparência de um mapa.

PLANO B

A função tem como entrada um argumento de `data.frame` e um argumento de espécie (no qual o usuário deve escrever o nome da espécie que desejar), e tem como objetivo retornar ao usuário (saída) informações sobre uma possível especialização/preferência da espécie em uma (ou mais de uma) fitofisionomia(s). O usuário irá inserir um `data.frame` no qual hajam, pelo menos, 3 colunas: Espécie, fitofisionomia e abundância. Por indexação dos nomes dessas colunas, será criado um novo `data.frame`, composto por essas 3 colunas. (Essa criação de um novo `data.frame` dá liberdade para que o usuário utilize a função, ainda que seu `data.frame` original tenha colunas com variáveis que aqui não serão utilizadas).

Criado um novo `data.frame`, a função fará um teste de significância, que avaliará se, dada a abundância total da espécie, a proporção dessa abundância que ocorre em determinada(s) fitofisionomia(s) (aquela(s) com maior proporção desta abundância total) poderia ser gerada ao acaso.

Em outras palavras, a função irá responder se podemos afirmar que uma espécie é especialista de habitat (sua ocorrência em maior quantidade naquela fitofisionomia não é devido ao acaso), ou se a maior ocorrência em um determinado habitat/fitofisionomia não é nada além do que poderíamos esperar pelo acaso.

Argumentos de entrada:

função (dados, espécie)

a) dados - `data.frame` com, pelo menos, as seguintes colunas: espécie, fitofisionomia, abundância

Verificando os parâmetros:

O objeto dados é um `data.frame`? Se não for, retorna uma mensagem de erro.

O objeto espécie é uma coluna do `data.frame`? Se não, retorna uma mensagem de erro.

O nome fornecido ao argumento "espécie" (nome da espécie) corresponde a pelo menos uma linha do `data.frame` (a espécie buscada está no `data.frame`)? Se não, retorna uma mensagem de erro.

Pseudo-código:

Comandos de instalação dos pacotes. Comandos para carregar os pacotes.

Comando para criar um novo data.frame, por indexação das colunas “espécie”, “fitofisionomia” e “abundância” do data.frame fornecido à função.

Comando para calcular a abundância total de cada uma das espécies em todas as fitofisionomias.
Comando para calcular a abundância de cada espécie em cada fitofisionomia.

Abre um loop, no qual são geradas muitas simulações de distribuição de valores de abundâncias das espécies para cada fitofisionomia. (Com isso, será gerado um cenário nulo de distribuição de abundâncias por fitofisionomia para cada espécie).

Para espécie escolhida (de acordo com o argumento da função), a distribuição de abundâncias observada será então comparada com aquela gerada pelo cenário nulo, por meio de um teste de significância.

Comando para geração de histogramas, com distribuição dos valores de ‘proporção da abundância encontrada na fitofisionomia com maior abundância’, gerados pelo cenário nulo, e uma indicação (um traço, por exemplo), do valor de proporção encontrado.

Caso haja proporções iguais em duas fitofisionomias diferentes, serão gerados dois histogramas

Saída:

1) Dataframe, contendo: Nome da espécie, fitofisionomias em que ela ocorre pelo menos uma vez, abundância em cada fitofisionomia, fitofisionomia(s) em que ela ocorre com maior frequência (fitofisionomia(s) preferencial(is)), e uma resposta à pergunta “Ela é especialista desta(s) fitofisionomia(s)?” (Que poderia ser escrita também da seguinte forma: Essa distribuição de abundâncias é devido a alguma preferência da espécie – Ou poderia ser simplesmente gerada ao acaso?) , que deverá ser “SIM” ou “NÃO”, a depender do resultado de teste de significância.

2) Histograma(s) resultante(s) do teste de significância.

Observações/Contextualização:

Tal função deve ser muito útil para automatizar, de forma rápida, uma informação sobre uma espécie, com base num data.frame fornecido: Se ela tem preferência por uma fitofisionomia, e com que grau de certeza isso pode afirmado.

Olá Carolina. As suas duas funções estão boas com um nível de desafio adequado. Acho que tudo ficaria mais interessante se combinássemos as duas propostas. Porque não criar um mapa de distribuição de abundância e fazer um teste de chi-quadrado. Pensei assim, o usuário da função entra com os mesmos dados, só que você plota dois gráficos o do esperado e o do obtido, assim a pessoa tem uma visualização clara do que poderia ser esperado ao acaso e do que poderia ela observou lado-a-lado ou sobreposto. Tudo que você precisa fazer a mais é acessar a informação de distribuição esperada e obtida na tabela de chi-quadrado e plotar cada uma separadamente.

Muito legal! Qualquer dúvida, manda um email ou me chama na 243.

— *Bruno Travassos* 2019/06/14 13:43

Código da função

```
fun.carol<-function(dat, sp, n=1000)
{
  #dat é data.frame?
  if(class(dat)!="data.frame")
  {
    stop("dat não é da classe data.frame")
  }
  #A especie escolhida consta no data.frame fornecido?
  if(sum(levels(dados$especie)==sp)==0)
  {
    stop("especie não consta no data.frame fornecido")
  }

  #Selecionando dados da espécie
  dat.sp=dat[dat$especie == sp,]
  dat.sp
  #fitofisionomias:
  fitofis<-dat$fitofisionomia
  fitofis
  levels(fitofis)
  #abundância total de cada espécie
  abund.sp<- tapply(dat$abundancia,dat$especie, sum)
  abund.sp
  #valores de abundancias da especie desejada:
  abs.sp<-tapply(dat.sp$abundancia, dat.sp$fitofisionomia,sum)
  abs.sp
  #abundancia total da especie desejada:
  abt.sp<-sum(abs.sp)
  abt.sp
  #quantidade de fitofisionomias
  as.factor(dat$fitofisionomia)
  qtd.fit<-length(levels(dat$fitofisionomia))
  qtd.fit
  #quantidade de fitofisionomias -1, para gerar cenário nulo
  qtd.fit.menos.um<-qtd.fit-1
  qtd.fit.menos.um
  #gero cenarios nulos num for rodando 1000 vezes esse cenário nulo
  resulta=matrix(NA,n,qtd.fit)
  for(i in 1:n)
  {
    resulta[i,]<-
diff(sort(c(0,round(runif(qtd.fit.menos.um,0,abt.sp),0),abt.sp)))
  }
  resulta
  #Quantas vezes, nesse cenario nulo eu ter valores
  #maiores ou iguais ao observado nas fitofisionomias?
  #faz sentido procurar valores maiores ou iguais????
```

```
resulta2<-rep(NA,length=qtd.fit)
for(d in 1:qtd.fit)
{
  resulta2[d]<-sum(resulta[,d]>=abs.sp[d])/n
}
resulta2
#Valores esperados são a média dos valores gerados
#pela distribuição aleatória
dist.esperada<- (round(apply(resulta,2,sum)/n))
dist.esperada
dist.observada<-tapply(dat.sp$abundancia, dat.sp$fitofisionomia, sum)
dist.observada<-as.numeric(dist.observada)
#Essas diferenças entre esperado e observado dizem algo sobre a
preferência da espécie? por alguma fitofisionomia?
dist.esperada
dist.observada
#Fazer gráfico em pizza desses valores esperados e observados
graf.esperado<-pie(dist.esperada, main="distribuição esperada",
labels=levels(fitofis))
graf.observado<-pie(dist.observada, main="distribuição observada",
labels=levels(fitofis))
#nomeando as fitofisionomias nas distribuição esperada:
dist.esperada<-list("distribuição esperada"=dist.esperada)
a= levels(fitofis)
dist.esp=data.frame(Fitofisionomia=a,dist.esperada)
dist.obs=data.frame(Fitofisionomia=a,dist.observada)
resulta3=data.frame(Fitofisionomia=a,P.valor=resulta2)
return(list("distribuição observada"=dist.obs,"distribuição esperada"=
dist.esp, "p-valores associados a n simulações" = resulta3 ))
}
```

Página de ajuda

fun.carol

Distribuição esperada (aleatória) e distribuição observada dos indivíduos da espécie entre as fitofisionomias.

Description:

Produz valores e gráficos da distribuição esperada em um cenário nulo e da distribuição observada. Tais distribuições são comparadas, gerando p-valores associados.

Usage:

```
fun.carol(dat,sp,n)
```

Arguments:

dat: Data.frame com os dados a serem analisados.

sp: Espécie

n: número de simulações para geração de cenário nulo

Details:

A função trabalha extrai, do data.frame fornecido, a distribuição dos indivíduos de uma espécie escolhida dentre as fitofisionomias contidas no data.frame.

É gerado um cenário de distribuições aleatórias (distribuição esperada) desses indivíduos, com n simulações aleatórias. O valor de n é escolhido pelo usuário, e tem 1000 como valor default.

Tais informações (distribuição esperada e distribuição observada) são transpostas em dois gráficos, facilitando sua visualização.

São gerados p-valores associados à distribuições observadas .

Os p-valores associados informam ao usuário a probabilidade de que a distribuição observada seja devida ao acaso, informando, portanto, sobre possíveis preferências de fitofisionomia da espécie em questão.

Value:

distribuição esperada.

distribuição observada.

p-valores associados à distribuições observadas .

um gráfico em pizza da distribuição esperada.

um gráfico em pizza da distribuição observada.

Author:

Carolina Freitas Schlosser

Example:

```
#especie<-rep(c("sapo","cavalo","boi"),each=10)

#abundancia<-
rep(c(1,0,2,2,1,3,0,4,3,1,5,4,5,7,8,6,8,9,9,4,8,10,5,9),length=30)

#fitofisionomia<-sample(rep(c("campo aberto","campo
sujo","cerrado","cerradao"),length=30))

#dado.fit<-data.frame(especie,abundancia,fitofisionomia)

#fun.carol(dado.fit, "cavalo", 300)
```

Arquivo da função [funcaocarolina.r](#)

From:

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

Permanent link:

http://ecor.ib.usp.br/doku.php?id=05_curso_antigo:r2019:alunos:trabalho_final:carolinafsch:start



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