CAROLINA TORRES GUTIERREZ

Bióloga, formada na Pontificia Universidad Javeriana / Bogotá, Colômbia. Estudante de Doutorado em Saúde Pública, Faculdade de Saúde Pública. USP. Linha de Pesquisa: Mosquitos de importância médica (Diptera: Culicidae).

Meus Exercícios

exec

Proposta Trabalho Final

A proposta se baseia em importar dados e efetuar estatística descritiva e testes estatísticos simples (primeiros passos para evidenciar associações entre variáveis), utilizando um conjunto de dados de coleta de mosquitos da família Culicidae (hematófagos) provenientes de diferentes localidades da Colômbia. Os dados incluem: 1) Classificação taxonômica dos indivíduos coletados (Gêneros/espécies), 2) Altitude da localidade (desde os 400m até aprox. 1.400m) 3) Tipo de criadouro (classe: artificial, natural) 4) Categoria da localidade (silvestre, rural, urbana)

A análise dos dados fornecerá a freqüência dos gêneros e espécies por criadouro e a distribuição dos táxons no interior de uma faixa altitudinal, além da relação entre as variáveis de criadouros e categoria das localidades. Será construída uma página de ajuda para abordar análises básicas com este tipo de insetos.

PLANO B

Analise da Proposta inicial, reduzindo o número de variáveis.

Comentários

Daniel (Musgo): Tenho algumas sugestões que talvez sejam interessantes: Você poderia incluir argumentos na sua função para possibilitar que o usuário escolhesse o tipo de teste que gostaria de fazer. Outra coisa interessante seria investigar a distribuição de frequências dos táxons nos diferentes pontos de amostragem sob uma ótica de competição entre espécies. Você poderia criar um cenário nulo de distribuição espacial dos insetos e contrastar com os dados coletados, talvez em associação com gráficos explicativos. Talvez com isso a sua função poderia ser aplicada para outros tipos de dados também, não somente para os culicideos.

##############################

Acho as dicas do Daniel pertinentes e interessantes, entretanto a principio fique só com a parte básica da proposta. Uma análise mais exploratória de correlação entre a abundância (e riqueza) dos animais (não precisa ser culicideos) em relação a alguma caracterísitca do habitat (no seu caso tipo de criadouro, categoria da localidade ou alturas) pode ser algo mais geral. Sugiro que seja uma função para uma análise exploratória da ocorrencia de espécies (abundancia e riqueza) em relação a algumas carcterísticas do habitat ou localidade de coleta... — *Alexandre Adalardo de Oliveira 2012/04/04 09:45*

TRABALHO FINAL

```
 package:None                R Documentation

Description:

Este é um exercício prático para explorar dados e estimar riqueza de
espécies dos dados de uma base de dados que consiste em colunas com nomes de
gênero, subgênero e espécie, e colunas com informação referente
a localidade de coleta (urbana,rural),altitude (metros sobre nível do mar) e
lugar de coleta para imaturos de culicídeos.
A base de dados que foi utilizada tem as dimensoes de 853 linhas vs. 6
colunas (registros de espécies e pontos onde foram coletadas). As colunas:
gênero, subgênero, espécie, altitude,
criadouro e localidade. A analise está subdividida em varios passos(steps)
que exploram os dados e comparam a riqueza de espécies por localidade assim
como a riqueza de espécies contida
em cada um dos criadouros amostrados. De forma complementar, sao dadas
funçoes simples para comparar a relação entre variáveis como: altitude vs.
espécies; localidade vs. espécie e criadouro vs. localidade.
Para esclarecer a análise feita, são anexados os arquivos com todos os
códigos utilizados (parte exploratória ), as funções para a análise (em
passos ou “steps”) e a base de dados original.

Usage:

     Anexploratoria(x=dados)

Arguments:
x= database em (.csv)

Details:

     A estimativa de riqueza utilizada na função corresponde com a fórmula:
Riqueza= Número de espécies por unidade amostral/Raíz quadrada do total de
```

indivíduos por unidade amostral. Sabe-se de outras fórmulas de riqueza propostas em textos de Ecologia,
no entanto, para o objetivo de fazer uma análise exploratória básica, a fórmula indicada se considerou adequada.

As unidades amostrais consideradas em esta função correspondem com as categorias diferentes para “Localidade” que são: “urbana” e “rural”.
Para cada passo ou STEP da função foi incluída uma observação entre símbolos “#” que dá a explanação devida.

Nome do database analisado= dadosculex (veja arquivo anexo)

O database têm as seguintes colunas:

dadosculex$Gênero: informa o gênero do indivíduo (para o caso, a maioria foram larvas de mosquito do gênero Culex)
dadosculex$subgênero: informa o subgênero do individuo (para alguns grupos esta categoria taxonômica é de importância)
dadosculex$Espécie: informa a espécie (nome composto por genero+subgenero+epiteto especif.) do indivíduo coletado
dadosculex$Localidade: informa a localidade (seja rural ou urbana) de coleta
dadosculex$Altitude: informa a altitude onde foi coletado cada indivíduo (exemplo: 1900 metros acima do nível do mar)
dadosculex$Criadouro: informa o tipo de criadouro onde se coletou cada individuo (lembrar que os mosquitos apresentam fases imaturas aquáticas que geralmente sao coleçoes contidas em ambientes silvestres ou antrópicos)

Value:

A função retorna uma lista de resultados por cada step ou passo da análise:
1. STEP1= Tabelas de relação entre variáveis: gênero vs. localidade; espécie vs. criadouro; gênero vs. altitude; subgênero vs. criadouro; e finalmente uma tabela de proporçoes que ajuda a observar a frequencia das espécies em cada criadouro.

2. STEP2= Criam-se objetos diferentes para cada uma das espécies presentes na matriz de dados e depois é calculada a abundância total.

3. STEP3= Calculam-se valores de riqueza de espécies para as duas categorias de "localidade"= urbana vs. rural

4. STEP4= Se faz um sumário das espécies vs. os criadouros vs. número de indivíduos em cada criadouro. Se estima a Riqueza de espécies para cada um dos criadouros na matriz de dados.

5. STEP5= Dados complementares. Ou seja, sao funçoes simples que ajudam a investigar relaçoes entre as diferentes variáveis, e que sao dadas de forma adicional à funçao exploratória.
Os resultados destas analises complementares apresentam gráficos

```
de relaçao entre variáveis.

6. STEP 6= Analise entre Altitude vs. Espécies (apresenta um boxplot)

7. STEP6.1= Analise de Localidade vs. Espécies (apresenta grafico de barras)

8. STEP6.2= Analise de Criadouro vs. Localidade (apresenta grafico de
barras)

9. STEP6.3= Analise de Altitude vs. Criadouro (apresenta um boxplot)

10.STEP6.4= Analise de Alittude vs. Localidade (boxplot)

11. STEP 6.5= Analise de Localidade vs. Altitude (An. variância)

12. STEP 6.6.= Analise de Espécie vs. Altitude (An. variância)


Note:

   Para uma melhor compreensao desta analise exploratória, confira o arquivo
anexo “Analisexploratoria.r”, com todos os códigos utilizados.

Author(s):

Carolina Torres Gutierrez
aniloract@gmail.com

References:
Montes, Joyce. (2005). Fauna de Culicidae da Serra da Cantareira, São Paulo,
Brasil.
Revista de Saúde Pública, 39(4), 578-584. Retrieved April 23, 2012,
from
http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0034-891020050004000
10&lng=en&tlng=pt.
http://dx.doi.org/10.1590/S0034-89102005000400010.


Examples:

  Use o arquivo "dadosculex.csv" e siga os códigos do arquivo
“Analisexploratoria.r”. Depois teste a funçao descrita aqui.
```

CODIGO DA FUNÇÃO

```
dadosculex <-read.table (file="dadosculex.csv", header=TRUE, sep=",")
dadosculex
```

```
head(dadosculex)
summary(dadosculex)
str(dadosculex)
tamanho= dim(dadosculex)
tamanho

### STEP 1 ######COMEÇANDO A EXPLORAR O DATABASE POR CADA COLUNA E FAZENDO
TABELAS ENTRE COLUNAS###


anexploratoria<- function(x, rmNA=TRUE)

{
dadosculex= read.table (file="dadosculex.csv", header=TRUE, sep=",")
dadosculex=x
genero= dadosculex$Gênero
localidade= dadosculex$Localidade
especie= dadosculex$Espécie
criadouro= dadosculex$Criadouro
altitude= dadosculex$Altitude
subgenero=dadosculex$Subgênero
analis1= table (genero, by= localidade)
analis2= table(especie, by= criadouro)
analis3= table (genero, by= altitude)
analis4= table(subgenero, by= criadouro)
analis5= prop.table(table(criadouro))
resultados1= list(analis1, analis2,analis3,analis4,analis5)

### STEP 2 ###### SE CALCULAM VALORES DE ABUNDANCIA ### ######
y1=dadosculex[dadosculex[ ,3] == "Culex (Culex) quinquefasciatus",3]
y2=dadosculex[dadosculex[ ,3] == "Culex (Culex) stenolepis",3]
y3=dadosculex[dadosculex[ ,3] == "Culex (Culex) coronator",3 ]
y4=dadosculex[dadosculex[ ,3] == "Culex (Culex) brevispinusus",3 ]
y5=dadosculex[dadosculex[ ,3] == "Culex (Culex) sp.",3 ]
y6=dadosculex[dadosculex[ ,3] == "Culex (Culex ) acharistus",3 ]
y7=dadosculex[dadosculex[ ,3] == "Culex (Carrollia) secundus",3 ]
y8=dadosculex[dadosculex[ ,3] == "Culex (Carrollia) bihaicolus",3]
y9=dadosculex[dadosculex[ ,3] == "Culex (Carrollia) sp.",3]
y10=dadosculex[dadosculex[ ,3] == "Culex (Microculex) aureus",3 ]
y11=dadosculex[dadosculex[ ,3] == "Culex (Microculex) neglectus",3 ]
y12=dadosculex[dadosculex[ ,3] == "Culex (Microculex) stonei",3 ]
y13=dadosculex[dadosculex[ ,3] == "Culex (Microculex) sp.",3 ]
y14=dadosculex[dadosculex[ ,3] == "Culex (Microculex) davisi",3 ]
y15=dadosculex[dadosculex[ ,3] == "Culex (Microculex) worontzowi",3]
y16=dadosculex[dadosculex[ ,3] == "Aedes sp.",3]
y17=dadosculex[dadosculex[ ,3] == "Culex sp.",3]
Abundancia=
length(c(y1,y2,y3,y4,y5,y6,y7,y8,y9,y10,y11,y12,y13,y14,y15,y16,y17))
total.especies= aggregate(especie, list(especie,localidade), length)
raizdeAbun= sqrt (Abundancia)
localidadeurbana=dadosculex$Localidade=="urbana"
```

```
localidaderural=dadosculex$Localidade=="rural"
### STEP 3 ######SE CALCULAM VALORES DE RIQUEZA #########
###dados de Localidade urbana ###
total.indi.urbanos= subset(especie, localidadeurbana)
total.especies.urbanas=unique(total.indi.urbanos)
totalindividuosurbanos= length(total.indi.urbanos)
totalespurbanas= length(total.especies.urbanas)
riqueza.urbana= totalespurbanas/raizdeAbun

###dados de Localidade rural ###
total.indiv.rurais= subset(especie, localidaderural)
totalindividuosrurais= length(total.indiv.rurais)
total.especies.rurais= unique(total.indiv.rurais)
totalesprurais= length(total.especies.rurais)
riqueza.rural= totalesprurais/raizdeAbun
resultadosurbanos=
list(totalindividuosurbanos,totalespurbanas,riqueza.urbana)
resultadosrurais= list(totalindividuosrurais,totalesprurais,riqueza.rural)
riqueza= list(riqueza.urbana,riqueza.rural)

resultados2= list(Abundancia,total.especies)
resultados3= list(resultadosurbanos,resultadosrurais)
resultados4 = list(resultados1,resultados2,resultados3)

############STEP 4, ESTIMATIVA DE RIQUEZA POR CADA TIPO DE CRIADOURO
(LEMBRAR QUE A VARIAVEL CRIADOURO FAZ PARTE DOS LUGARES DE COLETA PARA ESTA
DATABASE############

total.criadouros= aggregate(especie, list(especie, criadouro), length)
resultados4.1 = total.criadouros
resultados4.2= riqueza

armadilha=dadosculex$Criadouro=="Armadilha ovos"
############
total.armadilha= subset(especie, armadilha)
ttindivarmadi= length(total.armadilha)
raizdettarmadilha= sqrt(ttindivarmadi)
numberspeciesa= 3
riqueza.armadilha= numberspeciesa/raizdettarmadilha
############
bromelia= dadosculex$Criadouro=="Bromelia"
total.bromelia= subset(especie, bromelia)
ttindivbromelia<- length(total.bromelia)
raizdettbrom= sqrt(ttindivbromelia)
numberspeciesb= 10
riqueza.bromelia= numberspeciesb/raizdettbrom
############
descartavel= dadosculex$Criadouro=="descartável"
total.descartavel= subset(especie,descartavel)
```

```
ttindivdescartav=length(total.descartavel)
raizdettdescart= sqrt(ttindivdescartav)
numberspeciesc= 3
riqueza.descart= numberspeciesc/raizdettdescart
############
pantano=dadosculex$Criadouro=="Pântano"
total.pantano= subset(especie, pantano)
ttindivpantan= length(total.pantano)
raizdettpantan= sqrt(ttindivpantan)
numberspeciesd= 1
riqueza.pantan= numberspeciesd/raizdettpantan
############
tanque= dadosculex$Criadouro=="Tanque cimento"
total.tanque= subset(especie,tanque)
ttindivtanq= length(total.tanque)
raizdettanq= sqrt(ttindivtanq)
numberspeciese= 4
riqueza.tanq= numberspeciese/raizdettanq
############
balde= dadosculex$Criadouro=="balde"
total.balde= subset(especie,balde)
ttindivbalde= length(total.balde)
raizdettbalde<- sqrt(ttindivbalde)
raizdettbalde
numberspeciesf= 1
riqueza.balde= numberspeciesf/raizdettbalde
############
cocho=dadosculex$Criadouro=="Cocho"
total.cocho= subset(especie, cocho)
ttindivco= length(total.cocho)
raizdettco= sqrt(ttindivco)
numberspeciesg= 1
riqueza.cocho= numberspeciesg/raizdettco
############
planta=dadosculex$Criadouro=="Folha planta"
total.planta= subset(especie, planta)
ttindivplan= length(total.planta)
raizdettpla= sqrt(ttindivplan)
numberspeciesh= 1
riqueza.planta= numberspeciesh/raizdettpla
############
pegada=  dadosculex$Criadouro=="Pegada gado"
total.pegada= subset(especie,pegada)
ttindivpeg= length(total.pegada)
raizdettpeg= sqrt(ttindivpeg)
numberspeciesi= 2
riqueza.pegada= numberspeciesi/raizdettpeg
############
bambu= dadosculex$Criadouro=="Bambu"
total.bambu= subset(especie, bambu)
ttindivbamb= length(total.bambu)
```

```
raizdettbamb= sqrt(ttindivbamb)
numberspeciesj= 4
riqueza.bambu= numberspeciesj/raizdettbamb
############
solo= dadosculex$Criadouro=="depressão no solo"
total.solo= subset(especie, solo)
ttindivsolo= length(total.solo)
raizdettsolo= sqrt(ttindivsolo)
numberspeciesk= 3
riqueza.solo<- numberspeciesk/raizdettsolo
############
heliconia= dadosculex$Criadouro=="Heliconia"
total.helicon= subset(especie, heliconia)
ttindivhelicon= length(total.helicon)
raizdetthelicon= sqrt(ttindivhelicon)
numberspeciesl= 1
riqueza.helicon= numberspeciesl/raizdetthelicon
############
pneu= dadosculex$Criadouro=="Pneu"
total.pneu= subset(especie, pneu)
ttindivpneu= length(total.pneu)
raizdettpneu= sqrt(ttindivpneu)
numberspeciesm= 5
riqueza.pneu= numberspeciesm/raizdettpneu
############
tipocriadouro= list ("armadilha", "bromelia", "descartavel","pantano",
"tanque", "balde", "cocho", "planta", "pegada", "bambu", "solo", "helicon",
"pneu")
valoresriqueza= list(riqueza.armadilha, riqueza.bromelia, riqueza.descart,
riqueza.pantan, riqueza.tanq, riqueza.balde, riqueza.cocho, riqueza.planta,
riqueza.pegada, riqueza.bambu, riqueza.solo, riqueza.helicon, riqueza.pneu)
riquezaporcriadouro= list (tipocriadouro, by= valoresriqueza)
resultados.totais= list(resultados4, resultados4.1, resultados4.2)
resultados5 = riquezaporcriadouro
resultados7= list(resultados.totais, resultados5)
return (resultados7)
}

anexploratoria(dadosculex)


###### STEP 5 ###TENTANDO ACHAR CORRELAÇAO ENTRE AS VARIAVEIS
#####################
############A PARTIR DE ESTE PASSO SERAO ANALISADAS AS VARIAVEIS
SEPARADAMENTE ########################

relacaovariaveis<- function(x, rmNA=TRUE)

{
```

```
dadosculex= read.table (file="dadosculex.csv", header=TRUE, sep=",")
dadosculex=x
especie= dadosculex$Espécie
criadouro= dadosculex$Criadouro
altitude= dadosculex$Altitude
localidade= dadosculex$Localidade
table1= table (especie, altitude)
proporcoes= prop.table(table(especie, localidade))
table3= table(especie)
table4= prop.table(table(especie))
table5= prop.table(table(especie,criadouro))

resultados8= list(table1,proporcoes,table3,table4,table5)
return(resultados8)
}

relacaovariaveis(dadosculex)

###### STEP 6 ### INVESTIGANDO SE EXISTE CORRELAÇAO ENTRE AS VARIAVEIS
#####################

###### TESTA-SE RELACAO ENTRE ALTITUDE VS.ESPECIES ######

relacaovariaveis2<- function(x, rmNA=TRUE)

{
dadosculex= read.table (file="dadosculex.csv", header=TRUE, sep=",")
dadosculex=x
especie= dadosculex$Espécie
altitude= dadosculex$Altitude
graf1= boxplot(altitude ~ especie)
resultado9= graf1
return(resultado9)
}
relacaovariaveis2(dadosculex)


###### STEP 6.1 TESTA-SE RELACAO ENTRE LOCALIDADE VS.ESPECIES ######

Relacao3<- function(x, rmNA=TRUE)

{
dadosculex= read.table (file="dadosculex.csv", header=TRUE, sep=",")
dadosculex=x
especie= dadosculex$Espécie
localidade= dadosculex$Localidade
graf2= barplot(table(especie, localidade), beside = T, legend = T, ylim =
c(0, 400), col = terrain.colors(17))
resultado10= graf2
return(resultado10)
}
```

```
Relacao3(dadosculex)

###### STEP 6.2 TESTA-SE RELACAO ENTRE CRIADOURO VS. LOCALIDADE ######

Relacao4<- function(x, rmNA=TRUE)

{
dadosculex= read.table (file="dadosculex.csv", header=TRUE, sep=",")
dadosculex=x
especie= dadosculex$Espécie
localidade= dadosculex$Localidade
criadouro= dadosculex$Criadouro
graf3= barplot(table(criadouro,localidade), beside = T, legend = T, ylim =
c(0, 400), col = terrain.colors(13))
resultado11= graf3
return(resultado11)
}
Relacao4(dadosculex)

###### STEP 6.3 TESTA-SE RELACAO ENTRE ALTITUDE VS. CRIADOURO  ######

Relacao5<- function(x, rmNA=TRUE)

{
dadosculex= read.table (file="dadosculex.csv", header=TRUE, sep=",")
dadosculex=x
altitude= dadosculex$Altitude
criadouro= dadosculex$Criadouro
graf4= boxplot(altitude ~ criadouro)
resultado12= graf4
return(resultado12)
}
Relacao5(dadosculex)

###### STEP 6.4 TESTA-SE RELACAO ENTRE ALTITUDE VS. LOCALIDADE  ######

Relacao6<- function(x, rmNA=TRUE)

{
dadosculex= read.table (file="dadosculex.csv", header=TRUE, sep=",")
dadosculex=x
altitude= dadosculex$Altitude
localidade= dadosculex$Localidade
graf5= boxplot(altitude ~ localidade)
resultado13= graf5
return(resultado13)
}
Relacao6(dadosculex)
```

```
###### STEP 6.5 ANALISE DE VARIANCIA ENTRE LOCALIDADE E ALTITUDE  ######

Analise1<- function(x, rmNA=TRUE)

{
dadosculex= read.table (file="dadosculex.csv", header=TRUE, sep=",")
dadosculex=x
altitude= dadosculex$Altitude
localidade= dadosculex$Localidade
aov1 <- aov(altitude ~ localidade)
sumario= summary(aov1)
resultado14= sumario
return(resultado14)
}
Analise1(dadosculex)

###### STEP 6.6 ANALISE DE VARIANCIA ENTRE ESPECIE VS. ALTITUDE  ######

Analise2<- function(x, rmNA=TRUE)

{
dadosculex= read.table (file="dadosculex.csv", header=TRUE, sep=",")
dadosculex=x
altitude= dadosculex$Altitude
especie= dadosculex$Espécie
aov2 <- aov(altitude ~ especie)
sumario= summary(aov2)
resultado15= sumario
return(resultado15)
}
Analise2(dadosculex)

########################FINAL DA ANALISE ########################
```

---

dadosculex.csv

analisexploratoria.r

codigo_funcao.r

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

Permanent link:
http://ecor.ib.usp.br/doku.php?id=05_curso_antigo:r2013:alunos:trabalho_final:ctorres:start

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